A LINGUISTIC ENVIRONMENT FOR COMPARATIVE ROMANCE SYNTAX[1]

MAURICE GROSS
University of Paris

Romance languages traditionally constitute a privileged field for diachronic studies. There is a fundamental reason for this: the unique size and variety of available documents. The recent development of linguistics has had little effect on the field, presumably because acceptability judgments, the main basis of formalization, cannot be carried out on historical data, but it is precisely the activity of systematic manipulation of strings submitted to linguistic judgment that has proved revolutionary.

Nonetheless, we think that a well motivated activity of formalization could modify the field drastically, especially in the domain of syntax, if precise and extensive descriptions were available.

We propose here a general program of synchronic study of Romance languages and we will attempt to justify it through the experience acquired in constructing a lexicon-grammar of French,[2] and in similar enterprises for Italian,[3] Portuguese[4] and Spanish.[5]

We first present the principles of construction of lexicon-grammars, with examples taken from the French system.

1. INTRODUCTION TO THE EXPERIMENTAL PROCEDURES

The descriptions presented here have been made in Z.S. Harris' theoretical framework. We recall its main characteristics (Z.S. Harris 1982).

Linguistic data have to be distinguished among a vast body of observations or comments that can be made about language. Selecting data is both a theoretical and an empirical problem for which traditional grammar is often a starting point. Traditional grammar has constituted a body of theoretical concepts for which a certain consensus exists. This consensus is founded on the (mixed) success met by grammars used in teaching activities.

Structural grammars have sharpened the concepts of traditional grammars in two main directions:

- by requiring that grammars be formal, namely, that rules apply mechanically and not according to the talent of the grammarian/scholar or of the pupil;
- by eliminating unreliable semantic definitions. This step is a necessity, if *reproducibility* in the application of the rules is to be achieved.

Transformational grammar, by introducing the notion of systematic relation between sentences, brings complex mechanisms of interaction of rules into play; as a consequence, a formal activity of model construction becomes important.

The theoretical concepts we will rely on have their origin in this general background. However, we intend to effectively introduce the lexical dimension in syntax, which means that we will seek to provide a large coverage both of the lexicon and of the rules of grammar. In order to do so, we must guarantee some notion of additivity for the data. This preoccupation is rather new, since the concept of accumulation of facts appears to be absent from most current work.

We have adopted Z.S. Harris' level of representation. Although it is quite close to surface forms, it is abstract enough to represent most syntactic phenomena adequately and in a minimal way. It could easily be agreed upon by linguists interested in carrying out methodical descriptions to adopt this formalism, since the formal notions under study have been accepted, explicitly or not, by all linguists. Moreover, an overt recognition of this situation would allow linguists to build reliable grammars in common, and would not prevent further theoretical elaboration by individuals.

For example, we represent simple declarative sentences as

$$N_0\ V\ Compl$$

N_0 is the subject and *Compl* is a variable ranging over the sequence of complements. The status of *Compl* is provisional, since the analysis of complements is far from having received a satisfactory solution.[6] Complements are made explicit, as in the structure

$$N_0\ V\ N_1\ Prep\ N_2$$

where complement phrases are numerically indexed for reference by rules. Parentheses are used when it becomes necessary to mark the boundaries of a phrase, as for example in

$$N_0\ V\ (Na\ of\ Nb)_1\ Prep\ N_2$$

where complement number 1 is composed of a head phrase *Na* and of a noun complement *of Nb*. The symbol *S* represents a sentence. The symbol "=:" is used to specify structures: N_1 =: *Nhum* means that the first complement is headed by a human noun. Relations between sentences are indicated by

468

the "=" sign. Many relations we are dealing with are transformations in the naive sense.

This type of representation has proven to be sufficient in most situations where we are concerned with demonstrations of the *existence* of relations between sentences. We think that many relations are still being overlooked and that many which are under current discussion are ambiguous, that is, not specified well enough. The activity of proof construction that we advocate is independent of theoretical activities such as defining an abstract structure from which two related sentences can be derived, or such as looking for criteria that separate relations at the grammar level or at the lexicon level.

It seems to us that the main reason why reasonably complete formal grammars have not yet been built, although methods for doing it are available, is the inflation of nonoperational concepts in linguistic discussions. Some of these superfluous concepts are formal, and have been discussed in Gross 1978; others are more traditional, but are so deeply anchored in habit through early education that linguists are reluctant to eliminate them (Cf. 1.3).

An important issue in the construction of a grammar is that of homogeneity of data. Theoreticians have taken for granted the validity of intuitive data. On the other hand, socio-linguists have focused on extra-linguistic factors of heterogeneity in language structure. These two approaches reflect the difficulty of agreeing on a balanced level of empirical data. We will describe various procedures that have been adopted for the construction of our lexicon-grammar, and will indicate how these procedures get around some of the difficulties encountered.

1.1 THE LEXICON

A preliminary question is the selection of a working lexicon. Since our first study was on verbs, we had to

select a list of verbs. Commonly available dictionaries contain a wide variety of verbs, which can be highly literary, obsolete or entirely technical. The study was aimed at characterizing sentence forms built on a fixed and relatively large lexicon of verbs, hence, judgments of acceptability were to be performed on a large scale. Consequently, verbs whose meaning or use were unclear to the linguists who carried out the intuitive evaluations could be *a priori* eliminated. Going over dictionary lists of about 12,000 (morphological) verbs resulted in a list of about 7,000 (morphological) verbs from which linguists could with reasonable ease build simple sentences of varied forms that gave rise to replicatable judgments. This initial figure grew progressively to about 10,000, because of the ambiguities of the morphological entries. Consider for example the form *voler*; this verb enters into the following two structures (among others):

N_0 *vole* =: *Cet oiseau vole lentement*
(This bird flies slowly)

=: *Cet avion vole à une altitude de 1.000 m.*
(This plane flies at an altitude of 1,000 m.)

N_0 *vole* N_1 *à* N_1 =: *Max a volé un livre à Jo*
(Max stole a book from Jo)

These two structures cannot be reduced to each other: omitting the two complements in the second structure (if possible at all) does not lead to the first structure, which has different distributional properties, that is, different meaning. However, we do not rely on meaning to separate such ambiguous entries. A given morphological verb will be separated into several entries, when no rules of the grammar can relate its various uses to each other (Cf. the example of *partager*—'to share' in 3.2.4).

Obviously then, the size of the lexicon of verbs depends on the rule content of the grammar used to describe elementary sentences.

Consider again the example *voler*—'to fly' for which we have given two examples. Intuitively, the meanings of these examples are quite close, but there are also intuitive differences between them.

The introduction of nominalization relations with a support verb (Cf. 2.3) provides syntactic arguments for separating the two uses of the verb into two entries:

Cet avion vole à une altitude de 1.000 m.
(This plane flies at an altitude of 1.000 m.)

= *Cet avion est en vol à une altitude de 1.000 m.*
(This plane is in flight at an altitude of 1.000 m.)

The support verb *être en*—'to be in' does not apply to the other example:

**Cet oiseau est en vol (E + lent)*
(This bird is in slow flight)

In the same way, using the support verb *faire*—'to make', we observe the following difference:

Notre avion a volé directement de Lyon à Nice
(Our plane flew non-stop from Lyon to Nice)

= *Notre avion a fait un vol direct de Lyon à Nice*
(Our plane made a non-stop flight from Lyon to Nice)

Les oies volent sur des distances sans s'arrêter
(Geese fly over distances without stopping)

=**Les oies font des vols sur des distances sans s'arrêter*
(Geese make continuous flights over distances)

Other uses of the verb *voler*—'to fly' will have to be distinguished, as for example

Max vole sur Boeing 321
(Max flies a Boeing 321)

In practice, every single verb raises similar questions, and individual solutions have to be found for each verb. There is no shortcut, as we will see, since there are verbs which are close in meaning, but which raise different problems, as regards their separation into several entries.

Describing the lexicon-grammar of a language is then a complex, time-consuming enterprise, involving numerous specialists who work independently of each other. Moreover the descriptions they produce have to be cumulative: new compatible information is constantly being added to the lexicon-grammar. At each step of adjunction, the total construction should remain coherent. Hence, it is excluded that consideration be given to theoretical proposals such as those of generative grammar. The high pace of production of new theories and of their revisions is such that no lexicon-grammar could ever be constructed, if it were necessary to incorporate ephemeral theoretical proposals and to take into account their respective merits.[7] A theoretical modification that puts into question a mass of carefully accumulated data may well occur, and may imply rewriting a whole lexicon-grammar, but such a revision would be carried out only under highly compelling evidence, and presumably with the consensus of the research community. This attitude of careful and controlled revision of data is the rule of the hard Sciences.

1.2 ACCEPTABILITY OF SENTENCES

The proposed method of making acceptability judgments is standard. Judgments made on sequences of words should be binary: sequences are accepted as sentences or not. In the grammar we do not use more than the two values: acceptable and "*" (i.e. unacceptable). In discussions,

we often introduce an interrogation mark, thus allowing four values instead of two. But we exclude these expository values from the formal representation which is kept strictly binary. Interrogation marks are to be considered as informal comments about the formal descriptions (for a further elaboration of the notion of acceptability, cf. Z.S. Harris 1968).

There are various well-known reasons in favour of this position:

- the formal nature of the theory that aims at building a *characteristic* device (whether a formal grammar, an automaton or a characteristic function),
- the hypothesis that a first approximation of a description should not include socio-linguistic data in the grammatical representations. Thus, we make no provision for notions such as idiolect, multiple grammars for a single speaker or separate grammars for separate groups of speakers. Nor do we distinguish between styles, such as literate speech and slang, for example.

The binary hypothesis, which is that of Z.S. Harris since 1952 and of Chomsky 1957, acknowledges that syntactic phenomena will appear as extremely complex, even when made abstract by the elimination of numerous delicate nuances. Hence, introducing additional ill-defined parameters of study in the first stage of the description will only result in blocking further progress. This view has proved to be entirely correct.

We accept the distinction between synchronic and diachronic data, but only from the operational point of view of making acceptability judgments, not for *a priori* theoretical reasons (M. Gross 1975:225-228).

473

For many word sequences, one hesitates on the acceptability value to be assigned. We will indicate various procedures that allow linguists to obtain reproducible decisions. However, these procedures may sometimes fail. In general, a failure to reach a decision about acceptability affects situations that are not well understood linguistically, which involve for example productivity, or the instability of not well established rules. Often, these fuzzy judgments affect sentences of a certain complexity, and this entails that they have a high degree of ambiguity. Since we are unable to obtain data in such areas of the grammar, we consider that we cannot deal with the problem, at least for the time being. Consider an elementary example that bears on the determiner of the subjects of numerous verbs. In the pair of sentences

(1a) *Bob hit Joe*
=(1b) *Joe was hit by Bob*

no acceptability problem arises. Consider now the close pair

(2a) *Bob hit a waiter*
=(2b) *A waiter was hit by Bob*

We feel that the passive form (2b) is dubious when compared to (1b). There seems to be a phenomenon linked to the indefinite nature of the subject: with most verbs, and without context, indefinite subjects are of intermediary acceptability, and this remark is true for active forms as well.[8] Compare the following three sentences in this respect:

Joe is working hard
A waiter is working hard
A waiter is standing behind the counter

Since the phenomenon is not understood, we will not attempt to carry out acceptability judgments in order to separate the types of verbs involved, that is, we will not introduce in a lexicon-grammar a binary property that classifies verbs according to the felt difference. Notice that we do not say that the problem is not worth studying;[9] rather, we consider that the question has not yet reached a stage where reproducible judgments can be made which lead to a formal representation.

Recognizing such a situation has then an important practical effect: when a form such as Passive is studied with a view to its systematic representation, one should avoid testing forms that contain an indefinite subject.

Some indirect procedures used to make acceptability decisions are standard. For example, if the accepted theory generates a dubious form, one may decide to consider it acceptable. However, the matter can be more complex. In the example of the passive forms, theory A could link the application of the Passive rule to the determiner of the direct object. Such a step might be costly, since Passive would no longer apply in a "formal" way, that is, independently of the context of the *NP*'s involved. But under a theory B where Passive is formal, the dubious passive sentences would have to be considered as entirely acceptable, which does not reflect the actual status of the sentences. We may be helped out of this dilemma by the distinction between acceptability and grammaticality, as discussed by Chomsky 1957;[10] it is then possible to construct forms that are unacceptable in a clear-cut way, though they are grammatical.

In this context, one must also pay attention to the degree of abstraction of the original forms (whether deep or basic) from which accepted sentences are derived. The

degree of abstraction can vary between almost acceptable strings of words such as (2b) to sequences of abstract symbols for which the notion of acceptability is irrelevant. In this respect, we favour concrete theories (e.g. 3.1.4).

There are other ways of reaching an acceptability decision that depend on the particular problem under analysis. For example, consider the dubious form (2b) and its cleft transform:

(3b) *It is a waiter that was hit by Bob,* [*not a baker*]

If (2b) is not completely acceptable, we have to account for the fact that (3b), which is derived from (2b) has normal acceptability, that is, is as acceptable as

(3a) *It is Joe that was hit by Bob*

We then have the two syntactic paradigms or derivations:

(1a) = (2a) = (3a)

(1b) = (2b) = (3b)

They differ only in the determiner of the direct object. This lack of parallelism can only be described by means of a special device for indexing basic forms and their transforms. Such an addition to a theory would be avoided if (2b) were accepted. In this example, the syntactic paradigm forces us into regularizing (2b).

Unclear forms can also be normalized with respect to the lexicon, instead of relative to the grammar, as above.

In studying the distribution of a rule such as Passive over a full lexicon, the acceptability of numerous verbs of varied semantic types with respect to the rule has to be judged. These verbs belong to many different semi-technical domains and to different ways of speaking. Inevitably, the judgment of a given linguist comes up

against certain limitations. The experience gained at the L.A.D.L. about this type of work is the following:

Judgments about a form (e.g. Passive) are performed (more or less independently) by the members of a team of four or five linguists. One requirement is that quick and clear-cut "yes" (acceptable) or "no" answers have to be given. Also, there must be a consensus on the answer. No double-blind experiments or other techniques used by psychologists can be appealed to, for they would slow down the process.

Seeking the judgments of non-specialists of syntax has not been found to be useful. Laypeople tend to introduce stylistic, semantic and situational factors in their evaluations of sentences, parameters which are not directly relevant to syntactic acceptability. In contrast, linguists tend to use independent empirical and theoretical information that orient their manipulations of strings toward explicit decisions of acceptability. Also, linguists are careful to avoid interferences between independent phenomena. As a consequence, they build sentences that have been simplified to the greatest extent for their experiments. Laypeople often react in nonrelevant ways, when faced with what they consider to be artificial forms. For example, they will attempt (explicitly or not) to "balance" a simple sentence by changing tenses or determiners, or by introducing modifiers or adverbs, which may be the type of constructions linguists want to avoid in their design of an acceptability experiment. Carrying out acceptability judgments on syntactic forms needs special training and experience. Needless to say, eliciting acceptability judgments in a poorly known language from a poorly known informant is an entirely different field of linguistics. Results obtained in this way will have to

be improved upon before they can be submitted to deep theoretical analysis. 477

There is an empirical observation of some importance related to this situation. Sharp judgments of acceptability tend to become diluted when the sequences of words under examination are lengthened, while the global syntactic shape is kept identical. Changes occur in both directions: sharply unacceptable strings become only dubious, and clear sentences become cloudy.

There is another important empirical observation that helps linguists to achieve coherent decisions. The use of corpus for checking acceptabilities, that is, attestations, is time consuming. In the case of a phenomenon dependent on multiple conditions, the yield of the examination of large amounts of texts can be quite meager. There is never any prior indication that one or more novels will provide an interesting set of, say, passive forms with indefinite subjects. Moreover, no results about unacceptable strings can emerge from a text. Hence, a corpus-based approach can only be considered as an appendix to the combinatorial technique that consists in varying forms and lexical items, and in testing their acceptability. However, certain traditional studies based on corpus have revealed an interesting feature of texts: forms that have been considered as dubious by linguists are attested. For example, we have studied the distribution of sentential complements with respect to the French lexicon (M. Gross 1975). We then often faced constructions such as

Max autorise Bob à ce qu'il vienne demain
(Max authorizes that Bob comes to-morrow)

Max insiste sur ce que Bob vienne demain
(Max insists on that Bob arrives to-morrow)

that we found of dubious acceptability; the second example is practically unacceptable. Nonetheless, theoretical reasons and analogies lead us to accept them, that is, to mark them "+" (acceptable), in the lexicon-grammar. Interestingly, in a study of texts on a similar problem (the occurrence of certain subjunctive constructions) Nördahl 1969 found attestations of what we considered as very dubious forms.

The two preceding empirical observations converge. We have to recognize that acceptability judgments systematically underestimate the set of acceptable forms. We draw from these observations a working principle: When a linguist faces a difficult binary decision, namely, when (s)he evaluates a dubious string, (s)he should consider it as acceptable.[11]

Our effort of abstraction thus bears on the data, and not on the theory. We are attempting to define a level of the French language that possesses a certain generality of structure. A consequence of this approach is that variations observed in dialects will be described as departures from this abstract variety of French that will serve as a basis of reference. This approach is thought to lead, on the one hand to a simple grammar of reference for French, and on the other, to simple rules that will relate the varieties of French. In contrast, rules that would describe directly a level of French closer to observations (as found in a written or spoken corpus) are likely to fuse together idiosyncrasies and regularities. Such rules would be highly complex and they would tend to obscure basic processes of the language.

An obvious consequence of this approach is that the grammar to be built will carry a complex relationship to actual speech.

Again, these considerations are just the standard precautions taken in any experimental study in the hard Sciences. We have only sharpened procedures that are already agreed on by many linguists.

1.3 A REANALYSIS OF THE CONCEPT OF OBJECT

A substantial component of the lexicon-grammar of French is the set of the verbs taken with their sequence of object complement(s). The programme of construction of a lexicon-grammar of verbs basically requires us to make precise and operational the definition of object, since the term, as used in traditional and generative grammars, is inadequate.

Traditional definitions of transitive and intransitive verbs always mix syntactic criteria together with semantic intuitions about verbs and object complements. Object is another ill-defined concept which is not independent of the notion of transitivity. The formal version of object given in Chomsky 1965 is not essentially different from these, since it is connected to the notion of predicate or verb phrase, for which no operational definition can be found.[12] As a consequence, these notions cannot be used for the methodical determination of verb-object combinations.

In conformity with our general approach, we used an elementary test for the classification of complements: the shapes of the associated interrogative pronouns. The pronouns *qui*—'who/whom', *que/quoi*—'what' correspond to what we call direct and indirect objects, while the pronoun *où*—'where', *quand*—'when', *comment*—'how', etc., also called interrogative adverbs, are associated with circumstantial complements.

This test resembles the school test used in recognition

drills of complement types in text examples. Since the theoretical relations between the shapes of pronouns and the traditional abstract definitions of object are quite obscure, this test always tended to act as a substitute for the definitions. We have systematized this practice, but in order to obtain results that are reproducible over a large lexicon, we had to sharpen the test (i.e. our definition) in several respects.

For example, there exist complements that answer to both types of interrogative pronouns (object and circumstantial):

Max tire son argent d'une grosse ferme
(Max gets his money from a big farm)

De quoi Max tire-t-il son argent?
(From what does Max get his money?)

D'où Max tire-t-il son argent?
(Where does Max get his money from?)

In such a situation, the complement *de* N_2 =: *d'une grosse ferme* is ambiguous and some decision must be made as to how to represent its double status.

Other unexpected situations result from the strict application of the test. Consider the application of our definition of direct object to the sentences

Max est un idiot
(Max is an idiot)

Max a une ferme
(Max has a farm)

Both have a direct object, since the following dialogs are acceptable:

--Qu'est Max?

--Un idiot

(--What is Max? --An idiot.)

(5) { *--A qui Max tient-il?*
--A Léa

(6) { *--A quoi Max tient-il*
(7) *--A (ce qu'on écoute Léa, écouter Léa)*

These complements do not answer to circumstantial pronouns.

But now consider the following sentences that are identical in form to (4):

(8) *Max cherche à ce (qu'on écoute Léa, écouter Léa)*
(Max wants (people) to listen to Lea)

The only question that these sentential complements accept is:

(9) *Que cherche Max?*

which is a direct question: (3) and (9) are identical question forms. On the other hand, the answer to (9) is the indirect form (7).

Our definition thus has an undesirable consequence, since we are led to call direct objects certain complements which take *Prep =:* à and which are called indirect objects in other formally identical contexts. Other syntactic properties such as the possibility for a complement with *Prep =:* à to take the form of the pre-verbal pronoun *y* are not observed; cf. the contrast

Max y tient vs. **Max y cherche*

This discrepancy is limited to sentential complements. With nouns, the preposition à is not acceptable:

**Max cherche à (Léa, ce livre, etc.)* [13]

Another small set of examples is represented by

Max apprend à lire
(Max learns to read)

(10) {*--Qu'est-ce que Max apprend?*
{*--A lire (*-- Lire)*

Note the interdiction:

**A quoi est-ce que Max apprend?*

On the basis of such observations, we might feel tempted to abandon the test as a basis of the classification. However, a study of the lexicon shows that the phenomenon is limited to a small number of verbs: fewer than ten, as compared to the several hundred verbs and adjectives that take *Prep =: à* under the same conditions. Moreover, we did not feel it worthwhile to eliminate such exceptional verbs from the definition, for example by using the properties of the pre-verbal pronoun *y*. In order to do so, we would have had to complicate the test, which would have little result, except, possibly, for the introduction of a slightly better fit with the questionable set of traditional practices. Hence, we consider these few verbs as exceptions to the definition of object. How to explain their presence in the language is an open question.

Our experience with the use of the interrogation test as a tool of classification has shown that it is both general and reliable. In some cases however, it may have to be sharpened. For example, the shapes of the three questions (2), (3) and (10) are different:

- the difference between (2) and (3) is related to the category of the object noun (human or non-human). In (3), the pronoun *il* which duplicates the subject is forbidden with human objects:

  **Que Max aime-t-il?*

  Qui Max aime-t-il?

The permutation of the subject seen in (3) is forbidden if the object is human. Such a forbidden permutation yields the following form:

Qui aime Max?
(Who does Max love?)

which is accepted only as a question on the subject. Questions on non-human subjects generally lead to unacceptable forms such as

**Qu'amuse Max?*

- the periphrastic question (10) has more generality, for there are no restrictions on the categories of the nouns appearing in it;
- another possibility of applying an interrogation test is by means of questions without permutation of the interrogative pronoun. The intonation of such questions is in general clearly felt. The forms are:

Max aime (qui, quoi)?
Max tient à (qui, quoi)?

The pronoun *que* is not allowed in these forms, hence, they are more regular; they can be analyzed as the result of the substitution of only two pronouns (i.e. *qui, quoi*) in the complement phrase.

Exactly which question form is chosen has an impact on the classification. Suppose we decide to use the question without permutation in the definition of objects. We then observe that certain complements which could not enter into the interrogative forms (2) or (3) can now be submitted successfully to the test. For example, the sentence

(11) *Ils ont élu Max président*
(They elected Max president)

is usually analyzed with *Max* as a direct object and *prés* -

ident as an attribute of the object. We thus have

--Qui ont-ils élu président?

--Max

(--Who did they elect president? --Max.)

The unacceptable forms

**(Qui, que) ont-ils élu Max?*
(*(Who, what) did they elect Max?)

forbid us to call *président* a direct object. However, the following form without permutation:

Ils ont élu Max quoi?

is accepted.[14] If we adopt this version of the test, we are forced to call *président* a direct object, and *élire* a double transitive verb. In fact, we have adopted this solution, and accordingly, we call *élire* a verb with two direct objects or a double transitive verb, a terminology that has no effect on the formal analysis of sentence (11). Traditionally, the categories double direct object or doubly transitive are excluded from French grammar. The reason seems to be that a transitive verb must have a passive form. Hence, a double transitive verb should have two passive forms, as for example *to give* in English. But we have observed that Passive and Interrogation are independent syntactic properties, even though they both yield acceptable forms with the majority of transitive verbs. Indeed, we have found such pairs as

Ce problème concerne Max

Ce problème regarde Max
(This problem concerns Max)

which are generally closely synonymous sentences. According to the interrogation test, *Max* is a direct object in

both sentences, but when we apply the Passive rule, we observe

Max est concerné par ce problème

**Max est regardé par ce problème*

With respect to other properties (relativization, preverbal pronominalization, etc.) no difference is observed.

The following example has a passive form:

(12) *Max spent two hours looking at the stars*
= Two hours were spent by Max looking at stars

But the complement *two hours* yields an unacceptable sentence when it is submitted to the question with *what*:

**What did Max spend looking at stars?*

Again, we see that Passive and Interrogation are independent rules. The same situation of independence would be found if we were to choose as a basis for the definition of objects the shapes of relative pronouns or of preverbal pronouns. Thus, had the direct object been defined by the relative pronouns *qui*—'who/whom', *que*—'which', the complement *two hours* in (12) would have become a direct object, since we accept

The two hours which Max spent looking at stars
[were not lost]

In fact, these observations of the independence of syntactic properties show that such notions as direct object or transitive verb, which have been proposed on the basis of limited observations by grammarians, have no place in a formal grammar (M. Gross 1969). When more complex facts than those found in textbooks had to be accounted for, traditional grammarians progressively voided definitions of their operational content, by introducing abstract definitions such as those found in today's manuals.

Now that we have a reasonably complete picture of a language with respect to the syntactic components of the notion object, we can see that the latter covers a large variety of semantic situations; the attempt to give a precise description of it forces us to limit the representation of properties only to the reproducible linguistic phenomena, that is, to combinatorial features such as the shape of pronouns (interrogative, relative or pre-verbal), the passive forms, etc. Semantic notions such as "active" verb, "static" verb, or complement of "result" have an entirely different status and cannot be represented in a lexicon-grammar. The relations, if any, between meaning and syntactic properties are quite obscure.

We have discussed several phenomena which involve terminological issues and we have concluded that the elimination of notions such as object or transitivity from the grammar is necessary. However, a different attitude could be adopted. Our discussion can also be viewed as a proposal for a precise redefinition of the traditional notions that we have criticized. For example, one might find it desirable to keep using the terms direct and indirect objects, transitive and intransitive verbs, thus preserving a certain continuity between the older studies and the modern ones. Also, if some day, improved grammar rules are introduced in schools, totally new terms will not be necessary, and many examples teachers are using will still be valid.

Only a caucus of linguists who have at their disposal grammars with full coverage of the lexicon can make a rational choice among concepts, with a view both to the circulation of their results in the community and to such applications as language teaching and automatic syntactic analysis.

2. THE PRESENT STATUS OF LEXICON-GRAMMAR DESCRIPTIONS

2.1 LEXICON-GRAMMAR OF VERBS

Lexicon-grammars are presently being built for several languages (Cf. notes 2, 3, 4, and 5).

For French, the situation is the following: The lexicon of verbs has reached a certain level of completeness (1.1). About 10,000 verbs have been classified according to a system of about fifty classes, which are defined in terms of the object sequence of verbs. Thus, there are classes of verbs without direct object (intransitive verbs) which enter into a structure

$$N_0\ V\ Prep\ N_1$$

where $Prep\ N_1$ is different from the indirect object complements (*à* + *de*) N_1.

Structures with one object are of the three main types:

$$N_0\ V\ (E + à + de)\ N_1$$

Structures with two objects are essentially

$$N_0\ V\ N_1\ (E + à + de)\ N_2$$

Other structures are used in the classification. Consider the relation (A) = (B):

(A) $N_0\ V\ N_1\ de\ N_2$ =:

Max a chargé le camion de caisses
(Max loaded the truck with boxes)

(B) $N_0\ V\ N_2\ Loc\ N_1$ =:

Max a chargé des caisses dans le camion
(Max loaded boxes onto the truck)

As usual with syntactic relations (e.g. transformations), one finds verbs that enter into only one of the two members.

The verb *ranger*—'to stow' is such a case and enters only into (B):

**Max a rangé le camion de caisses*
(*Max stowed the truck with boxes)

Max a rangé des caisses dans le camion
(Max stowed boxes in the truck)

Verbs like *ranger* are then classified with respect to structure (B(which contains a nonobject complement.

Observations on this construction leads to the result that verbs with two prepositional objects are exceptional. Only about 20 verbs enter into the structure with objects in à and *de*:

N_0 *V a* N_1 *de* N_2 =:

Max a parlé à Bob ce de livre
(Max talked to Bob about this book)

The structure

N_0 *V* à N_1 à N_2 =:

Cet outil sert à Max à cela
(This tool helps Max for that)

is limited to fewer than ten verbs and adjectives, and the only known case of

N_0 *V de* N_1 *de* N_2 =:

Max a hérité d'une maison de sa tante
(Max inherited a house from his aunt)

is not universally accepted.

No example of a verb with three objects (according to our definition) has been found so far.

The main class types are given in Tables la and lb. Each class is defined by one type of structure, in general. The types have been further subclassified according to

such criteria as the nature of the various arguments:

N_i =: *Nhum + que S + V-inf Compl + etc.*

Table 1a

Classes of Verbs with Sentential Subject and/or Complement(s)

Tables	Structures	Numbers
1	N_0 *U Prép V° Ω*	80
2	N_0 *V V° Ω*	150
3	N_0 *V* N_1 V^1 *Ω*	45
4	*Qu P V* N_1	550
5	*Qu P V Prép* N_1 *Il V Prép* N_1 *Qu P*	180
6	N_0 *V Qu P*	450
7	N_0 *V a ce Qu P*	170
8	N_0 *V de ce Qu P*	225
9	N_0 *V Qu P à* N_2	360
10	N_0 *V Qu P Prép* N_2	180
11	N_0 *V* N_1 *à ce Qu P*	180
12	N_0 *V Qu P* N_0 *V* N_1 *de* $V^1\Omega$	60
13	N_0 *V* N_1 *de ce Qu P*	140
14	N_0 *V à ce Qu P Prép* N_2	20
15	N_0 *V de ce Qu P Prép* N_2	70
16	N_0 *V Prép Qu P Prép* P_2 *ce Qu P*	60
17	*Il V Prép ce Qu P Prép* N_2	10
18	N_0 *V Prép* N_1 *Prép* N_2 *Prép Qu P*	20
19	*Qu P V* N_1 *Prép* N_2	50
	TOTAL	3000

Table 1a corresponds to the classes of verbs published in Gross 1975. These verbs all have at least one sentential subject or complement.

Table 1b is divided in three sections that roughly correspond to three publications:

- Boons, Guillet, Leclère 1976a (to be revised) describes the simple transitive verbs
- Boons, Guillet, Leclère, 1976b, the intransitive verbs
- Guillet, Leclère, Boons, 1983, verbs with two complements, the second is Locative.

490

Table 1b

Classes of Verbs Without Sentential Complement

Tables	Structures	Examples	N°
31I	il V	il pleut	37
31H	NOhum V	Max bêtifie	380
31R	NO-hum V	l'eau bout	185
33	NO V à N1	Max vaque à ses travaux	90
34LO	NO V Loc N1	Les rats grouillent dans le pré	
	N1 V de NO	Le pré grouille de rats	100
35S	NO V avec N1	Max flirte avec Léa (SYM)	80
35L	NO V Loc N1	Max aborde au quai (ACTIF)	190
35ST	NO V Loc N1	Max gît sur le sol (STATIQUE)	
35R	NO V Prép N1	Max abuse du tabac	285
35RR	NO V Prép N1 Prép N2	(not fully known)	
32R1	NO V N1	Ceci confirme Ida dans ses idées	60
32R2	NO V N1	Max brandit le livre	440
32R3	NO V N1restr	Max embrasse cette carriere	150
32A	NO V N1app	Max construit une maison	100
32RA	NO V N1	Le temps jaunit le papier	300
32C	NO V N1	Max a abimé le livre	370
32CL	NO V N1	Max embrasse Ida (sur le nez)	95
32CV	NO V N1 en N2	Max tranforme Ida en grue	50
32H	NO V N1hum	Max a violé Ida	450
32NM	NO V N1	Max pèse 100 kilos (*PASSIF)	70
32PL	NO V N1pl ob1	Max entasse les caisses	100
32L	NO V N1lieu	Max gagne la porte	60
36DT	NO V N1 à N2hum	Max donne le pin à Luc	200
36S	NO V N1 avec N2	Max compare Luc (avec+et) Bob	95
36SL	NO V N1 Loc N2	Max colle le pin (sur+avec+et) le fer	90
36R	NO V N1 à N2	Max conjugue ce V au futur	60
37E	NO V N1 de N2	Max vide la cuve du fuel	230
37M1 to 37M6	NO V N1 de N2	Max remplig la cuve de fuel	880
38LD	NO V N1 Loc N2	Max met le pain sur le lit	215
38LS	NO V N1 de N2	Max ôte le pain du lit	90
38LH	NO V N1 Loc N2	On a congédié Max de son poste (Nihum)	180
38LR	NO V N1 Loc N2	On marine le veau dans le vin	168
38LO	NOLoc V N1	Max avale le vin	125
38PL	NO V N1 en N2pl	Max coupe le pain en deux bouts	80
38R	NO V N1 Prép N2	Max entame son speech par un toast	150
38RR	NO V N1 Prép N2, Prép N3	(Not fully known)	
39	NO V N1 N2	On a élu Max président	60
		TOTAL	6000

Each class is represented in the form of a table, i.e., a matrix whose rows correspond to verbs and whose columns are structures into which each verb may enter. A "+" sign is placed at the intersection of a row and a column when the verb enters into the corresponding structure, and a "-" sign otherwise (cf. Annex).

Columns are associated with *single* structures, hence, transformations correspond to *sets* of structures (e.g., pairs of columns). Some nontransformational relations are represented in the same way.

At the present stage of the lexicon-grammar, about 500 properties are formally represented, this figure may vary slightly according to the way they are counted. Since properties vary with the number and nature of complements, the number of relevant properties is variable with each class. On the average, there are about thirty properties attached to a class or table.

There are practically no two verbs that have identical properties (i.e., identical rows of "+"s and "-"s), nor are there pairs of columns with identical content. Hence, the corresponding transformations all have exceptions. Nonetheless, certain classes of verbs are remarkably homogeneous, both syntactically and semantically, i.e., on the basis of their lexical content. In such classes, syntactic properties and meaning are strongly correlated. However, we do not consider that these situations exemplify the traditional view according to which syntactic properties are predictable from the meaning (e.g. the prediction of the use of French subjunctive from the obvious meaning of the verb: verbs of "will," of "doubt," etc.). There are several reasons for refusing the traditional association.

First, these "natural" classes have emerged in limited

numbers. Thus, the majority of verbs are parts of such meaning-form associations, but they correspond to cases which are isolated both semantically and syntactically. For them, the association holds trivially: the unique meaning of the verb determines its unique syntactic properties. In this case, we have no need for abstract notions of meaning, since the association between *word* and syntactic forms is sufficient.[15]

Second, there may be a general way of explaining the existence of these "natural" classes. We have proposed an analysis of these classes which we call Fusion and which we can illustrate with the following example:

Consider the verbs of saying in French. There are over 200 of them, when enumerated on an intuitive basis. The class can be easily augmented, for example, by adding to it the verbs used specially for the sounds emitted by particular animals (e.g. wild fowl). Other sound verbs are less productive. The difference concerns such verbs as

Max a miaulé à Bob de partir
(Max mewed to Bob to leave)

Max a tambouriné à Bob de partir
(Max drummed to Bob to leave)

All these verbs can be said to be extensions of the verb *dire*—'to say/to tell', in the following sense:

- semantically, they all mean *dire* in a certain manner (Zwicky 1971);
- syntactically, they all possess most of the properties of *dire*

One of the properties that define a subclass of verbs of saying is the fact they can also be used as verbs of command, but under different syntactic conditions. The meaning of saying is associated with the two related structures:

493

(1) N_0 *à* N_2 *que S =:*

Max a dit à Léa qu'il était content

=(2) N_0 *V à* N_2 V^0 *Compl =:*

Max a dit à Léa être content

(Max told Lea that he was satisfied)

The sentence *S* is in the indicative, and the reduced infinitive form V^0 *Compl* has N_0 (=:*Max*) for subject. The meaning of command corresponds to the two related structures:

(3) N_0 *V à* N_2 *que Ssubj =:*

Max a dit à Bob qu'il parte
(Max told Bob that he should leave)

=(4) N_0 *V à* N_2 *de* V^2 *Compl =:*

Max a dit à Bob de partir
(Max told Bob to leave)

There are verbs which are verbs of saying only, and enter into the forms (1) and (2) (e.g. *raconter*—'to tell'). Similarly, there exist verbs strictly of command, which are limited to structures (3) and (4) (e.g. *ordonner*—'to order').

More than 150 verbs have both meanings, and at the same time the four syntactic properties (1)-(4), among other common ones. Also, verbs of saying have individual properties that distinguish them from each other.

Extensions of *dire* often have an intransitive proper meaning, such as

Un chat a miaulé
(A cat mewed)

and the corresponding verb of saying can be viewed as a metaphoric extension of this proper use. We formalize these remarks

- first, by allowing human subjects for these verbs that are not properly accepted. This step is the metaphoric extension leading to

  Max a miaulé
  (Max mewed)

- second, by applying the rule of Fusion:

  Max a dit à Bob de partir,
  en (miaulant, tambourinant)

  = *Max a (miaulé, tambouriné) à Bob de partir*

This rule has a morphological effect: the verb *dire* is replaced by *miauler* or *tambouriner*, while all properties of *dire* are left unchanged. One can illustrate this effect by means of the feature representation of verbs (Chomsky 1965). Suppose we have the representations

dire	*miauler*	*tambouriner*
F1	*G1*	*H1*
F2	*G2*	*H2*
.	.	.
.	.	.
.	.	.
Fl	*Gm*	*Hn*

where the *Fi*'s, *Gj*'s and *Hk*'s are syntactic marks, that is "+" marked columns of the lexicon-grammar, (M. Gross 1975). In these examples, the features 1 to 4 might correspond for example to the four structures (1) to (4). After Fusion applies, we have

miauler	*tambouriner*
F1	*F1*
F2	*F2*
.	.
.	.
.	.
Fk	*Fk*
G1	*H1*
Gm	*Hn*

We now see that some natural classes can be analyzed with the help of a rule and of other classes; in the example, we used the class of verbs of animal sounds, but more generally, the class of verbs of emission of signals, whether auditive or visual is involved (e.g. *faire signe à Jo de partir*—'to wave to Jo to leave'). Locative causative verbs (Guillet, Boons, Leclère to appear) and verbs of motion (Lamiroy 1983) have also been analyzed in this way.

On the one hand, Fusion accounts for the change of syntactic properties, on the other, its range of application appears to be semantically determined in a transparent way. It is in this sense that Fusion formalizes commonly observed superficial form-meaning associations.

2.2 LEXICON-GRAMMAR OF FROZEN EXPRESSIONS

Frozen or idiomatic expressions are of varied syntactic types. We describe their syntactic properties, without attempting to represent their syntactic peculiarities.

There are idiomatic simple sentences such as

(1) *La maison a pris feu*
(The house caught fire)

(2) *Max se bat contre des moulins à vent*
(Max tilts at windmills)

(3) *Bob a croisé le fer avec Luc*
(Bob crossed swords with Luke)

(4) *Max connaît Bob de vue*
(Max knows Bob by sight)

(5) *Max brûle la chandelle par les deux bouts*
(Max burns the candle at both ends)

These sentences can be analyzed according to their complement sequence. Thus, (1) has the structure subject-verb-direct complement, that is, $N_0\ V\ N_1$. We will designate frozen noun phrases by the symbol C_i parallel to N_i in the free case. (1) is then represented as $N_0\ V\ C_1$. Accordingly, the other examples have the following structures:

Table 2

Classes of Frozen Sentential Expressions

Tables	Structures	Examples	
C1	N_o V C_1	Il a loupé le coche	2 168
CAN	N_oC($_1$C à de N)	Cela a délié la langue de Max (lui)	477
CDN	N_oV($_1$C de N)	Il bat le rappel de ses amis	404
CP1	N_oV Prép C_1	Il charrie dans les bégonias	1 212
CPN	N_oV Prép ($_1$C de N)	Il abonde dans le sens de Max	207
C1PN	N_oV C_1Prép N_2	Il a déchargé sa bile sur Max	1 648
CNP2	N_oVN_1Frép C_2	Ils ont passé Max par les armes	1 260
C1P2	N_oVC_1Prép C_2	Il met de l'eau dans son vin	716
C5	Que P V Prép C_1	Que Max reste milite en sa faveur	140
C6	N_oV Qu P Prép C_2	Il a pris du bon côté que Max reste	274
C7	N_oVC_1 à ce Qu P	Il a dit non à ce que Max reste	108
C8	N_oV C_1de ce Qu P	Il se mord les doigts de ce qu'il est resté	253
CADV	N_oV Adv	Cela ne pisse pas loin	180
CX	N_oV X	Il est parti sans laisser d'adresse	80
CO	C_oV Ω*	La moutarde monte au nez de Max	1 135
A1	N_o avoir C_1	Il a eu le mot de la fin	124
A1PN	N_o avoir C_1 Prép N_2	Il a barres sur Max	92
A1P2	N_o avoir N_1 Prép C_2	Il a Max en horreur	73
A12	N_o avoir C_1 Adj_1	Il a la vue basse	89
A1P2	N_o avoir C_1 Prép C_2	Il a mal aux cheveux	242
EO1	C_o de N être Adj	La barbe de Max est fleurie	319
EOP1	C_o être Prép C_1	Les rieurs sont du côté de Max	150
		TOTAL	11 351

*We use the symbol Ω as an equivalent for the symbol *Compl*.

(2) := N_0 V Prep C_1

(3) := N_0 V C_1 Prep N_2

(4) := N_0 V N_1 Prep C_2

(5) := N_0 V C_1 Prep C_2

Subjects can also be frozen, as in

(6) C_0 V Compl =:

Les dents de Jo s'entrechoquent
(Jo's teeth knock together)

There are other expressions which are globally similar to these, and whose (frozen) C_i's accept free noun complements, as for example in

Max mangera dans le creux de la main de Bob
(Max will eat out of Bob's hand)

Max a coupé l'herbe sous le pied de Bob
(Max cut the ground from under Bob's feet)

These are represented as

(7) N_0 V Prep $(C \text{ de } N)_1$

(8) N_0 V C_1 Prep $(C \text{ de } N)_2$

Structures (1) to (8) exemplify the principles on which our classification of frozen sentences has been built. Table 2 provides numbers for the main syntactic classes. Table 2 corresponds mainly to "full" verbs, that is, verbs that differ from the "empty" ones *être*—'to be', *avoir*—'to have', *faire*—'to make'. Frozen sentences with empty verbs appear in a different component of the grammar, for they have characteristic syntactic properties (2.3).

In each class, various syntactic properties correspond to the columns: Passive, Pronominalizations, etc. (cf. Annex).

Frozen sentences are also being described by the same

498

Table 3

Classes of Frozen Adverbials

Tables	Structures	Examples	Numbers
PO	Adv	soudain	174
PC	Prép C	en bref	435
PDC	Prép Dét C	contre toute attente	422
PCDC	Prép C de C	en désespoir de cause	240
PCDN	Prép C de N	au moyen de N	276
PAC	Prép Adj C	de sa belle mort	310
PCA	Prép C Adj	à gorge déployée	253
PCPN	Prép C Prép N	par rapport à N	48
PCPC	Prép C Prép C	des pieds à la tête	125
PCONJ	Prép C Conj C	en tout et pour tout	112
PV	Prép V Compl	à dire vrai	90
PECO	(Adj) comme C	comme ses pieds	197
PVCO	(V) comme C	comme un cheveu sur la soupe	194
PPCO	(V) comme Prép	comme dans du beurre	23
PF	(frozen) S	Dieu seul le sait	192
		TOTAL	3 101

method for Italian, and English (Elia to appear, Freckleton 1983, Machonis 1982).

The main global observation made on frozen sentences is that their grammar is exactly the same as the grammar of free sentences, provided they are represented as described here. It is essential that the representation adopted eliminates any attempt to reduce frozen forms to analogous free ones. For example, we do not try to package *kick* and *the bucket* into one syntactic unit that would become a verb analogous to the intransitive verb *to die*. We analyze *the bucket* as the direct complement of the verb *to kick* and the full sentence has the structure $N_0\ V\ C_1$. The fact that this structure has no passive form is no different for us than the restrictions on Passive observed on free forms. In fact, many frozen sentences $N_0\ V\ C_1$ (more than half) accept Passive. Exactly the same situation has been found to hold for all transformational relations.

We have also studied frozen adverbials, but more from a simple morphological point of view. The main types are given in Table 3. Again, the internal properties of these phrases have not been found to differ from the properties of comparable free forms.

We have tabulated 13,000 frozen sentences as compared to 10,000 free ones. Although there may be a few discrepancies between the ways each type is counted, it is already clear that there are many more frozen forms than free ones, from the lexical point of view. The lexicon-grammar of free verbal sentences presented in 2.1 has a certain degree of completeness, whereas the lexicon-grammar of frozen forms is still far from such a status.[16]

Compound nominals constitute another important type of frozen construction. Compound nouns form the most produc-

tive part of most languages, since the vast majority of technical and scientific nouns are produced by composition of simple nominals.

Thus, the frozen component of the lexicon-grammar of a language appears to be an essential part of it, and not a collection of exceptional items, as traditional and generative studies tend to suggest.

These preliminary observations on frozen forms should have consequences for the current views on learning, and in particular, about the amount of rote learning involved.

2.3 SUPPORT AND OPERATOR VERBS

There is at least a third component of the lexicon-grammar of French, and of other languages as well, which involves special verbs called support verbs (*Vsup*) and operator verbs (*Vop*) (M. Gross 1981, La Fauci 1979). Examples of *Vsups* are

(1) *This book is of a certain importance*

(2) *This text is in contradiction with ours*

(3) *This book had a certain impact on Max*

(4) *Bob has hatred for such people*

(5) *Max made a mistake*

(6) *Max took Bob's advice*

Here one cannot argue that the verbs *to be Prep, to have* and *to make* are regular verbs with selectional restrictions. It is intuitively clear that the main predicative element is the complement noun in all these examples.

One way of representing this information is by introducing nominalization relations between sentences (Z. S. Harris 1964), such as, for example

(1) = *This book is important*

(2) = *This text contradicts ours*

(3) = *Bob hates such people*

501

However, the phenomenon is more general than this, as examples (3), (5) and (6)—which bear on nonderived nouns—indicate.

Also, many other verbs appear to have the same character:

(7) *This book grew in importance*

(8) *This book has attained a certain importance*

(9) *This text enters into contradiction with ours*

(10) *Bob lost all hatred for such people*

(11) *Max nurtures hatred for such people*

In addition to the intuition that the main verb has no essential semantic role, whereas a deep relation holds between subjects and supported complements, these structures have remarkable syntactic properties that distinguish them from regular sentences. Consider for example the sentence

Max discusses hatred for such people

which is superficially similar to (11). Let us apply clefting to the phrase *for such people* which plays the same role of noun complement of *hatred* in both sentences. We obtain

It is for such people that Max nurtures hatred

**It is for such people that Max discusses hatred*

Operator verbs are connected with support verbs. Consider the sentences

(12) *The media gave a certain importance to this book*

(13) *Your remark puts this text in contradiction with ours*

(14) *This event caused Bob's hatred for such people*

We have reasons to consider that the relation observed in (1), (2) and (4) respectively between subjects and complements has become a relation between both complements in

(12), (13) and (14). Here, the three verbs *to give, to put* and *to cause* are causative operators that have been applied to sentences containing a support verb.

Many verbs now believed to be ordinary and represented as such in the lexicon-grammar will have to be reanalyzed as *Vsups* or *Vops*. The extent of this reanalysis cannot yet be estimated, since there is no unique and simple method of distinguishing among these categories. In the same way, certain frozen sentences appear to be cases of lexically restricted *Vsups* and *Vops*.

There are various relations between the two new categories, for example

- relations among *Vsups* that are semantic variants (modulo aspect or other modalities):

  (1) = (7) = (8)

  (2) = (9)

  (4) = (10) = (11)

- relations between *Vsups* and *Vops*:

  (1) = (12)

  (2) = (13)

  (3) = (14)

Hence, operator and support verbs emerge as a constituent of the grammar of simple sentences, that is, a quite different component from that of ordinary verbs or of frozen sentences (Danlos 1981, Giry-Schneider 1978, Labelle 1974). The existence of more complex examples than those given here (G. Gross 1982, M. Gross 1981) involving many as yet unrecognized structures may force us to revise current views about simple sentences entirely.

3. COMPARING LANGUAGES

As already noted, we have to work within a stable theoretical framework. In this respect, the main feature of Z. S. Harris' framework is that it is built up from a sharply limited set of abstract entities. These notions are well founded, and are as language-independent as possible, in the present state of knowledge. Hence, comparing languages that have been described individually within such a general theory is a concrete activity. In contrast, it is hard to see how comparisons could be made between facts stated within the variety of theories in current use. For example, within a given language, it is already hard to see how analyses in terms of case grammar could be compared with descriptions in the generative semantics style or in the formalism of relational grammar. The same appears to be true for any two variants of generative grammar. It is sometimes possible to detect common features of a given phenomenon across theories, but in general, this is done by abandoning the specific character of the theories involved, which, more or less explicitly, implies returning to naive transformational theory.

Lexicon-grammars built according to the procedure outlined in 1 should be less subject to the artificial formal problems whose origin is often in the theoretical background itself. Also, since these facts have a certain homogeneity, they should lend themselves more directly to comparisons. However, new problems arise concerning the manner and the level at which languages should be compared.

Limiting the size of the abstract theoretical apparatus is a concern not only of recent formal theories, but of traditional grammar as well. Because it is inherited from Latin-Greek, the metalanguage of traditional grammar has often been considered to be universal, hence more suitable

for comparisons. However, we have seen in 1.3 that it is not entirely exempt of redundancies and inconsistencies, besides being unformalizable.

Families of languages have been constituted up until now mainly on the basis of morphological criteria. Only recently, syntactic criteria have been considered on a systematic basis (Greenberg 1966, Klima 1962). Many questions have been raised about the validity of such classifications (e.g. Meillet 1952). The sharpening of linguistic studies has revealed the existence of numerous new properties that all appear to be valid parameters of classification.

As is usual in taxonomic enterprises, seemingly characteristic features may become more numerous than the items to be classified. On the other hand, using single features for classificatory purposes may lead to arbitrary classifications, if the status of the feature with respect to all others is not known. An equilibrium has to be found in terms of "natural" classes, a notion that cannot be defined *a priori*.

For example, consider the feature "zeroable subject pronouns." Italian, Spanish and Portuguese share this syntactic property, whereas French does not have it; in this respect, French is closer to English or German. If we were to use the causative construction as an essential classificatory feature, Japanese (Kuroda 1965) and French (Kayne 1977) would be classed in a family that would not include English, nor some other Indo-European languages. We can easily multiply such examples and thus propose a set of properties with which we could group together or separate at will any arbitrary combination of languages. This seems to us to be the present state of language classification, a state of flux caused by a surge of knowledge

505

which, presumably, will continue to increase, especially if collective efforts of description are set up.

Nonetheless, we think that there is enough evidence of similarity among Romance languages to consider them a privileged family.[17] Our program consists in pursuing independent constructions of lexicon-grammars for these languages, in order to systematize the search for similarities and differences, both from the syntactic and the lexical points of view.

3.1 COMPARING STRUCTURES AND RULES

3.1.1 STRUCTURES WITH DIRECT OBJECTS

Consider again the notion of direct object. In French, as well as in other languages, this notion implies the absence of a preposition before a complement. But to be made operational, the concept must be corrected in various ways:

- determiners or predeterminers may begin with a preposition. As a result, a direct complement may resemble superficially an indirect one. A well-known example is the French partitive determiner:

  Max lit de la science-fiction
  (Max reads science-fiction)

There are many prepositions occurring in this position, for example:

Max lit aux alentours de dix livres par jour
(Max reads about ten books a day)

In Spanish, the notion direct object does not seem to conform to the formal definition used in other languages. Numerous Spanish verbs take a human direct object preceded by *Prep =: a*, whereas a non human object in the same position would appear without a preposition. One possibility

for making these constructions comparable to those of French, Italian and Portuguese[18] is by classifying the complements *a N* of Spanish according to the shape of the corresponding pronoun:

- *a N = lo(s), la(s)* would correspond to direct objects,
- *a N = le* would correspond to indirect objects.

Other criteria could be used, such as Question or Passive; in general, each yields a slightly different list of transitive verbs.

3.1.2 *EXTRAPOSITION*

There is no fundamental difference between comparing structures and comparing rules. Since a transformational rule can be seen as a pair of structures, comparing two rules amounts to comparing the corresponding pairs of structures.

Consider for example the rule of Extraposition, as described in English and in French. This rule has a member which is a structure with an impersonal subject pronoun. Examples are

Que Max soit parti a été noté
= *Il a été noté que Max était parti*

That Max left was noticed
= It was noticed that Max left

One feature of the relation or rule is that the subject is moved to a post verbal position. Another point is that the subject position must be filled by an impersonal pronoun, which is in the singular and agrees with the verb, even when the "source" subject is in the plural.

Let us now consider the situation in Italian, Spanish and Portuguese:

- subject pronouns are optional, whereas in English and French they are obligatory. Often, sentences without any subject pronoun are to be preferred to those containing one;
- a full noun phrase subject can easily be permuted; indeed, the post-verbal position for the full subject is sometimes preferred. The rule of permutation involved in such sentences does not have the restrictions found with English and French Extraposition.

As a consequence of these two features, certain sentences of Italian, Portuguese or Spanish of the form

$$V\ Compl\ N_0 \quad \text{or} \quad V\ N_0\ Compl$$

may be obtained from the form $N_0\ V\ Compl$ by the general rule of permutation, or else one could argue that a rule of Extraposition similar to the English or French one has applied, but that the (unattested) impersonal pronoun has been erased by the general rule of subject pronoun deletion, which is obligatory here. Hence, given sentence forms such as

(It) Arrivano tre treni
(Arrived three trains)

A Bob piace che Max venga
(To Bob pleases that Max comes)

(Po) Agrada ao Bob que o Max venha

(Sp) Cundió el pánico

one cannot declare categorically that they have undergone Extraposition.

One general feature could be retained to define Extraposition across languages: the change of the verb number from plural to singular when a plural subject has moved to the right of its verb:

(Po) *Vendem-se cada vez menos andares*

Vende-se cada vez menos andares

(Sp) *Liegaron tres trenes*

**Liegó tres trenes*

This definition meets with other problems, in Italian (Elia, D'Agostino, Martinelli 1981) and in Portuguese (Malaca 1981).

Thus, although the comparison of Extraposition in English and in French is a well defined problem, many non-trivial questions have to be answered in each of the other Romance languages before the actual comparison can be performed.

3.1.3 *SUBJECT OF INFINITIVE CONSTRUCTIONS*

Consider the case of the personal infinitive in Portuguese, which, traditionally, is considered as an idiosyncrasy of that language. However, there is a way of analysing it which is close to the description of certain French infinitive forms.[19] Examples in Portuguese are:

Tu fazeres isso a Maria choca o Pedro
(That you did this to Mary offended Peter)

O Pedro reconheceu ter a Maria razaõ
(Peter acknowledges that Mary was right)

French examples are:

(1) *Max a dit à Jo de partir*
(Max told Jo to leave)

(2) *Max a promis à Jo de partir*
(Max promised Jo to leave)

Intuitively, we feel that in (1), *Jo* is the subject of the infinitive verb *partir*, whereas in (2), *Max* is the subject. This intuition can be given a formal basis, by selecting an intrinsically reflexive verb in the infinitive comple-

ment, as in

(3) *Max a dit à Jo de se conduire bien*
(Max told Jo to behave herself nicely)

Max a promis à Jo de se conduire bien
(Max promised Jo to behave himself nicely)

In these sentences, the reflexive pronoun must agree in person and number with the intuitive subject:

Tu as dit à Jo de se conduire bien
(You told Jo to behave herself nicely)

**Max t'a dit de se conduire bien*
(*Max told you to behave himself nicely)

Max t'a dit de te conduire bien
(Max told you to behave yourself nicely), etc.

Tu as promis à Jo de te conduire bien
(You promised Jo to behave yourself nicely)

**Tu as promis à Jo de se conduire bien*
(*You promised Jo to behave herself nicely)

Max t'a promis de se conduire bien
(Max promised you to behave himself nicely), etc.

Hence, there is a need to describe person-number agreement between infinitive verbs and their subjects, and there is a formal basis for doing so in languages other than Portuguese.

If we now compare the Portuguese examples with (3) and (4), the difference between them should no longer appear so great. In French as in Portuguese, infinitive verbs have a subject: with certain main verbs in Portuguese, this fact must be marked, whereas in French, it is in general unmarked. The difference is then limited to the morphological position of the person-number marks and possibly to special lexical conditions in Portuguese.

Let us suppose we operate in a theoretical framework where syntactic features such as gender, number and person are attached to verbs, and where segmentation rules intro-

duce the corresponding morphemes. In such a framework, the difference between French and Portuguese will be circumscribed to the conditions under which segmentation rules operate. One would presumably regard this as a surface, hence, superficial difference.

On the other hand, one may, with good reasons, deny any resemblance between the two types of infinitive forms we have just reduced to the same formalism; this is tantamount to claiming that Portuguese has a specific construction which is foreign to other languages.

It is thus clear that attitudes toward phenomena may vary considerably, according to the theoretical options one adopts.[20]

3.1.4 *SENTENTIAL PHRASES IN* QUE

The following example suggests that comparisons between structures should be performed at a more abstract level than that of observed sentences.

In French, we find two related types of sentential phrases:

- *que S* in subject and direct complement positions,
- *ce que S* in indirect positions, mainly with *Prep =: à, de*.

But there exist reasons to reconstruct *ce* in direct positions (M. Gross 1968). For example, passive forms such as

Max est étonné de (ta présence, ce que tu partes)
(Max is surprised (at your presence, that you leave))

are related to the active forms obtained by undoing the Passive transformation:

Ta présence étonne Max
(Your presence surprises Max)

(1) **Ce que tu partes étonne Max*

511

but instead of the subject *ce que S* in (1), we observe

Que tu partes étonne Max
(That you leave surprise Max)

Thus, assuming that *ce* is a determiner for all *que S* structures regularizes the set of sentences of French. Notice that in other contexts, *ce* is a demonstrative adjective, that is, a definite determiner.

In Spanish, corresponding sentential forms are observed with definite articles:

el que S, lo de que S

We then see that the regularization needed for French becomes essential when French and Spanish sentential constructions are compared.

It should also be emphasized that the French unacceptable—hence, abstract—form (1) is highly concrete, since it differs from an observed form by an element appearing elsewhere, in similar syntactic positions. Note that the same observations apply to both French and Spanish. In contrast, if we had used an abstract symbol for the complementizer (e.g. *Comp*) in order to describe *que S* phrases, comparisons would have had an unverifiable character: for example, the same symbol (*Comp*) would have applied to Italian where no determiner is found in front of a *che S* phrase.

3.2 COMPARING LEXICONS

3.2.1 *INDIVIDUAL LEXICAL ITEMS*

One could compare individual entries. For example, given two languages, one might compare two corresponding verbs together with their syntactic properties, as they are represented in lexicon-grammars. However, the selection of such pairs runs into various difficulties.

Semantically equivalent verbs should be compared; in that case, the translation process between the two languages raises various questions. The comparison of the meanings of two verbs is the comparison of the distributions of the nouns they accept in subject and complement positions. Too few distributional descriptions of verbs are available so far, so that synonymy remains a fairly unreliable intuitive concept. But even if elaborate descriptions were available, it would not always be clear which parts of the distributions should be distinguished for comparison. Consider for example the verb *to eat*, whose direct complement is usually a noun of solid food. However, a few nouns of liquids are accepted:

(Fr) Max mange de la soupe
(Max eats soup)

(It) Max mangia la zuppa

(Po) O Max come a sopa

(Sp) Max come sopa

Should one consider the few nouns of liquids that are compatible with *eat* as a special subclass that may or may not translate as nouns of solid food do, or should one consider that they are an integral part of the distribution of the direct complement? So far as the subjects of *manger* — 'to eat' are concerned, we note that German has separate verbs for animal and human subjects. Could it be the case that the Romance languages are also sensitive to this difference, but in a way that is not superficially apparent?[21]

Should one compare verbs that are morphologically equivalent, i.e., which clearly have the same historical root? However, it often happens that the current meaning has drifted away from the older meaning; e.g., consider the Italian verb *divertir* — 'to amuse', in French *amuser*.

French has also the verb *divertir*, with the same meaning as *amuser*; however, whereas It. *divertire* and Fr. *amuser* belong to standard speech, Fr. *divertir* is literary. Should one then associate the pair *divertire-divertir* or the pair *divertire-amuser*? That is, should we neglect the difference in level, even though we know that in French the literary level is often associated with the use of the word in Classical French?

In some cases, which are not too rare in the Romance family, both conditions are met: morphologically close verbs appear to be synonymous across languages. Then, the comparison of the syntactic properties of such forms should be revealing, even at the elementary level of checking the similarities and differences of such properties. However, such comparisons may have to be limited to words which clearly correspond to each other.

3.2.2 *CLASSES OF LEXICAL ITEMS*

A more classical question can be raised about nouns. Consider the following nouns which all mean *bill* or *check* in English:

(Fr) le compte but more commonly *l'addition*
(It) il conto
(Po) a conta
(Sp) la cuenta

Clearly, these words are morphologically related, reflecting their common Latin origin. But a question arises: Why did French and Italian choose the masculine gender, whereas Portuguese and Spanish adopted the feminine? Such questions have been dealt with in traditional grammars, but we think that only a systematic cooperative enterprise between linguists can lead to a sufficient accumulation of standardized data (here, probably data about nominaliza-

tions of verbs), from which general principles of evolution of genders may possibly be discovered.

The difficulty of comparing single lexical items or single rules may be avoided in part by comparing classes of words, especially in the cases where natural classes have been observed.

For example, in the lexicon-grammar of French, a set of about 150 verbs of saying are taken to constitute a natural class for the following reasons:

- they enter into the syntactic paradigm defined by the structures (1) to (4) in 2.1,
- all the verbs of the paradigm share the meaning of *dire*-- 'to tell',[22]
- when other syntactic properties of these verbs are described, it is found that many are shared by most members of the class (M. Gross 1975: table 9).

If such a situation is found in different languages, a comparison between whole classes should be more meaningful than comparisons between single members of the class. Thus, in Italian, the situation is the same as in French: we have the paradigm

Max ha detto a Jo che è stato maiato
Max ha detto a Jo di essere stato malato
Max ha detto a Jo che se ne vada
Max ha detto a Jo di andarsene

In Portuguese and in Spanish, reduced infinitive forms are not as common as in French or in Italian:

Max (le) dijo a Jo que había estado enfermo
**Max (le) dijo a Jo (de) haber estado enfermo*
Max (le) dijo a Jo que se fuese
**Max (le) dijo a Jo (de) irse*

(Po) Eles anunciaratam ao Max que a Maria obtido una bolsa

Eles anunciaratam ao Max terem obtido una bolsa

Another example of a natural class is the class of verbs of motion. Consider the syntactic paradigm that defines this class:

N_0 *V* V^0*-inf Compl =:*

Max (court, entre,...) parler à Bob
(Max (runs, enters,...) to talk to Bob)

{ -- *Où* N_0 *V-t-il?*
{ -- V^0*-inf Compl =:*

{ --*Où Max (court, entre,...)-t-il?*
{ --*Parler à Bob*

We see that the infinitive complement is locative, in the sense that it answers a question in *où* —'where'.

These two conditions are fulfilled by more than 100 verbs in French. All are related semantically to motion, and they share a significant number of other syntactic properties. For example, the infinitive complement cannot be derived from a full sentential complement *que S*, whereas this type of derivation is justified for most infinitive complements in French. The infinitive locative form cannot contain a negation, nor other modalities such as *pouvoir*—'can' or *devoir*—'must', etc.

The determination of a complete list of verbs of motion runs into a difficulty. Some verbs come very close to the definition given above, but their acceptability in those structures that define the verbs of motion is mixed. Also, semantically, they are somewhat removed from the concept of motion. For example, we have

Max (s'étend, s'asseoit, s'effondre) sur son lit lire son journal
(Max (lays down, sits down, collapses) on his bed to read his newspaper)

Notice that we have improved the acceptability of our examples by introducing the locative complement *sur son lit*—'on his bed'; without it, the sentences are dubious.[23]

In these examples, the infinitive form does not answer the question in *où*. However, the combination locative-infinitive does answer this question:

-- Où Max s'étend-il?

-- Lire son journal

--Où Max s'étend-il?

--Sur son lit lire son journal

Other syntactic properties of these verbs are similar to those of the proper verbs of motion (e.g. negation, modals). Semantically, these verbs do not correspond to a motion of their subject, but rather to a change of posture of the subject. The locative complement is a complement of destination, as with verbs of motion.

In Italian (Elia 1983; Elia, D'Agostino, Martinelli 1981), a very similar situation has been observed. The only difference is that the infinitive complements which answer the question in *dove*--where must take the preposition *a*, and not "zero" as in French. On the whole, the syntactic properties of *V-inf Compl* are the same as in French, and the set of verbs is about the same, with their special meaning of motion. Verbs of change of posture are more easily accepted in the infinitive construction:

Max va a vedere Bob

Max si distende sul letto a leggere il giornale

In Spanish (Lamiroy 1981, 1983), the situation is closely parallel to that of French and Italian, both lexically and syntactically. The infinitive form that answers

the question *a donde* — 'where' is preceded by *Prep =: a*. But Spanish is different in an important respect from both French and Italian: there exists a full sentential complement *a que S* which can be considered as the source of the infinitive form:

Max viene a que le paguen

Max viene a pagar

Spanish thus becomes a model of regularity on two accounts:

(i) unlike French and Italian, the infinitive form has an attested source,

(ii) unlike French but like Italian, the infinitive form takes *Prep =: a*, which can be considered a locative marker (Guillet, Leclère, Boons 1983). Notice that *a* disappears both in French and Italian from the interrogative pronoun, but not in Spanish *(a donde)*.

In Portuguese (Macedo to appear), the facts seem different. Verbs accept constructions similar to those of French, Italian and Spanish, but there is no analogous lexical extension:[24]

O Max (vai, vem) vir o Bob

Comparisons can also be conducted across languages among individual lexical items, and we will do it for the example of *marcher* (cf. 3.2.4). However, it seems to us that comparisons between classes of constructions are much more significant. Thus, if only the few lexical items appearing in our examples above are considered, no difference can be detected between Portuguese and the other Romance languages. Also, what we have called natural classes appears to involve productive processes, which by definition cannot be detected at the level of individual lexical items. Hence, if only comparisons between single lexical items are carried out, no difference would be perceived between words undergoing general processes, and isolated

elements accidentally preserved through time.

Consider the following example of corresponding classes in French and in Spanish. A class of verbs is defined in French by the following structure:

N_0 *V* N_1; N_0 =: *que S* + etc., N_1 =: *Nhum*

These verbs can have a sentential phrase as well as various types of nouns for subject. Their direct object must be strictly human, and there is no second indirect object (*à, de*) N_2; more than 500 verbs meet these conditions. A scan of the list reveals a remarkable semantic homogeneity: most of the verbs are psychological verbs. Also, they have common syntactic properties (M. Gross 1975: table 4, Ruwet 1972). Examples are

Que Bob parte (amuse, ennuie, dégoûte intéresse, etc.) Max

(That Bob leaves (amuses, annoys, disgusts, interests, etc.) Max)

In Spanish a similar list has been defined (Subirats 1983) by the following structure:

N_0 *V a* N_1; N_0 =: *que S* + etc., N_1 =: *Nhum*

and the complement *a* N_1 is an indirect object, not a direct one (i.e., it pronominalizes to *le*). The list of verbs is largely the same in both languages, and it translates word for word.

An obvious question can now be raised: why is there such a difference (or change) of structure between French and Spanish, whereas in Italian (Elia 1977) and in Portuguese (Macedo 1979), these same classes are structurally

similar to those of French, i.e., verbs have a direct object?

In order to answer this question, we examine the French verbs more closely and note that the condition N_1 =: *Nhum* is not strictly respected. With many verbs, some non-human nouns may be found in N_1, but these are few and are of a special type, viz., inalienable possessions of humans:

(1d) *Cette histoire réjouit le coeur de Bob*
(This story warms the cockles of Bob's heart)

(2d) *Ces nouvelles remontent le moral de Bob*
(This news boosts Bob's morale)

Often, such sentences are synonymous with others in which N_1 has been restructured:

(1a) *Cette histoire réjouit le coeur à Bob*

(2a) *Ces nouvelles remontent le moral à Bob*

The *à Nhum* complement can then be pronominalized to *Ppv =: lui*, exactly like an indirect object in *à*. The result is often more natural than the source:

Cette histoire lui réjouit le coeur

Ces nouvelles lui remontent le moral

Sentences (1a) and (1d), (2a) and (2d) are respectively synonymous with the following sentences where N_1 is human:

(1) *Cette histoire réjouit Bob*
(This story rejoices Bob)

(2) *Ces nouvelles remontent Bob*
(This news boosts Bob)

There may be theoretical problems with the exact formulation of the relations between these sentences, but there

is no doubt as to the existence of the relations:

(1d) = (1a) = (1)

(2d) = (2a) = (2)

In particular, a zeroing rule must affect the special objects.

In Spanish, the situation is about the same, except that the structure with special nouns is *direct*. We have

(3) *Esto alegra a Bob*

(3d) *Esto alegra el corazón de Bob*

(3a) *Esto alegra el corazón a Bob*

We are now in a position to propose an explanation for the difference observed between Spanish and the other Romance languages. In *all* Romance languages, the structure with a direct object will be considered basic:

N_0 V N_1 =: (1d),(2d),(3d) with special N_1

These basic forms are close to frozen sentences (class CAN: table 2, 2.2). The zeroing operation applies to the (d)-forms in French, Italian and Portuguese:

(1d) = (1)

where it applies to the (a)-forms in Spanish:

(3a) = (3)

One of the questions brought up by this proposal concerns the separation between diachronic and synchronic occurrences of the phenomenon. For certain verbs that do not seem to accept any restricted N_1s, the zeroing operation may have applied at an earlier stage of the language; with other verbs, there are productive metaphors that seem quite recent, and to which zeroing does not (yet) apply.

Because the facts concern large classes of verbs, we

are confident that this analysis will account for most of the differences observed. Moreover, its concrete nature allows an effective verification, for example by going through texts for occurrences of unobserved restricted *N*'s.

We feel that the accumulation of such cases is the only path by which we can arrive at principles of comparison, or at least hope to discover significant lexical and syntactic features in families of languages. In this way, we may be able to reduce the arbitrariness of the criteria so far invoked to constitute families of languages.

3.2.3 *FROZEN EXPRESSIONS*

Frozen expressions seem to have anecdotal origins. As a rule, the particular relation between the words composing a frozen expression and its meaning is arbitrary. Suppose now that a frozen expression translates word for word in some other language, i.e., the translation is a frozen expression, and that the two frozen expressions have the same meaning. Under such peculiar conditions, the double correspondence between the two expressions requires an explanation.

Preliminary explorations have shown that such correspondences are not at all rare; examples for varied syntactic structures are given in table 4.

Some explanations for the similarities are available. For example, the origin of certain expressions can be found in the Bible, as proverbs or metaphors, others may have been propagated by translations of universal authors, such as Cervantes, Dante, Molière and Shakespeare. However, a question remains: given the arbitrariness, that is, the exceptional character of these expressions, they would appear to be a supplementary burden on the learner in each language. Then their stability through generations of speakers of separate languages is as yet inexplicable.

Table 4

C1: N_0 *V* C_1 =:

N_0 *perdre la tête*
N_0 *perdere la testa*
N_0 *perder la cabeza*

CAN: N_0 *V* $(C(de,a)N)_1$ =:

N_0 *éclairer les idées (de, à) Max*
N_0 *schiarire le idee (di, a) Max*
N_0 *aclatar las ideas (de, a) Max*

CDN: N_0 *V* $(C\ de\ N)_1$ =:

N_0 *poser la première pierre de cet édifice*
N_0 *porre la prima pietra di questo edificio*
N_0 *poner la primera piedra de este edificio*

CP1: N_0 *V Prep* C_1 =:

N_0 *se noyer dans un verre d'eau*
N_0 *perdersi in un bicchiere d'acqua*
N_0 *ahogarse en un vaso de agua*

CPN: N_0 *V Prep* $(C\ de\ N)_1$ =:

N_0 *tomber dans les bras de Max*
N_0 *cadere nelle braccia di Max*
N_0 *caer en brazos de Max*

C1PN: N_0 *V* C_1 *Prep* N_2 =:

N_0 *se casser la tête pour résoudre ce problème*
N_0 *rompersi la testa per risolvere il problema*
N_0 *romperse la cabeza por solucionar el problema*

Table 4 (cont'd)

CNP2: N_0 *V* N_1 *Prep* C_2 =:

N_0 abandonner Max à son sort
N_0 abbandonare Max alla sua sorte
N_0 abandonar Max a su suerte

C1P2: N_0 *V* C_1 *Prep* C_2 =:

N_0 ne comprendre rien à rien
N_0 non capire niente di niente
N_0 no entender nada de nada

CO: C_0 *V Compl* =:

Le sang monter à la tête (de, à) Max
Il sangue salire a la testa (di, a) Max
Subirse la sangue a la cabeza (de, a) Max

CADV: N_0 *V Adv* =:

Max aller loin
Max andare lontano
Max llegar lejos

CV: N_0 *V V-inf Compl* =:

N_0 laisser à désirer
N_0 lasciare a desiderare
N_0 dejar que desear

3.2.4 *EXCEPTIONS*

A similar area of comparative studies involves exceptional lexical items.

Suppose that one has found reasons to consider a given lexical item of a language as an exception to certain syntactic rules of that language. Suppose that in another language the same exception is found, i.e., corresponding rules have an exception that translates into the exception of the other language. Such a situation requires an explanation. Notice that an explanation using the notions of borrowing or of inheritance from a common ancestor, say Latin, cannot be maintained in most cases, because of the nature of an exception. By definition, an exception is an unexpected situation, that is, a situation in which rules that have a wide range of application should apply but don't. Hence, one would expect that at the time of the borrowing and/or of learning, the rules would introduce an analogical pressure that would act so as to regularize the exception. Such a hypothesis is all the more plausible that rules and exceptions are unconscious for the speaker. But from numerous observations, we conclude that teaching and external corrections, which sometimes maintain exceptions, are virtually without effect on the use of most exceptional constructions.

Consider, for example, the French verb *marcher*— 'to walk!. Intuitively, it appears to be a verb of motion, not very different semantically from *courir*—'to run'. However, the verb *marcher* does not enter into the infinitive construction discussed in 3.2.2:

**Max marche voir Bob*

whereas *courir* does:

Max court voir Bob

525

Thus, *marcher* has to be considered an exception with respect to our observations on verbs of motion. It is interesting to note that in Italian, *camminare* is a corresponding exception:

**Max cammina a vedere Bob*

Max corre a vedere Bob

The exceptional nature of this verb appears to be linked to the meaning of its locative complements. With verbs of motion, locative complements are regularly interpreted as the destination or the origin of the motion. With *marcher* and *camminare*, the locative complement cannot have this interpretation, and may only have the meaning of the scene where the motion takes place:

Max marche à son bureau

Max cammina al suo ufficio

It seems to us that this coincidence of facts cannot be accidental, and so requires a common explanation. Notice however, that French *cheminer*, which is the morphological correspondent of *camminare*, is a semi-exception in this respect, since it is accepted in the infinitive construction, but its locative complement is "scenic," and not destination:

Max chemine retrouver Ida

**Max chemine à son bureau*

Another such example of exceptional properties might be found in the analysis of the French verb *partager* — 'to share'. This verb enters into the sentence forms:

(1F) *Max a partagé une tarte avec Bob*
(Max shared a pie with Bob)

(2F) *Max a partagé une tarte entre Bob et Jo*
(Max shared a pie between Bob and Jo)

In (1) *Max* received a share of the pie, but not in (2). The difference in the interpretation of the relations between the arguments of this verb forces us to consider that we are dealing with two different verbs *partager*. In fact, a rule relating (1) and (2) would seem to be *ad hoc*, because its application would be limited practically to this verb only. In this respect, French *diviser* — 'to divide' enters into (2), but not into (1).

The situation in Italian seems to be similar: *dividere* has the properties of *partager:*

(1I) *Max ha diviso il dolce con Bob*

(2I) *Max ha diviso il dolce tra Bob e Jo*

Notice that Italian *condividire* is not similar to French *diviser* (and English *to share*), since it enters only into (1I). It is not known whether *dividere* is the only verb in Italian with both structures, but the present classification of Italian verbs already suggests that these verbs are quite rare.[25]

4. CONCLUSION

We have outlined what may seem an ambitious program of study for the Romance languages. But it should be noticed that none of the demands we make and none of the questions this program raises are really new, for practically all the components of the program have been discussed at various times. The novelty lies rather in the systematization of the descriptions. Once we have come to realize the need of working with as complete a picture as possible of the lexical conditions that must be met by syntactic rules, we are confronted with the problem of constructing lexicon-grammars for the Romance languages.

Once such lexicon-grammars are available, comparisons can be made much more efficiently and meaningfully. How-

ever, the essential motivation for these comparisons should not be forgotten. Each Romance language has had an evolution of its own over more than one thousand years. When compared today, these languages share many lexical and syntactic features, whereas they differ in many others. The following two related questions can then be raised: Why have languages preserved so many properties and why have they diverged on others? A reasonably complete knowledge of the present (synchronic) situation can provide important clues and valuable hypotheses about the evolution of Romance languages.

Reconstructing the numerous steps through which the Romance languages have proceeded seems to be the only way to confirm hypotheses. The coherence and abundance of Romance data are such that the hope of reaching a global and precise understanding of this evolution is well founded.

In turn, the systems of evolution that will eventually be constructed for Romance languages could serve as the starting models for the description of the history of other groups of languages.

ANNEX

Intransitive Verbs

Table 35 R

(From Boons, Guillet, Leclère 1976a)

N_0 = N hum	N_0 = N pc	N_0 = N-hum	N_0 = N nr	N_0 = V Ω	N_0 = V-n	N_0 = N plur obl		N_0 V	N_0 est V-ant	N_0 est V pp	N^0 pc lui V	N_0 de N^0 pc		N_1 = N hum	N_0 V Prép N pc de N_1	N_0 lui V Prép N^1 pc	N_0 V Prép N pc de N_0	N_0 V Prép N^0 pc	N_1 = N-hum	N_1 = le fait que P	Ppv	N_1 = V-n	N_1 = N plur obl	N_0 V de N^0 pc Prép N_1	N hum V sur ce point	N_0 est V-ant Prép N_1	il V N_0 Ω
+	−	−	−	−	−	−	dîner	+	−	−	−	−	de	−	+	−	−	−	+	−	−	+	−	−	−	−	+
−	+	+	−	−	−	−	éclater	+	−	+	+	−	en	−	−	−	−	−	+	−	−	−	+	−	−	−	+
+	−	+	−	−	−	−	écoper	+	−	−	−	−	de	+	−	−	−	−	+	−	+	−	−	−	−	−	−
+	+	−	−	−	−	−	s'emparer	−	−	−	−	−	de	+	+	−	+	−	+	−	+	−	−	−	−	−	−
+	−	−	−	−	−	−	s'emporter	+	−	+	−	−	contre	+	+	−	+	−	+	+	−	−	−	−	+	−	−
+	+	+	+	+	−	−	l'emporter	+	−	−	−	−	sur	+	+	−	+	−	+	+	−	−	−	−	+	−	−
+	−	+	+	+	−	−	évoluer	+	−	+	−	−	vers	−	−	−	−	−	+	−	−	−	−	−	+	−	+
−	+	+	−	−	−	−	exploser	+	−	+	+	−	en	−	−	−	−	−	+	−	−	−	+	−	−	−	+
+	−	−	−	−	−	−	se fendre	+	−	−	−	−	de	−	−	−	−	−	+	−	+	−	−	−	−	−	−
+	−	−	−	−	−	−	festoyer	+	−	−	−	−	de	−	+	−	−	−	+	−	+	−	−	−	−	−	−
+	−	−	−	−	−	−	finasser	+	−	−	−	−	avec	+	+	−	+	−	+	−	−	−	−	−	+	−	−
+	+	+	+	+	−	−	en finir	+	−	−	−	−	avec	+	+	−	+	−	+	+	−	−	−	−	+	−	−
+	−	−	−	−	−	−	fondre	−	−	−	−	−	en	−	−	−	−	−	+	−	−	−	+	−	−	−	−
+	+	+	−	−	−	−	fondre	−	−	−	−	−	sur	+	+	+	+	−	+	−	−	−	−	−	−	−	+
+	−	−	−	−	−	−	fuir	+	−	−	−	−	devant	+	−	−	−	−	+	+	−	−	−	−	−	−	+

Verbs with Sentential Complements

(From M. Gross 1975)

529

jacasser	intimer	interdire	inspirer	insinuer	inoculer	infliger	indiquer	inculquer	imputer	imposer	hurler	hululer	hoqueter	hocher la tête		
+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	N_{hum}	Sujet
–	+	+	+	+	+	+	+	+	+	+	–	–	–	–	N_{nr}	
–	+	+	+	+	+	+	+	+	+	+	–	–	–	–	*le fait Qu P*	
–	–	–	–	–	–	–	–	–	+	+	–	–	–	–	V^1 Ω	
–	+	+	+	+	+	+	+	+	+	+	–	–	–	–	V^2 Ω	
+	–	–	–	–	–	–	–	–	–	–	+	+	+	–	N_0 *est V-e*	Compl. direct
+	–	–	+	–	–	–	–	–	–	–	+	+	+	+	N_0 V	
+	–	–	–	–	–	–	–	–	–	–	+	+	–	–	N_0 *pousse Dét V-ment*	Vop
–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	N_0 *pousse Dét V-⊕*	
+	–	–	–	–	–	–	–	–	–	–	+	+	+	–	N_0 V *contre* N_{hum}	
+	–	–	–	–	–	–	–	–	–	–	+	+	+	–	N_0 V *après* N_{hum}	
+	+	+	+	+	–	–	+	+	–	+	+	+	+	–	N_0 V N_1	
+	–	–	–	+	–	–	+	+	+	–	+	+	+	+	*que P*	Compléments
–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	V^0 Ω	
+	–	–	–	+	–	–	+	+	+	–	+	+	+	–	*Aux* V^0 Ω	
–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	*de* V^0 Ω	
–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	(N_1) *(être Adj)*	
–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	*(être Adj) (Qu P)*	
+	–	–	–	+	–	–	+	+	–	–	+	+	+	+	*que* N_0 V *(être Adj)*	
+	–	–	–	+	–	–	+	+	–	–	+	+	+	+	(N_1) *(Adj)*	
–	–	+	–	–	–	–	–	–	–	+	–	–	–	–	*(Adj) (Qu P)*	
–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	*que* N_0 V *(Adj)*	
–	–	+	–	–	–	+	–	–	–	+	–	–	–	–	*Nég, interrog. = subj*	
+	–	–	–	+	–	–	+	+	–	–	+	+	+	+	*impératif = subj*	
–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	*que Psubj*	
+	+	+	+	+	+	+	+	+	–	+	+	+	+	+	*de* V^0 Ω	
–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	*de* V^2 Ω	
+	+	+	+	+	+	+	+	+	–	+	+	+	+	+	*si P ou si P*	
+	–	–	–	–	–	–	+	–	–	–	+	+	+	–	*ce (ci + la)*	Pron.
+	+	+	+	+	+	+	+	+	+	+	+	+	+	–	*ppv*	
–	–	+	–	–	–	+	+	–	–	+	–	–	–	–	N_{hum}	Noms
+	+	+	+	+	+	+	+	+	+	+	+	+	+	–	N_{-hum}	
–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	*le fait Qu P*	
–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	N_0 V *à* N_2	Comp. indirects
+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	N_{hum}	
+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	*ppv*	
–	–	–	–	–	–	–	–	–	+	+	–	–	–	–	N_{-hum}	
–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	*ppv*	
–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	*là*	
–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	[extrap] [passif]	
+	–	–	–	–	–	–	+	–	–	–	+	+	+	+	*de* N_{hum}	
–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	*à quoi* N_0 V N_1	
–	+	–	–	+	–	–	+	–	–	–	+	+	+	–	N_0 V *à* N_2 : "P"	
–	+	+	–	+	–	–	+	+	–	+	+	+	+	–	"P" . V N_0 *à* N_2	
+	+	+	+	+	+	+	+	+	+	+	+	+	+	–	[passif par]	
–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	[passif de]	

530

Verbs With Sentential Complements

(From Elia 1978)

Sujet				Italien	N_0 V contro Num	N_0 V(presso + con) Num	Complément direct						Compl. Indirect				Passivo	Français	Différences de classement
							complétives						Noms						
Num	Nnr	il Fatto Che F	V^1 Ω				che Fcong	Fcong	V^0 Ω	di V^0 Ω	(N_1) (V^1-inf)	se F o se F	cio	Ppv	Num	N-um	il Fatto Che F	? Che F	N_0 V di Num	N_0 V da N_2	dal Fatto Che F	a Num			
+	−	−	−	ABOLIRE	−	−	+	−	−	−	−	−	+	+	+	+	+	+	−	−	−	−	+	ABOLIR	
+	+	−	−	ABORRIRE	−	−	+	−	−	−	−	−	+	+	+	+	+	+	−	−	−	−	+	ABHORRER	
+	−	+	−	ABROGARE	−	−	+	−	−	−	−	−	+	+	+	+	−	+	−	−	−	−	+	ABROGER	
+	−	−	−	ACCETTARE	−	−	+	−	−	+	−	−	+	+	+	+	+	+	−	−	−	−	+	ACCEPTER	
+	−	−	−	ACCOGLIERE AVVM	−	−	+	−	−	+	−	−	+	+	+	+	+	−	−	−	−	−	+	ACCUEILLIR ADVM	12
+	−	−	−	ADORARE	−	−	+	−	+	−	−	−	+	+	+	+	+	+	+	−	−	−	+	ADORER	
+	−	−	−	AGGIORNARE AVVF	−	−	+	−	−	−	−	−	+	+	+	+	+	+	−	−	−	−	+	AJOURNER	
+	−	−	−	AGOGNARE	−	−	+	−	−	+	−	−	+	+	+	+	+	+	−	−	−	−	+	CONVOITER	
+	−	−	−	AMBIRE	−	−	+	−	+	−	−	−	+	+	+	+	+	+	+	−	−	−	+	AMBITIONNER	
+	−	−	−	AMMIRARE	−	−	+	−	−	+	−	−	+	+	+	+	+	+	−	−	−	−	+	ADMIRER	
+	+	−	−	ANNULLARE	−	−	+	−	−	+	−	−	+	+	+	+	+	+	−	−	−	−	+	ANNULER	
+	−	−	−	ANTICIPARE AVVF	−	−	+	−	−	+	−	−	+	+	+	+	+	+	−	−	−	−	+	ANTICIPER ADVFUT	
+	−	−	−	APPETIRE	−	−	+	−	−	−	−	−	+	+	+	+	+	+	+	−	−	−	+	DESIRER	
+	−	−	−	APPROVARE	−	−	+	−	−	−	−	−	+	+	+	+	+	+	−	+	−	−	+	APPROUVER	12
+	−	−	−	ASPETTARE	−	−	+	−	−	+	−	−	+	+	+	+	+	+	−	+	−	−	+	ATTENDRE	
+	−	−	−	ASPETTARSI	−	−	+	−	−	+	−	−	+	+	+	+	+	+	−	+	−	−	+	S'ATTENDRE	7
+	−	−	−	ATTENDERE	−	−	+	−	−	−	−	+	+	+	+	+	+	+	−	+	−	−	+	ATTENDRE	
+	−	−	−	ATTENDERSI	−	−	+	−	−	−	−	+	+	+	+	+	+	+	−	+	−	−	+	S'ATTENDRE	7
+	+	−	−	AUTORIZZARE	−	−	+	−	−	−	−	−	+	+	+	+	+	+	−	−	−	−	+	AUTORISER	
+	−	−	+	AUSPICARSI	−	+	+	−	−	−	−	−	+	+	+	+	+	+	−	−	−	−	−	SOUHAITER	
+	−	−	−	AUGURARSI	−	−	+	−	−	−	−	−	+	+	+	+	+	+	−	−	−	−	−	SOUHAITER	
−	−	−	−	AVERE L'IMPRESSIONE	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	AVOIR L'IDÉE	
+	−	−	−	AVERE A CUORE	−	−	+	−	−	−	−	−	+	+	+	+	+	+	−	−	−	−	+	AVOIR A COEUR	

531

Frozen Adverbs *(M. Gross)*

Table: PD

SUJETS N_0 =: Nhum N_0 =: N-hum	VERBES	ADVERBES FIGES		
+ –	VENIR	DANS	LA	"PERIODE"
+ –	PARTIR	SUR	CE	–
+ –	DEMONTRER N A N	PAR	L'	ABSURDE
+ –	PARTIR	DANS	L'	AFFIRMATIVE
+ –	DIRE NA N	EN	L'	AIR
+ –	TRICHER	A	POSS-O	AISE
+ –	ARRETER S	A	L'	ALLER
+ –	VENIR	A	TOUTE	ALLURE
+ –	ESPERER N	DE	TOUTE POSS-O	AME
+ –	ARRANGER N	A	L'	AMIABLE
+ –	GAGNER N	A	L'	ARRACHE
+ –	VENIR	CONTRE	TOUTE	ATTENTE
+ –	PARTIR	A	L'	AUBE
+ +	VENIR	PAR	L'	AUTOBUS
+ –	PAYER N	A	L'	AVANCE
+ –	CONSULTER N	A	L'	AVENIR
+ –	CONSULTER N	DANS	L'	AVENIR
+ –	CHOISIR N	A	L'	AVEUGLETTE
+ –	DISCUTER	–	TOUT	AZIMUT
+ –	BOIRE N	AVANT	LA	BAGARRE
+ –	SPECULER	A	LA	BAISSE
+ –	PARLER	–	TOUT	BAS
+ –	TRICHER	DE	PLUS	BELLE
+ –	FONCER	A	TOUTE	BERZINGUE
+ –	AGIR	A	LE	BESOIN
+ –	CUIRE N	A	LE	BEURRE
+ –	FONCER	A	TOUTE	BITURE
+ –	CUIRE N	A	LE	BOIS
+ –	ACCEPTER N	EN	TOUTE	BONNE FOI
+ –	RIRE	DE	TOUTE POSS-O	BOUCHE
+ –	LUTTER	JUSQU'A	LE	BOUT
+ –	CUIRE N	SUR	LA	BRAISE
+ –	FONCER	A	TOUTE	BRIDE
+ –	CUIRE N	A	LA	BROCHE
+ +	VENIR	PAR	LE	BUS
+ –	CUIRE N	A	LE	BUTAGAZ
+ –	CUIRE N	A	LE	BUTANE
+ –	DORMIR	EN	TOUT	CAS
+ –	CUIRE N	SOUS	LA	CENDRE
+ –	REMBOURSER N	A	LE	CENTUPLE

Frozen Expressions

(M. Gross 1982)

Table C1

N0 - Hum	N0 - Hum		N_0 V Passif		Npc* Autres déterminants	
+	-	CONNAITRE	-	+	LE	- - COUP
+	-	CONNAITRE	-	-	POSS-Ø	- - DOULEUR
+	-	CONNAITRE	-	+	LE	- - TRUC
+	-	NE CONNAITRE PAS	-	-	POSS-Ø	- - BONHEUR
+	-	NE CONNAITRE QUE	-	-	-	- - CA
+	-	CONSERVER	-	-	POSS-Ø	- - CHEMISE
+	-	SE CONTEMPLER	-	-	LE	+ - NOMBRIL
+	-	COUPER	-	+	DET	- - CORDON OMBILICAL
+	+	DEBLOQUER	-	+	det	- - SITUATION
+	-	DETENIR	+	-	LA	- - VERITE
-	+	DISTILLER	-	+	LE	- + VENIN
+	+	DOMINER	-	+	LE	- - LOT
+	-	DRESSER	-	+	POSS-Ø	- + BATTERIES
+	-	ENDOSSER	-	+	LE	- - HARNOIS
+	+	ENFONCER	-	+	LE	- - CLOU
+	-	ETRE . N PAS	-	-	UNE	- - LUMIERE
+	-	ETRE . N PAS	-	-	-	- - MANCHOT
-	+	ETRE . N PAS	-	-	LA	- - MORT
+	-	ETRE . S DIT	-	-	-	- - TOUT
+	-	FAIRE	-	-	UN	- - BRIN DE TOILETTE
+	-	FAIRE	-	-	-	- - GRISE MINE
+	-	FAIRE	-	-	-	- - HARA-KIRI
-	+	FAIRE	-	-	-	- - JURISPRUDENCE
+	-	FAIRE	-	+	UNE	- + MINUTE DE SILENCE
-	+	FAIRE	-	-	-	- + NOMBRE
+	-	FAIRE	-	+	DET	- - OPERATION PORTE OUVERTE
+	-	FAIRE	-	-	DU	- - QUARANTE CINQ FILLETTE
+	-	FAIRE	-	-	-	- - TAPIS
+	-	FAIRE	-	-	-	- - TINTIN
+	-	FAIRE ENTENDRE	-	-	POSS-Ø	- - VOIX
+	-	FAIRE PASSER	-	-	DET	- - ENFANT
+	-	FAIRE SAUTER	-	-	DET	- - ENFANT
-	+	FERMER	-	+	POSS-Ø	- - PORTES
+	-	FLETRIR	-	+	DET	- - CRIME
+	-	FORCER	-	-	LA	- + CHANCE
+	-	FORMER	-	+	LE	- - CARRE
+	-	FORMER	-	+	DET	- - NUMERO
+	-	FORMER	-	+	DET	- - NUMERO DE TELEPHONE
+	-	FORMER	-	+	LES	- - RANGS
+	+	FRANCHIR	-	+	DET	- - CAP

Frozen Expressions of English

(P. Freckleton and P. Machonis)

Table C1P2

NN OO == NN H UH MU M				NOVC1				PASSIF	
++	SEPARATE	THE	MEN	-	FROM	THE	BOYS	-	57
+-	SET		FOOT	-	ACROSS	THE	THRESHOLD	-	58
+-	SET	POSS-O	HAND	-	TO	THE	PLOUGH	-	47
+-	SET	POSS-O	HOUSE	-	IN		ORDER	+	50
++	SET	N'S	MIND	-	AT		EASE	+	110
++	SET	N'S	MIND	-	AT		REST	+	109
++	SET	N'S	TEETH	-	ON		EDGE	+	96
++	SET	THE	THAMES	-	ON		FIRE	-	59
++	SET	THE	WORLD	-	ON		FIRE	-	60
+-	SHED	A	TEAR	+	FOR		NELSON	+	61
+-	STICK	POSS-O	FOOT	-	IN	POSS-O	MOUTH	-	101
+-	TAKE	THE	BIT	-	IN	POSS-O	MOUTH	+	5
++	TAKE	THE	BREAD	-	OUT	N'S	MOUTH	-	83
+-	TAKE	THE	BULL	-	BY	THE	HORNS	-	64
+-	TAKE		IT	+	ON	THE	CHIN	-	62
+-	TAKE	THE	LAW	-	INTO	POSS-O OWN	HANDS	-	30
+-	TAKE	POSS-O	LIFE	-	IN	POSS-O	HANDS	-	63
+-	TAKE	A	LOAD	-	OFF	POSS-O	FEET	-	91
+-	TAKE	THE	LORD'S NAME	-	IN		VAIN	+	93
-+	TAKE	A	LOT	-	OF		DOING	-	111
+-	TAKE	THE	NAME OF THE LORD	-	IN		VAIN	+	94
++	TAKE	A	TURN	-	FOR	THE	BETTER	-	97
++	TAKE	A	TURN	-	FOR	THE	WORSE	-	98
+-	TAKE	THE	WORDS	-	OUT OF	N'S	MOUTH	-	92
++	TELL		TALES	+	OUT OF		SCHOOL	-	65
+-	TELL!		IT	-	TO	THE	MARINES	-	66
+-	TELL!		IT	-	TO		SWEENEY	-	67
+-	THROW		CAUTION	-	TO	THE	WINDS	+	69
+-	THROW		DISCRETION	-	TO	THE	WINDS	+	70
+-	THROW	POSS-O	HAT	-	IN(TO)	THE	RING	+	103
+-	TOSS	POSS-O	HAT	-	IN(TO)	THE	RING	+	104
++	TURN	THE	SCALE(S)	+	IN	N'S	FAVOUR	+	112
+-	WASH	POSS-O	(DIRTY) LINEN	-	IN		PUBLIC	-	102
+-	WEAR	POSS-O	HEART	-	ON	POSS-O	SLEEVE	-	74
+-	WORK	POSS-O	FINGERS	-	TO	THE	BONE	-	75

NOTES

1) E.R.A. No 247 of the Centre National de la Recherche Scientifique. I am indebted to A. Elia and M. Salkoff for their remarks. They have helped me to greatly improve a preliminary version of the manuscript.

2) Boons, Guillet and Leclère 1976a, 1976b; Giry-Schneider 1978; M. Gross 1975; G. Gross 1982; Labelle 1974; Meunier 1977, 1981; de Négroni 1978.

3) Elia 1979; Elia, D'Agostino, Martinelli 1981.

4) Malaca 1981; Macedo 1981.

5) Subirats 1981. There also exist studies on Korean verbs Hong 1982, on Arabic verbs: Chad 1981, on English verbs: Salkoff 1983 and on Malagasy: Rabenilaina 1979.

6) The symbol *Compl* has nothing to do with the notion of predicate, whether traditional or formal: *VP*, *PredP*. We are not aware of evidence that would justify the use of the notion of predicate in a grammar, that is, in a theory.

7) Actually, if theoreticians took upon themselves the burden of verifying their proposals with respect to systematic data, that is, with respect to a lexicon-grammar, we can safely predict that the rate of production of new theoretical devices would be considerably diminished.

8) In object positions, the situation is different.

9) Quite the contrary. Also, from a not strictly syntactic point of view (dialectal, philosophical or other), this question might be interesting in its present empirical state.

10) This important distinction appears to have been lost in recent years by generative grammarians.

11) This principle holds mainly for simple sentences. Long complex sentences, such as those frequently found in texts, should be either simplified or abandoned.

12) Or equivalently, for which too many incompatible definitions have been proposed.

13) The sentences with direct objects

Max cherche (Léa, ce livre)

are accepted, but with a different verb: *chercher*—'to look for'.

14) Notice that the interrogative pronoun is the non-human form *quoi*, while *a priori*, *président* is categorized as human.

15) This association is exactly what we call an element of the lexicon-grammar.

16) In the same way, we count about 1,200 "regular" adverbs, that is, of the form Adjective-*ment* (Adjective-*ly* in English), whereas frozen adverbials already number 3,000.

17) The place of English among the Romance languages is not settled in our view.

18) These languages contain dialectal examples of the general phenomenon observed in Spanish.

19) And closer to that of other Indo-European languages with infinitive forms analogous to that of French (Macedo 1981, Raposo 1973, Rouveret 1978).

20) Of course, diachronic considerations should also intervene. Studies on Old French by Chaurand 1983 and on Latin by Lemaire 1979 are suggestive in this respect. Cf. also Lightner 1983 for some morphological aspects.

21) A possible trace of the difference could be the use of the suffix *-vore* (-vorous) in sentences such as

Le chat est carnivore
(The cat is carnivorous)

22) No verb having the four syntactic properties has a different meaning, but there exist verbs of saying that enter into a different syntactic paradigm: *parler à quelqu'un de quelque chose* — 'to talk to someone about something'.

23) One problem about the introduced locative complement is its scope: Does it bear on the main verb, on the infinitive verb, or on some combination of both (Danlos, to appear)?

24) The data appear to be close to those of English, where *to go* seems to be the only verb that has some of the properties observed in the Romance languages

Bob went to see Max

In German (Caroli 1983 and in Dutch (Lamiroy 1983), the situation is of a different nature. Corresponding to the verbs of motion of French or Spanish, one observes combinations of a small number of verbs and a variety of locative particles.

25) The situation seems different in Spanish where (1) and (2) translate with two different, but related, verbs:

(1S) *Max ha compartido el pastel con Bob*
(2S) *Max ha repartido el pastel entre Bob y Jo*

REFERENCES

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

———— 1976b. *La structure des phrases simples en français. Classes de constructions transitives.* Rapport de recherches du L.A.D.L. No. 6. Paris: LADL, Université Paris 7, 85 pp., tables et index: 58 pp.

Caroli, Folker. 1983. "La notion de directionnalité dans les verbes de déplacement en allemand." In *Actes du 1er Colloque Européen sur la grammaire et le lexique comparés des langues romanes.* Amsterdam: John Benjamins B.V.

Chad, Mohamed. 1981. *Les constructions intransitives de l'arabe classique.* Thèse de 3ème cycle. Paris: LADL, Université Paris 7, 250 pp.

Chaurand, Jacques. 1983. "Les verbes supports en ancien français : 'donor dans les oeuvres de Chrétien de Troyes," *Lingvisticae Investigationes VII:2*. Amsterdam: John Benjamins B.V.

Chomsky, Noam. 1957. *Syntactic Structures*. Det Hagen: Mouton, 116 pp.

———— 1963. "Formal Properties of Grammars." In *Handbook of Mathematical Psychology*, D. Luce, E. Bush, E. Galanter, eds. New York: John Wiley and Sons Inc. pp. 323-418.

———— 1965. *Aspects of the Theory of Syntax*. Cambridge, Mass: MIT Press, 251 pp.

Danlos, Laurence. 1980. *Représentation d'informations linguistiques: constructions N être Prép X*, Thèse de 3ème cycle. Paris: LADL, Université Paris 7, 353 pp.

———— To appear. "Les compléments associés aux infinitives locatives," Paris: L.A.D.L.

Elia, Annibale, Maurizio Martinelli, Emilio d'Agostino. 1981. *Lessico e strutture sintattiche. Introduzione alla sintassi del verbo Italiano*. Naples: Liguori, 430 pp.

Elia, Annibale. 1977. "Une classe de verbes italiens qui acceptent une complétive sujet et un objet direct 'humain'." *Travaux de Recherches du LADL*. Paris: LADL Université Paris 7.

———— 1979. *Lexique des constructions complétives en italien. Verbes à un complément*. Thèse de 3ème cycle. Paris: LADL, Université Paris 7.

———— To appear. "L'infinitive en a dans les verbes locatifs intransitif italiens."

Freckleton, Peter. 1983. *Une comparaison des expressions de l'anglais et du français*. Thèse de 3ècycle in preparation. Paris: Université Paris 7, L.A.D.L.

Giry-Schneider, Jacqueline. 1978. *Les nominalisations en français. L'opérateur faire dans le lexique*. Geneva: Droz, 353 pp. Tables: 61 pp.

Greenberg, Joseph. 1966. *Universals of Language*. 2nd edition. Cambridge, Mass.: MIT Press.

Gross, Gaston. 1982. "Un cas de constructions inverses: donner et recevoir." *Linguisticae Investigationes VI:2*. Amsterdam: John Benjamins B.V., pp. 1-44.

Gross, Maurice. 1968. *Grammaire transformationnelle du français, Syntaxe du verbe*. Paris: Larousse, 183 pp.

———— 1969. "Remarques sur la notion d'objet direct en français. *Langue Francaise 1*. Paris: Larousse, pp. 63-73.

———— 1975. *Méthodes en syntaxe*. Paris: Hermann, 414 pp.

————1981. "Les bases empiriques de la notion de predicat sémantique." *Langages 63*. Paris: Larousse, pp. 7-52.

Guillet, Alain, Christian Leclère, Jean-Paul Boons. To appear. *Les verbes à compléments transitifs locatifs, La structure des phrases simples en français II*. Geneva: Droz.

Harris, Zellig S. 1976. *Notes du cours de syntaxe*. Paris: Seuil, 237 pp.

537

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

———— 1964. The Elementary Transformations, Transformations and Discourse Analysis Papers 54, in Harris, Zellig S. 1970 *Papers in Structural and Transformational Linguistics*. Dordrecht: Reidel.

———— 1968. *Mathematical Structures of Language*. New York: Wiley, 230 pp.

———— 1982. *A Grammar of English on Mathematical Principles*. New York: Wiley Interscience, 429 pp.

Hong, Chai-song. 1982a. "Remarques sur la construction des verbes de mouvement en coréen." *Lingvisticae Investigationes VI:1*. Amsterdam: John Benjamins B.V., pp. 83-144.

Kayne, Richard S. 1977. *Syntaxe du français, Le cycle transformationnel*. Paris: Le Seuil, 446 pp.

Klima, Edward S. 1962. Correspondence at the grammatical Level, Mechanical Translation Group. Cambridge, Mass.: *Research Laboratory of Electronics G.P.R.*, MIT.

Kuroda, Sige-Yuki. 1965. "Causative Forms in Japanese," *Foundations of Language* 1, pp. 30-50.

Labelle, Jacques. 1974. *Etude de constructions avec opérateur AVOIR* (nominalisations et extensions). Thesis. Paris: LADL.

Lamiroy, Béatrice. 1981. "Les prépositions a et para devant l'infinitif complément du verbe de mouvement en espagnol." *Lingvisticae Investigationes V:1*. Amsterdam: John Benjamins B.V., pp. 75-90.

———— 1983. *Les verbes de mouvement en français et en espagnol, Etude de syntaxe comparée de leurs infinitives*. Amsterdam: John Benjamins B.V.

Lemaire, Maryvonne. 1979. *Recherches sur la préfixation et sur la suffixation des verbes latins*. Thèse de 3ème cycle. Paris: LADL, Université Paris 7.

Lightner, Theodor M. 1983. *Introduction to English Derivational Morphology*. Amsterdam: John Benjamins B.V.

Macedo Oliveira, Maria Elisa de. To appear. "La syntaxe des verbes de mouvement en portugais."

———— 1979. *Syntaxe des verbes psychologiques du portugais*. Thèse de Doctorat. Paris: Université Paris 7, LADL.

———— 1981. "Les structures de complétives du portugais." *Lingvisticae Investigationes V:1*. Amsterdam: John Benjamins B.V., pp. 91-136.

Machonis, Peter A. 1982. "Transformations of English Idioms." Read at the Winter L.S.A. meeting.

Malaca Casteleiro, João. 1981. *Sintaxe transformacional do adjectivo, Regência das construções completivas*. Lisboa: Instituto Nacional de Investigaçaõ Cientifica, 561 pp.

Meillet, Antoine. 1952. *Linguistique historique et linguistique générale*. Paris: Klincksieck.

Meunier, Annie. 1977. "Sur les bases syntaxiques de la morphologie dérivationnelle." *Lingvisticae Investigationes I:2*. Amsterdam: John Benjamins B.V., pp 287-331.

Meunier, Annie. 1981. *Nominalisations d'adjectifs par verbes supports.* Thèse de 3ème cycle. Paris: LADL, Université Paris 7, 282 pp.

Négroni-Peyre, Dominique de. 1978. "Nominalisations par ETRE EN et réflexivation." *Lingvisticae Investigationes II:1.* Amsterdam: John Benjamins B.V., pp. 127-163.

Nördahl, Helge. 1969. *Les systèmes du subjonctif corrélatif.* Bergen et Oslo: Universitetsforlaget, 217 pp.

Rabenilaina, Roger-Bruno. "La neutralité de la diathèse en Malgache." *Lingvisticae Investigationes III:2.* Amsterdam: John Benjamins B.V., pp. 295-321.

Raposo, Eduardo Paiva. 1973. *Estudos sobre o Infinitivo em Portuguès.* Thesis. Lisbon: Faculté des Lettres.

Rouveret, Alain. 1980. "Sur la notion de proposition finie. Gouvernement et inversion." *Langages 60.* Paris: Larousse, pp. 75-107.

Ruwet, Nicholas. 1972. "A propos d'une classe de verbes 'psychologiques'." In *Théorie syntaxique et syntaxe du français.* Paris: Le Seuil, pp. 181-251.

Salkoff, Morris. 1983. "Bees are swarming in the garden: A systematic synchronic study of productivity." *Language* 59.288-346. Baltimore: The Waverly Press.

Subirats, Carlos. 1981. "Sur les applications théoriques de la construction des classes syntaxiques de verbes." *Linguisticae Investigationes V:2.* Amsterdam: John Benjamins B.V., pp. 329-374.

———— 1983. "Les verbes psychologiques à complétive sujet en espagnol." In *Actes du ler Colloque européen sur la grammaire et le lexique comparés des langues romanes,* ed. by A. Guillet & N. La Fauci. Amsterdam: John Benjamins B.V.

Treig, Thomas. 1977. "Complétives en allemand. Classification." *Rapport de Recherche LADL.* Paris: LADL P7, pp. 39-203.

Zwicky, Arnold. 1971. "In a Manner of Speaking." *Linguistic Inquiry II:2.* Cambridge, Mass.: MIT Press, pp. 223-233.