8 Constructing Lexicon-Grammars

MAURICE GROSS

1. INTRODUCTION

We present the state of construction of a lexicon-grammar of French, and, at the same time, principles applicable to other languages. The system which has been constructed can be considered as a specific linguistic theory systematically applied to a linguistic material that has a significant coverage.

The theory of lexicon-grammar is founded on the following axiom:

The linguistic unit of meaning is the elementary sentence.

The significance of this principle has to be made more precise. One obvious point is that the term 'elementary sentence' needs to be defined. We consider this requirement an empirical question more than a theoretical one, which can only be solved by the very process of the construction of the lexicon-grammar.

A consequence of our axiom is that words are not elementary units of meaning; this statement must be justified. That words are not elementary units of meaning is obvious for compound words, that is, combinations of words that are not compositional. Compound nouns for example have non-compositional meaning by definition. They are much more numerous than simple nouns in the lexicon of any language. They constitute the technical vocabularies (several millions of words). That sentences are elementary units of meaning is obvious in the case of idiomatic sentences. The importance of this remark stems from the observation that idiomatic sentences are more numerous than ordinary sentences. It is also clear that verbs cannot be considered without their subject and possible object(s), hence they are units. The same is true for *be Adjective* forms, and also for predicative

This work has been partly financed by the 'Programme de Recherches Coordonnées' in Computational Linguistics of the Ministry of Research and Higher Education. I am indebted to Jean Baudot, David Berlinski, Sige-Yuki Kuroda, Ted Lightner, Morris Salkoff, and Marco Schützenberger for important suggestions.

nouns and adverbs, although in an indirect way that will be discussed below. Converging observations led to this theoretical stand:

1. More syntactic properties of sentences than usually thought depend on the main verb. For example, determiners are mostly described as being constrained by their noun. However, they can be restricted by the *verb* to which their noun is attached:

Bob wants some beer. **Bob loves some beer.*
Bob hunts the wild goose. **Bob hunts a wild goose.*

In the same way, adjectival modifiers of nouns may have to be selected by the verb, as can be seen from the variations of meaning in the pairs:

Bob is building a future mansion. = *Bob is building a mansion.*
Bob is eating a future cake. ≠ *Bob is eating a cake.*

2. The systematic description of French verbs (or simple sentences) has shown that no two verbs have the same syntactic properties (Gross 1975; Boons *et al.* 1976*a*, 1976*b*; Guillet and Leclère 1991). As a consequence, verbs have to be described individually, and not in terms of intensional classes.

3. The proportion in the lexicon of idiomatic sentences, of metaphorical and technical sentences with no compositional meaning, is very high. All these sentences or sentence types have anecdotal origins. Hence, they have to be described individually, that is, without reference to classes of lexical combinations or of interpretation patterns.

4. The large number of verb-complement combinations that cannot be qualified in terms of semantic (i.e. selectional) restrictions leads to the notion of support verbs. The variety of support verbs also implies individual syntactic description of nouns.

The three preceding types of sentences can be clearly distinguished in European languages. We will base their presentation on French, but there is growing evidence that this division is more general.

1.1 *Some requirements for a grammar*

The notion 'rule of grammar' plays an important role in many approaches to syntax. But no specific notion of 'rule' seems to have been agreed on by the community of linguists, for the theoretical aspects of the notion change quickly, and its variability from author to author constitutes an opaque screen in the evaluation of possible advances. One elementary aspect of the notion 'rule' has always been neglected by linguists: the determination of complete domains of

application for proposed rules. We will explore this feature of grammars here. The notion of rule is hardly distinguishable from the broad concept of syntactic theory. Currently, rules or syntactic theories are being elaborated for a small number of examples of a given phenomenon, and then the occasional discovery of new examples and counter-examples functions either to reinforce or to invalidate theories. Such an activity is considered to be the normal scientific activity of syntacticians. This approach is based on the assumption that the syntactic phenomena being formalized are general or important, which, as we will see, need not be the case. As a matter of fact, such an assumption could only be founded on a good knowledge of the range of application of rules.

There have been many academic discussions about the status of counter-examples: is a counter-example a crucial experiment that invalidates a rule, or should it be neglected as uninteresting, that is, as not affecting the generality of the proposed rule or theory? We think that the basis of such a question is deeply rooted in the nature of language and, mainly, in its historical development. Linguistic forms have been invented and accumulated in the memory of man since prehistoric times. Hence, languages contain a large variety of devices, often hidden by certain general superficial features.[1] Some of these devices are productive, namely, rules govern them and extend them to open parts of the lexicon (Salkoff 1983). Other devices are fossils left over from formerly productive rules. This classical point of view offers a way of explaining many phenomena without resorting to abstract pseudo-mathematical mechanisms. In particular, the historical position provides a falsifiable way of dealing with exceptions; but then explanation will belong to the diachronic description of the language, in principle distinct from the synchronic activity of constructing a grammar. Thus, we consider that the nature of rules and exceptions can be properly dealt with only if a fundamental step in the design of a rule is taken, namely, only if one attempts to determine the lexical range of application rules. We will discuss the principles that we followed in the course of the construction of our lexicon-grammar, the coverage reached, and the consequences for syntactic theory. An example such as the distribution of prepositions with respect to verbs clearly illustrates some inadequacies of the field. Consider the following phenomenon: given a verb, it may have complements, with or without a preposition: *Max ate his egg* has no preposition (or has 'zero' preposition), whereas the sentences:

Max looked at his egg. *Max concentrates on his egg.*

[1] This viewpoint is largely Jesperson's (1961–5) and Lightner's (1983).

have prepositions *at* and *on*. Under certain conditions, the preposition is specific to the verb; thus, the following sequences are not accepted as sentences:

**Max ate about his egg.*
**Max concentrates at his egg.*
**Max dreamed for his egg.*

The same phenomenon exists in other European languages such as French, German, Italian, and is not, of course, limited to Indo-European languages. However, it has never occurred to any linguist that a reasonably complete list of verbs with the preposition(s)[2] introducing their object(s) could be informative, and even useful. Although, in the learning of a second language, one of the commonest and most persistent mistakes heard is misuse of prepositions, pedagogues have never asked for nor tried to construct, once and for all, such a list, not even for a language as widely taught as English. Dictionaries sometimes contain information relevant to this problem, but only rarely are their examples intended to demonstrate any special syntactic use of verb entries. A useful description, however, should have, facing each entry, all its constructions. Then, when a user of such a description hesitates about the use of a preposition Prep with a verb V, he can check whether Prep is in the list associated with V. In the absence of this Prep, the description guarantees it cannot be used with V.[3] Such a dictionary would provide a *characterization* of the use of Preps with Vs, a description that coincides with the aims of grammar as made explicit by structural linguists (Harris 1946, 1951) and adopted in generative linguistics.[4] First, we will expose the general assumptions under which we proceeded when we attempted to characterize large segments of French grammar.

[2] A precise definition of object complement is based on the set of prepositions involved and on the shape of the associated questions. Object complements will yield interrogative pronouns *who-qui* and *what-quoi*, but will not correspond to questions *where-où*, *when-quand*, *how-comment*, etc. Thus the two dialogues:

On what did Max concentrate? —*On his egg.*
* *Where did Max concentrate?* —*On his egg.*

indicate that TO CONCENTRATE is construed here with an object.

[3] Several dictionaries have been designed to provide systematic syntactic information: Hornby and Cowie (1989); Caput and Caput (1969); Bonnard *et al.* (1970); Summers *et al.* (1987); Benson *et al.* (1986). They have not yet reached the state where this procedure could be applied.

[4] A generative grammar is essentially a characteristic function defined on the free monoid generated by the vocabulary of the language. It should separate the sequences of words accepted as sentences from the others (Chomsky 1957; Chomsky and Schützenberger 1963).

1.2 *Transformational argumentation*

Relations between lexicon and grammar are often discussed in transformational grammar. Arguments relating grammatical rules and the size of the lexicon were used quite early. Thus, structural grammars were shown to be inadequate, in the sense that they did not relate active sentences and their corresponding passive forms. Let us examine the types of arguments involved in the demonstration of the existence of such relations. The following types of arguments can be developed in support of the active–passive relation:

1. *The qualitative argument*. One aspect of the rule that we call qualitative is well known: in a phrase structure grammar the highly complex distributions of nouns occurring in subject and object positions of a verb have to be described, once for the verb (e.g. TO BITE) and a second time for its passive form (e.g. TO BE BITTEN BY). In contrast, a single description is sufficient for both these forms in a transformational grammar. This type of argument leads to a qualitative change in the theory.

2. *The quantitative argument*. In English there are numerous verbs that enter into active–passive pairs (several thousand); hence the saving just described is amplified by this numerical factor. This quantitative effect on phrase structure grammars has never been seriously discussed, since no figures about the size of the lexicon are available in this context. We insist on the independence of the two arguments (1) and (2). Argument (1) would still be valid if there were only one passivizable verb, say, TO BITE. In that case, argument (2) would not carry much weight. We will distinguish a third type of argument:

3. *The argument of variety*. When a relation (i.e. roughly a transformation) such as passive applies to several lexical items, the existence of differences between these items may be important. Suppose that passive were to apply only to a set of verbs such as TO LIKE, TO LOVE, TO HATE, TO ADORE, etc. Since these verbs are quite similar semantically, that is distributionally, one could argue that passive applies in one and the same mode to all these verbs. But it happens that passive applies to a wide range of verbs: verbs with very different distributions in their subject and object (e.g. TO AMUSE, TO BITE, TO CORROBORATE), with a variety of possible second complements (e.g. TO GIVE, TO PUT, TO PREFER), etc. This feature of passive shows that its action is formal, namely independent of the lexical content of the noun phrases involved. This formal nature reinforces the demonstration that a syntactic, i.e. combinatorial, relation holds

between an active and a passive form.[5] We will make use of an analogous type of argument (4) when we discuss frozen forms.

4. *The argument of idiomatic invariance*. In many syntactic studies, these types of arguments are invoked, but in general they are not kept separate. For example, when the quantitative argument (2) is used, it is always in an implicit way, as with passive or the verbal forms that do not accept a negation. Linguists leave it to intuition that the rule operates on large numbers of lexical items. No systematic attempt has ever been made to determine the lexical range of a single rule such as passive. We will describe an attempt in French to construct complete rules, that is, an attempt to state both the rules and the lexical conditions under which they operate. In particular, the lack of numerical data associated with phenomena relating the lexicon to the grammar will be shown to have direct consequences on certain theoretical choices.

2. CONSTRUCTION OF A LEXICON-GRAMMAR

The terminology lexicon-grammar is motivated by the preceding discussion. We constantly face a theoretical choice: either we select a certain model of grammar and we duplicate lexical entries, thus increasing the size of the lexicon, or else we introduce a rule (e.g. passive), saving entries, but changing the shape of the grammar, and preventing the lexicon from growing in an undesirable way. Many discussions hinge on the same question: for example, does one introduce raising rules for TO SEEM, for TO BELIEVE, etc., or does one increase the size and structure of the lexicon? Since the size and shape of the lexicon and the size and shape of the grammar are so tightly interconnected, it is impossible to obtain a global view of one component if one ignores the other. Thus, the main motivation for constructing a lexicon-grammar is to determine systematically the rules of a language and where they apply. As has been commonly noted, rules involve sets of words. Often, a set of words determined by the range of application of a rule is called a syntactic class. But, so far in linguistics, syntactic classes have been defined only intensionally, never in extension. Moreover, it is not always clear that the examples given to demonstrate the existence of a rule show that the classes involved are not empty. In certain cases, reviewing what seemed to be

[5] Harris (1968) uses an interesting consequence of the formal character of the rule: when a rule applies to a semantically deviant form, the result has the same semantic deviance. Arguments (1), (3), and (4) (cf. s. 3.2) are based on this invariant character of syntactic rules with respect to meaning.

a fact when just a few sentences were concerned showed that acceptability judgements could not be reproduced over a larger part of the lexicon (Gross 1979). In such cases the rule and its classes have to be rejected as artefacts. This happens regularly with semantic notions.[6] A point of departure is argument (1) discussed in section 1.

The goal of general linguistics is to determine the universal characteristics of human language, irrespective of the accidents present in any particular language. The domain of activity of linguists thus consists in designing theoretical programs that will achieve that goal. In contrast, the study of particular languages is confined to activities called English grammar, Turkish grammar, etc., and is usually performed by specialists whose interests are mainly pedagogical. Since grammars of particular languages have never been particularly useful to their native speakers, the task of building them has not benefited from any academic or intellectual status that would have put grammarians in contact with standard scientific procedures. For the same reason, the field of descriptive grammar has not been organized in such a way that data could be accumulated in some consistent way. It is only recently, with the growing interest in computational linguistics, that such a research programme has been considered. We are describing the outcome of such a programme: we have shown with French that it is possible to build detailed grammars of particular languages. Building and then comparing particular grammars may hint at universal features that will then possess a minimal empirical basis. The particular grammar we have designed has nothing to do with a generative grammar. We will first give an overall picture of the lexicon-grammar of French in its present state; we will see how it supports our point of view, and will then discuss problems encountered and possible solutions for them.

3. A LEXICON-GRAMMAR OF FRENCH

Except for a few rules, linked more to the spelling of words than to syntax (e.g. the conjugation of strong verbs), there is not a single example where regular cases and exceptions have been compiled. This is rather puzzling since, in many cases, going through a dictionary and checking every relevant item leads to a reasonably complete picture of

[6] As a matter of fact, the separation advocated by structural and transformational linguists between form and meaning is better stated in terms of reproducible vs. non-reproducible intuitions, independently of their purported nature. The subjective character of most concepts of meaning can often be detected by varying the examples, and variations necessarily occur when one operates methodically on a lexicon.

the effect of a rule. We found this situation intolerable from both a theoretical and an empirical point of view. To remedy it for French, we used the procedure already suggested:

1. We emphasized the description of simple sentences, that is, sentences with one verb and its object complements. We extended the set of syntactic properties that depend on the verb by re-examining already known facts and by introducing new constraints found in the course of the work. At present, we deal with an order of magnitude of 400 properties of this type.[7]

2. These properties have been verified and represented for a set of about 12,000 verbs. This set has been limited to non-technical verbs, so that they could be submitted to the acceptability judgements of a team of six to ten linguists.[8] For many obsolete or technical verbs it was not possible to rely on intuitions of acceptability, and these verbs were not taken into account. We will see that there are several ways of counting verbs and properties, but changes in the counting conventions do not affect the overall picture.

3.1 *The lexicon-grammar of verbs*

3.1.1 *The main data*

The available description of simple French sentences can be presented in the shape of a matrix. Each column corresponds to one property, that is, to one sentence form. Each row contains one lexical entry. At the intersection of a row and a column, a '+' sign indicates that the corresponding verb enters into the corresponding construction, a '—' sign, that it does not (cf. Table 8.1, class 4). A row of the matrix represents the syntactic paradigm of the corresponding verb. Each member of the paradigm is noted as on the following example:[9]

N_0 *s'étonner de ce que* S

It contains an explicit lexical item (here a verbal form *étonner*)[10] and grammatical constants (*se, de, ce, que*); the only variables allowed are noun phrases (N_is) and sentences (S).

[7] Introducing nominalization and adjectivization relations would raise this number to about 600.

[8] For a discussion of certain experimental problems raised by acceptability, cf. Boons (1974).

[9] Notations are transparent; indices on noun phrases N_i are as follows: i = 0 for subjects, i = 1, 2 for first and second complements. The variable W stands for any sequence of complement(s). The sign =: is used to specify the lexical or structural content of a form; the sign = indicates a syntactic relation between two sentence forms, roughly, a transformation in Z. S. Harris's sense.

[10] We do not use any class symbol for verbs in such forms, except for abbreviatory purposes. We will see below that the notion of class has become quite complex, following an empirical observation on the syntactic diversity of verbs.

TABLE 8.1 *Sample of a table of verbs with sentential subjects*

Subject								Adjective				Direct object										
N_{hum}	N_{nr}	*le fait* Qu P	V' W		V 'concrete'	N_0 V	a — ant	a — able	a — eux	a — (E + at) eur	N_{hum}	N_{-hum}	*le fait* Qu P	N_1 *se* V *de ce* Qu P	N_1 *se* V *auprès de* N_{hum} *de ce* Qu P	N_1 *est* Vpp *de ce* Qu P	|passive *par*|	|passive *de*|	N_0 V N_1 *contre* N_{hum}	
+	+	+	+	irriter	+	+	+	−	−	−	+	+	−	+	+	+	+	−	+	
+	+	+	+	juguler	+	+	−	−	−	−	+	+	−	−	−	−	+	−	−	
−	+	+	+	lénifier	−	+	+	−	−	−	+	+	−	−	−	−	+	−	−	
+	+	+	+	léser	−	+	−	−	−	−	+	+	−	−	−	−	+	−	−	
+	+	+	+	lessiver	+	+	+	−	−	−	+	+	−	−	−	−	+	−	−	
+	+	+	+	leurrer	−	+	−	−	−	−	+	−	−	−	−	−	+	−	−	
+	+	+	+	magnétiser	+	+	+	−	−	−	+	+	−	−	−	−	+	−	−	
+	+	+	+	magnifier	+	+	+	−	−	−	+	+	+	−	−	−	+	−	−	
+	+	+	+	marquer	+	+	+	−	−	−	+	+	−	−	−	−	+	−	−	
+	+	+	+	martyriser	−	+	+	−	−	−	+	−	−	−	−	−	+	−	−	
+	+	+	+	mater	−	+	−	−	−	−	+	−	−	−	−	+	−	−	−	
+	+	+	+	mécontenter	−	+	−	−	−	−	+	−	−	+	−	+	+	−	−	
−	+	+	+	méduser	−	+	+	−	−	−	+	−	−	−	−	−	+	−	−	
+	+	+	+	ménager	−	−	−	−	−	−	+	+	−	−	−	−	+	+	−	
+	+	+	+	métamorphoser	+	+	−	−	−	−	+	+	−	−	−	−	+	−	−	
+	+	+	+	meurtrir	+	+	+	−	−	−	+	+	−	−	−	−	+	−	−	
+	+	+	+	miner	+	+	−	−	−	−	+	+	−	−	−	−	+	−	−	
+	+	+	+	mithridatiser	+	+	+	−	−	−	+	−	−	−	−	−	+	−	−	
+	+	+	+	mobiliser	−	+	−	−	−	+	+	+	−	−	−	−	+	−	+	
+	+	+	+	modérer	−	+	−	−	−	+	+	+	−	−	−	−	+	−	−	

Note: The structures represented are N_0 V N_1, with N_0 −: Qu P (Gross 1975: table 1)

Syntactic relations may hold between columns and thus induce a structure on the columns of the matrix. They are equivalence relations and are noted by the sign '=' (Harris 1968). We refer to the classes of structures they determine as syntactic paradigms. For example, we will write for the unexpected relation:[11]

1. N_0 *irriter* N_1 = N_1 *s'irriter de* N_0 *auprès de* N_2

The elements of a syntactic paradigm constitute our fundamental units, both syntactic and lexical. In the case of a verb this position is natural: it is hard to imagine a verb without its subject and main complement(s), but we will see that this solution can be extended to the other parts of speech in a natural way (section 3.3). As a consequence, the basic element of meaning we deal with is the simple sentence.

Precise delimitation of syntactic properties is necessary not only for theoretical reasons but for practical purposes as well. Thus, the data accumulated by different specialists must be consistent, their judgement must be made compatible. In this respect, the size of the enterprise has demanded special precautions. In particular, highly operational criteria had to be devised. Such requirements led us to base the classification largely on prepositional properties. The use of object prepositions with verbs, as summarily described in the introduction, is traditionally limited in French to the three prepositions 'zero', *à*, and *de*, which are considered as fundamental. We adopted the postulate that these prepositions were important, and that this importance had an empirical basis, as shown in the praxis of generations of grammarians and schoolteachers. Since these prepositions introduce a variety of complements including many types of adverbials, we limited ourselves to the following notion of object:

- a direct object answers the question pronouns *qui* (who) or *que* (what),
- an indirect object answers the question pronouns *à* (*qui*, *quoi*), *de* (*qui*, *quoi*) (Prep (*who*, *whom*, *what*)).

As a consequence of this preliminary choice, the structures to be examined a priori are the following:

- intransitive: N_0 V =: *Max dort.* (Max sleeps.)

[11] The pair of French sentences:

Cette décision irrite Bob. — (This decision irritates Bob.)
Bob s'irrite de cette décision auprès de son chef. (Bob tells his boss that this decision irritates him.)

present an important difference in structure and meaning. Such a difference often constitutes a reason to exclude the relation from the set of transformations of a language. However, investigating the lexicon-grammar of French has led us to establish a relation between them (Gross 1975).

- one object: N_0 V N_1 —: *Max voit Bob.*
  (Max sees Bob.)
  N_0 V *à* N_1 —: *Max obéit à Jo.*
  (Max obeys Jo.)
  N_0 V *de* N_1 —: *Max rit de Jo.*
  (Max laughs at Jo.)
- two objects: N_0 V N_1 N_2 —: *Ils ont élu Max président.*
  (They elected Max president.)
  N_0 V N_1 *à* N_2 —: *Max dit ceci à Bob.*
  (Max says this to Bob.)
  N_0 V N_1 *de* N_2 —: *Max reçoit un lit de Jo.*
  (Max receives a bed from Jo.)

N_0 V *à* N_1 *à* N_2, N_0 V *à* N_1 *de* N_2, N_0 V *de* N_1 *de* N_2, etc.

- three objects: etc.

Notice that a variable number of object(s) can be omitted, that is, verbs may enter into shorter (sub)structures. These possibilities are treated as properties of the longest structure.[12]

This description appears straightforward, and one may wonder why it has never been performed by lexicographers, for example. In fact, many delicate problems arise at each step of the work. Some questions seem anecdotal, others are likely to affect the whole structure of the classification. This is the case with the following example.

Consider the structure:

$$N_0 \text{ V } N_1 \textit{ de } N_2$$

It has one direct object and one indirect object in one of the main prepositions of French. Some verbs that enter into this structure also enter into a related construction which involves locative complements:

2. N_0 V N_1 *de* N_2 — N_0 V N_2 Loc N_1
   (*a*) *Max a chargé le camion de caisses.*
   (Max loaded the truck with boxes.)
   — (*b*) *Max a chargé des caisses* (*dans* + *sur*) *le camion*.[13]
   (Max loaded boxes (into + on) the truck.)

The first term of the relation (2*a*) has a direct object and an indirect object in *de*; it enters directly into one of our a priori classes. The second term (2*b*) contains a locative complement introduced by Loc —: *dans* + *sur* + etc., and belongs to the classification as a member of the syntactic paradigm of N_0 *charger* N_1 *de* N_2. But now if we consider sentences such as:

[12] Thus, the structure N_0 V N_1 *à* N_2 has three possible substructures (i.e. columns): N_0 V N_1, N_0 V *à* N_2, N_0 V that the entry verb may enter or not. The '+' and '—' signs then allow the representation of obligatory and optional complements.

[13] The '+' sign here is to be interpreted as an 'or'.

3. *Max a placé de lourdes caisses (dans + sur) le camion.*
   (Max placed heavy boxes (into + on) the truck.)
4. *Max a introduit les caisses dans le camion.*
   (Max introduced the boxes into the truck.)

they have the form N_0 V N_1 Loc N_2. The prepositions Loc are not part of our initial set of object prepositions; as a consequence the complement Loc N_2 must be left out of the description. For *placer* (to place) Loc N_2 is obligatory, but the structure has no place at all in our classification. For *introduire* (to introduce) Loc N_2 is more or less optional; hence this verb would be best described by the structure N_0 V N_1, that is, in a very incomplete way. This undesirable situation has been remedied by using a notion of analogic extension (Harris 1964). We consider (3) and (4) as extensions of the second member of the relation (2), even though (3) and (4) do not enter into the first member of (2):

**Max a (placé + introduit) le camion de caisses.*
(Max (placed + introduced) the truck with boxes.)

Analogic extension is not a purely intuitive notion; for example, the syntactic properties of the semantically analogous Loc N_2 in (2*b*), (3) and (4) are identical. The notion has proved useful in attempting to separate verb complements (or essential complements) from sentence complements, a classificational problem which is not quite settled (Guillet and Leclère 1991; for Italian, cf. d'Agostino 1983 and Martinelli 1984).

Thus, the complements represented in the classification have been generalized beyond the strict and obvious notion of object preceded by 'zero', *à*, or *de*.

3.1.2 *Some results*

Globally, the description of verbs corresponds to a binary matrix of 12,000 verbs by more than 400 properties.[11] In practice, we do not deal with a single matrix of 12,000 by 400. The matrix has been subdivided into about fifty submatrices of about the same size which constitute a systematic classification of French verbal constructions. Already at

[11] Publication of this description is under way. Among the material available so far, we have Gross (1975), which includes the complementizer system bearing on about 3,000 verbs. Boons *et al.* (1976*a*) includes 900 intransitive verbs, (1976*b*) includes 1,300 transitive verbs; Guillet and Leclère (1991) has 1,200 verbs with one direct object and locative second complements. The action of various support verbs on verbs and adjectives (cf. 3.3.) has been described by the same method in Giry-Schneider (1978, 1987); Gross (1989). Most of the material cited in this chapter exists in computer form, and can be provided on request.

this point the study of the classification yields a variety of observations, some global, others specific. We have noted for example:

1. About 300 verbs enter into the structure N_0 V *à* N_1, whereas more than 3,000 enter into N_0 V N_1. We can conclude from this count that the preposition 'zero' has more structural importance than the preposition *à*. Such a result is hardly surprising. However, by providing quantities for the structure of a language, one creates a qualitatively new situation, where statements about the generality of phenomena are no longer derived from vague intuitions acquired through unclear practices, but from an examination of the data.

2. Among the verbs that enter into the structure N_0 V *à* N_1, where N_1 —: N_{hum}, some give rise to a pre-verbal pronoun, others do not:

Max obéit à Bob. — *Max lui obéit.*
Max pense à Bob. — **Max lui pense.*[15]

All that grammars provide is a 'rule' and examples of pronominalized forms and of exceptions. We have determined the lexical extension of the two corresponding sets of verbs: there are about fifty verbs of the type 'obéir' (to obey) and twenty of the type 'penser' (to think). We do not consider that such figures justify the use of the terms rule and exception. We were not able to 'explain' or 'predict' the phenomenon, for example, on the basis of the meaning of the verbs (as is commonly attempted). We are then led to consider that two rules of pronominalization are involved, each applying to a given set of verbs. Other syntactic conditions (e.g. reflexivization) may interfere with the rule of pre-verbal pronominalization.

3. We started the classification on the assumption that the three basic prepositions were criteria of general significance. If we set aside about 1,500 intransitive verbs (since they have no object or take other prepositions, the criteria are not relevant) we find that more than 80 per cent of the verbs have at least one of the three types of objects.[16] The starting hypothesis is thus justified.

4. There are practically no verbs with three objects.[17]

5. We have not found a single verb entering into the structure N_0 V *de* N_1 *de* N_2.[18] Moreover, examples of structures with two prepositional objects are rare. One finds about fifty of them, and then often by stretching acceptability judgements:

[15] With N_1 —: N-hum, the non-human pre-verbal pronoun *y* is accepted by both verbs.

[16] An example of a verb unaffected by the criteria would be COMPTER SUR (to count on), which *must* have a prepositional complement *sur* N (on N): *Max compte sur toi* (Max counts on you), vs the substructure **Max compte* (*Max counts).

[17] One of our best examples of three essential complements is:

Max a parié (*cent francs*) (*avec Luc*) (*que Léa partirait*).

[18] The only candidate, HÉRITER (to inherit), is controversial

N_0 V *à* N_1 *à* N_2 —: *Cet outil sert à Max à cela.*
Cet outil est utile à Max à cela.
(This tool is useful to Max for that.)

N_0 V *à* N_1 *de* N_2 —: *Max parle à Bob de ce lit.*
(Max talks to Bob about this bed.)

We feel we have observed general features of French verbal structures.[19] We may compare them to standard morphological data of a similar formal nature. For example, it has been observed that the initial consonant clusters of English words are restricted (Harris 1951: 153). Rules that state such restrictions are called morpheme structure rules. They are established by studying the words of a dictionary. We are in a position to state analogous rules for the syntactic structure of simple French sentences:

The sequence of object complements in 'zero', à *and* de *is at most of length two; it cannot contain two prepositional objects.*

This rule holds for more than 99 per cent of cases. The significance of such a figure is not obvious. Verbs with two prepositional objects account for the 1 per cent that prevents the rule from being quite general. But are they exceptions? That is, are such structures entirely abnormal, or is it an accident that no more verbs enter into them? An answer to such a question may be looked for in various conceptual areas; it may be linked to a formal property of grammars, or may result from diachronic considerations.

We will also mention two global observations made on the whole classification that have implications for the notion of rule of grammar and for an approach by researchers to learning:

6. The availability of a systematic classification provides a basis for the rigorous definition of classes that intervene in rules. In order to estimate the variety of verbs represented in the classification, the following equivalence relation can be defined:

Two verbs are in the same class, if and only if they have identical rows of + and — signs.

Comparing rows two by two is a simple operation, given the classification in computer form. The result is:

For 12,000 verbs there are about 9,000 classes.

This result was obtained on the basis of roughly 300 standard syntactic properties studied by linguists. But many properties found later to be significant (cf. the examples of section 3.3) have to be added to the

[19] This organization has been applied to other languages than French (e.g. Italian: Elia 1984; d'Agostino 1983; Portuguese: Macedo 1979; Ranchhod 1990; and Spanish: Subirats 1986; for a general bibliography of lexicon-grammar cf. Leclère and Subirats 1991).

columns of the matrix. When one then studies the classes that contain more than one verb, one finds that these new properties further separate the elements of these classes. It is safe to say that the descriptions under way will not contradict the following assertion:

No two verbs have the same set of syntactic properties.

7. When we compared the columns of the matrices, that is the sentence structures, we found that no two columns had identical contents. Since transformations (and other relations) can be made to correspond to pairs of structures, this observation amounts to saying that all transformations have exceptions.

These observations should not be considered as a precise measure of irregularity. This situation does not prevent us from defining 'natural classes', that is, classes where verbs have many common properties and are separated only by a minority of the properties. We will return to this discussion in section 4.3.

Here are some more specific observations. Consider the empirical observation (4) that no French verb has three objects. It leads to the following constraint on grammars: simple sentences are often considered as predicates or functions of several variables. We observed an upper limit of three variables for each function. It is indeed possible to claim that theoreticians have already dealt with this question: distinguishing structures such as verb phrases (VP) and predicate phrases (PredP) is an attempt to separate 'essential' or verb complements from 'circumstantial' or sentence complements. Many criteria have been proposed for this separation, but none of them has ever been applied to a lexicon. We did apply, to about 12,000 verbs, precise and operational criteria in order to determine their objects, and we found no verbs with three objects. We think that this negative result is crucial. It cannot be obtained in any other fashion than by going through the complete lexicon of verbs. The generative approach has not gone further into the matter than have traditional grammarians, who have been aware for a long time of differences among the complements of verbs. Generative studies have sharpened certain criteria and have introduced new ones, but all of these criteria stand as mere *proposals* for a future classification of complements. Only experiments can show whether they are valid, that is, *reproducible* over large numbers of lexical items. No one can prejudge what the result of the classification will be. Once more, this result may not be surprising by itself; what we are pointing out is that, because of the procedure that has been followed, the status of the (known) facts has changed entirely. Instead of being an element of linguists' lore, it has reached the quantitative level.

Here are some observations of a different type. We studied various syntactic properties of sentential complements of the shape *que* S (that S) and infinitive complements of the form V W,[20] that is, sentences without overt subjects, and we noted that:

(*a*) sentential complements have exactly the shape *que* S (that S) in subject and direct object positions; they are found in indicative and/or subjunctive mood, depending on the main verb; in the presence of a (non-zero) preposition, sentential complements have the shape *ce que* S (*Max tient à ce que Bob parte*) vs. *que* S (*Max veut que Bob parte*—Max wants Bob to leave). There is no relation between mood and *ce*, that is, between mood and Prep;

(*b*) the subject of the infinitive verb of V W has a position that varies with the main verb, as in the English examples *Max told Bob to leave*, where *Bob* is *to leave*, and *Max promised Bob to leave*, where the subject of *to leave* is *Max* (Gross 1968; Rosenbaum 1967); the infinitive form can be preceded by the preposition *de* in direct object position, according to the verb: *Max envisage de partir* (Max considers leaving) vs. *Max souhaite partir* (Max wishes to leave).

All these properties figure in columns of the classificatory matrix.

If we set aside the features of (*a*) and (*b*), namely if we compare the distributions of (*ce*) *que* S and (*de*) V W without taking into account the parenthesized variant elements, we see that the verbs governing sentential and infinitive constructions coincide, with three sets of exceptions:

- first, a set of about 100 verbs, mainly such modals as POUVOIR (can) and such aspectuals as FINIR (to stop), take only the infinitive;
- another set of about 200 verbs that take only the infinitive is semantically homogeneous:[21] these verbs all involve a common notion of movement from one point to another (ALLER (to go), COURIR (to run), etc., and causatives of the former ENVOYER (to send), etc.);
- a set of about 200 other verbs takes only *que* S.

All other verbs, that is about 3,000, enter into both constructions. Each complementizer appears then to be largely predictable from the other, with constraints connected to the meaning of the main verbs.

[20] The symbol W stands for a variable sequence of complements.

[21] We determined semantically homogeneous sets of verbs as follows. The syntactically defined set of verbs was presented to readers. Reading the list evoked in their minds one 'idea' that they verbalized by means of terms such as *motion*, *movement from a departure point to a destination*. What gives value to this experiment is the fact that random sets of verbs of comparable sizes do not elicit, at least easily, similar responses. Moreover, for other sets defined syntactically the associate intuitions, when they exist, vary with the syntactic properties.

Meaning appears to play a role in the rule relating the two French complementizers. For example, the bulk of exceptions to reduction of *que* S to the infinitive V W is made up of verbs that trigger an intuition of *reasoning;* this intuition bears on verbs that appear principally in two structures:

5. *Que* S V *de ce que* S —:
   Que Max soit parti vient de ce que Bob l'a insulté.
   (That Max left comes from the fact that Bob insulted him.)

Other such verbs are: DÉCOULER, PROVENIR, RÉSULTER, etc. The second structure is causative with respect to the former:

6. N_0 V *que* S *de ce que* S —:
   Ida conclut que Max est parti de ce que sa voiture n'est plus là.
   (Ida concludes that Max left from the fact that his car is no longer there.)

Similar verbs are: DÉDUIRE, EXTRAPOLER, INDUIRE, INFÉRER, etc.

Only a small number of situations where form and meaning can be related in a clear way are known; no conclusion can be drawn from them, and many points are unclear as to their place in formal grammar, although the analysis by Fusion (section 4.3) may constitute an adequate explanation for their semantic homogeneity.

Relations between the complementizers of English may exist in a form analogous to what we have described above for French, but only a systematic study of the lexicon can reveal it. Notice that French traditional grammarians have overlooked important correlations in a similar way: for example, the study of the subordinating conjunction *que* traditionally constitutes a chapter which is independent of the chapter on infinitive constructions.

3.2 *The lexicon-grammar of frozen sentences*

By frozen sentences we mean for example idiomatic, metaphorical, or technical sentences such as (7) to (12):

7. *Max took the bull by the horns.*
8. *Max took Bob for a ride.*
9. *Max crossed swords with Bob.*
10. *Max cut the ground from under Bob's feet.*
11. *Bob's heart went out to the starved children.*
12. *The game is not worth the candle.*

Such sentences have a meaning that cannot be inferred from the meaning of the individual words that compose them. They differ from the usual sentences of linguistic discussion in that frozen positions do

not allow substitutions of phrases (i.e. selectional restrictions), as discussed in section 5.2.1 below. In (7), only the subject is a free position where substitutions occur; both complements are frozen in this respect. In (8), the subject and the direct object are free. In (9), the subject and the complement in *with* are free, while the direct complement is frozen. In (10), the subject is free, the direct complement the *ground* is frozen, and we observe a noun complement, *Bob*, which is free, although it modifies the frozen phrase *from under the feet*. In (11) the subject is frozen and, again, its noun complement is free; the complement is free. In (12), only the tense is free.

We studied the frozen sentences of French, and we classified them according to the number and the shape of their complements, and to the syntactic location of their frozen parts, that is, according to the factors described in section 3.1.1 (cf. Tables 8.2 and 8.3). Most studies on frozen sentences are aimed at reducing them to free structures. Typically, the absence of such a property as passive for the sentence:

13. *Max kicked the bucket*.

constitutes an argument for 'packaging' the sequence *kick the bucket* into a structure equivalent to that of a free verb similar to TO DIE. We are not interested in such attempts, and we analyse (13) as a structure N_0 V N_1, that is, as a subject, verb, direct complement form, which happens to be marked negatively for certain syntactic properties (e.g. passive).

So far, we have accumulated and analysed more than 20,000 sentences (Gross 1982). We describe frozen sentences in the same terms as free sentences of the lexicon-grammar. The definitions of the classes of frozen forms are largely based on the number and shape of the complements. We simply represent the positions of frozen nouns by the symbol C_i (C for constant) instead of N_i for variable noun. For example, (13) is represented as N_0 V C_1, and (7) as:

N_0 V C_1 Prep C_2 =: $(Max)_0$ *took* $(the\ bull)_1$ *by* $(the\ horns)_2$

Sentence (11) will be noted:

$(N's\ C)_0$ V Prep N_1

since in its subject position (i.e. N_0 or $(X)_0$) one finds a free noun phrase: N =: *Bob*, which is a noun complement of the frozen part: C =: *heart*.

We defined some classes (C_5 to C8) by means of the sentential properties of their N_is:

N_i =: *que* S + V-inf W

We also had to limit ourselves to minimal sentences, which implied separating frozen adverbials from 'essential' frozen complements. The

TABLE 8.2 *A classification of frozen verbal expressions*

Tables	Structures	Examples	Size
C1	N_0 V C_1	*Il a loupé le coche*	4,450
CAN	N_0 V (C *à* — *de* N)$_1$	*Cela a délié la langue de Max (lui)*	810
CDN	N_0 V (C *de* N)$_1$	*Il bat le rappel de ses amis*	610
CP1	N_0 V Prep C_1	*Il charrie dans les bégonias*	1,850
CPN	N_0 V Prep (C *de* N)$_1$	*Il abonde dans le sens de Max*	320
C1PN	N_0 V C_1 Prep N_2	*Il a déchargé sa bile sur Max*	2,010
CNP2	N_0 V N_1 Prep C_2	*Ils ont passé Max par les armes*	1,610
C1P2	N_0 V C_1 Prep C_2	*Il met de l'eau dans son vin*	1,040
CPP	N_0 V Prep C_1 Prep C_2	*Il tape du poing sur la table*	210
CPPN	N_0 V C_1 Prep C_2 Prep C_3	*Il se met le doigt dans l'œil jusqu'au coude*	370
C5	Que P V Prep C_1	*Que Max reste milite en sa faveur*	170
C6	N_0 V Qu P Prep C_2	*Il a pris du bon côté que Max reste*	300
C7	N_0 V C_1 *à ce* Qu S	*Il a dit non à ce que Max reste*	140
C8	N_0 V C_1 *de ce* Qu S	*Il se mord les doigts de ce qu'il est resté*	290
CPQ	N_0 V C_1 Prep *ce* Qu S	*Il partira à temps pour voir Luc*	330
CPPQ	N_0 V C_1 Prep C_2 Prep *ce* Qu S	*Il rend grâce au ciel de ce qu'il a réussi*	210
CADV	N_0 V Adv	*Cela ne pisse pas loin*	350
CV	N_0 V (Prep) VW	*Il est parti sans laisser d'adresse*	450
Co	C_0 V W	*La moutarde monte au nez de Max*	1,320
CoQ	C_0 V Prep N Qu S	*Peu lui importe s'il part*	300
CoE	(V + X) W	*Minute papillon!*	1,100
A1	N_0 *avoir* C_1	*Il a eu le mot de la fin*	280
A1PN	N_0 *avoir* C_1 Prep N_2	*Il a barre sur Max*	110
ANP2	N_0 *avoir* N_1 Prep C_2	*Il a Max en horreur*	90
A12	N_0 *avoir* C_1 Adj_2	*Il a la vue basse*	110
A1P2	N_0 *avoir* C_1 Prep C_2	*Il a mal aux cheveux*	340
Eo1	C_0 *de* N *être* Adj	*La barbe de Max est fleurie*	830
EoP1	C_0 *être* Prep C_1	*Le temps est à la pluie*	340
		Total of entries	20,340

TABLE 8.3 *A sample of a syntactic table of frozen verbal expressions whose structure is* N_0 V C_1 *(i.e. verbs with one direct frozen complement)*

N_0: N-hum	N_0: N-hum		N_0 V	Passive		N_1: N^0pc	Other Det	
+	−	*mijoter*	−	+	*un*	−	+	*mauvais coup*
+	−	*monter*	−	+	*une*	−	−	*crème*
+	−	*monter*	−	−	*la*	−	−	*gamme*
+	−	*monter*	−	−	*la*	−	−	*garde*
+	−	*monter*	−	+	Poss-O	−	−	*ménage*
+	−	*montrer*	−	−	Poss-O	+	−	*cul*
+	−	*montrer*	−	−	*les*	+	−	*griffes*
+	−	*mordre*	−	−	*la*	−	−	*poussière*
+	−	*se mordre*	−	−	*les*	+	−	*lèvres*
−	+	*se mordre*	−	−	*la*	−	−	*queue*
+	−	*moudre*	−	+	*du*	−	−	*vent*
+	−	*mouiller*	−	−	Poss-O	−	−	*culotte*
+	−	*mouliner*	−	+	*du*	−	−	*vent*
+	−	*mûrir*	−	+	Poss-O	−	+	*décision*
+	−	*mûrir*	−	+	Poss-O	−	+	*déclaration*
+	−	*mûrir*	−	+	Poss-O	−	+	*intervention*
+	+	*nettoyer*	−	+	*les*	−	−	*écuries d'Augias*
+	+	*nier*	−	+	*la*	−	−	*évidence*
+	−	*noircir*	−	+	*du*	−	−	*papier*
−	+	*nourrir*	−	−	Poss-O	−	−	*homme*
+	−	*nourrir*	−	−	*de*	−	−	*noirs desseins*
+	−	*noyer*	−	+	*le*	−	−	*poisson*
+	−	*numéroter*	−	−	Poss-O	−	−	*abattis*
+	−	*observer*	−	+	*le*	−	−	*silence*
+	+	*occuper*	−	+	*le*	−	−	*terrain*
+	+	*ouvrir*	−	+	*le*	−	−	*bal*
+	−	*ouvrir*	−	−	*le*	−	−	*ban*
+	−	*ouvrir*	−	−	*le*	+	−	*bec*
+	−	*ouvrir*	−	−	*la*	+	−	*bouche*
+	+	*ouvrir*	−	+	*la*	−	−	*chasse*
+	−	*ouvrir*	−	+	*le*	−	−	*gaz*
+	−	*ouvrir*	−	−	*de*	+	−	*grands yeux*
+	−	*ouvrir*	−	−	*la*	+	−	*gueule*
+	−	*ouvrir*	−	−	*les*	−	−	*guillemets*
+	+	*ouvrir*	−	+	*les*	−	−	*hostilités*

main criterion used for this purpose was the obligatory (vs. optional) character of the C-phrase:

- obligatory phrases are parts of frozen simple sentences;
- optional phrases are adverbials, in general.

In this way, the system of the frozen classes can be directly compared with the lexicon-grammar of free sentences:

1. Frozen expressions have often been considered exceptions by linguists. The notion of an exception appears to have the following basis:

- meaning is non-compositional;
- frozen expressions do not present most common syntactic properties.

But systematic studies have indicated that, from a lexical point of view, frozen expressions constitute a basic phenomenon. The size of the lexicon of verbs (i.e. 12,000 free sentences) and the size of the set of frozen sentences (i.e. 20,000) are in the same order of magnitude.[22] Thus, the phenomenon occupies at least the same size in the lexicon as what is distinguished as the normal case.

2. Since adverbs are most frequently optional, they have been the subject of separate studies that have led to conclusions analogous to those reaching for frozen sentences. Thus, in French, the regular type of manner adverbial has the shape of an adjective followed by the suffix *-ment* (the corresponding situations would occur in English with the suffix *-ly*).[23] Figures for regular and frozen adverbs are of the same type: about 2,000 'regular' adverbs (Molinier 1982) and more than 6,000 frozen adverbs (Gross 1990).

We also observed that the range of meaning of frozen adverbs is narrow: a large proportion of them mean 'a lot' or 'strongly', as in the pair:

Bob laughed fit to split his ribs. — *Bob laughed heartily.*

A typical and large class of such adverbs has the form of *like* phrases, as in:

Bob mange comme (quatre + un porc). *Bob eats like (hell + a horse).*

The classes of frozen adverbials are summarized in Table 8.4.

[22] This is a lower estimate in the sense that this coverage is not as complete as the coverage of the 12,000 free sentences. The lexicon-grammar contains what are commonly termed clichés, metaphors, and proverbial expressions that are not productive but numerous and will have to be counted as frozen forms. Constructions which have been found to be productive have not been listed as lexical items; they are given as processes, and as such do not noticeably increase the figures. Freezing compound nouns is probably the general way of producing new lexical items, at least in technical domains: *lunar module*, *electron gun*, etc.

[23] The derivation of these so-called regular adverbs is not always obvious. For example, *heartily* could well be a frozen form.

TABLE 8.4 *Classes of frozen adverbials*

Tables	Structures	Examples	Size
PADV	Adv	*soudain*	520
PC	Prep C	*en bref*	590
PDETC	Prep Det C	*contre toute attente*	750
PAC	Prep Adj C	*de sa belle mort*	670
PCA	Prep C Adj	*à gorge déployée*	710
PCDC	Prep C *de* C	*en désespoir de cause*	620
PCPC	Prep C Prep C	*des pieds à la tête*	240
PCONJ	Prep C Conj C	*en tout et pour tout*	290
PCDN	Prep C *de* N	*au moyen de* N	450
PCPN	Prep C Prep N	*par rapport à* N	140
PV	Prep V W	*à dire vrai*	240
PF	P (fixed phrase)	*Dieu seul le sait*	360
PECO	(Adj) *comme* C	*comme ses pieds*	300
PVCO	(V) *comme* C	*comme un cheveu sur la soupe*	330
PPCO	(V) *comme* Prep C	*comme dans du beurre*	40
PJC	Conj C	*et tout le tremblement*	150
		Total of entries	6,400

Source: Gross (1990).

Among others, the following empirical observation seems to us to be of some importance, although it is not clear how it could be accommodated in any of the current theories: the (frozen) nouns involved in frozen sentences are practically all non-human. In a sample of about 12,000 sentences we observed about 100 human nouns (vs. more than 12,000 non-human nouns). Such an observation simply cannot be conceived of without a framework that requires systematic enumeration of the elements of a lexicon-grammar.

We are now in a position to explain our use of the argument of idiomatic invariance (argument (4), section 1.2). From an experimental point of view, idioms present certain advantages over free forms. Consider for example the pairs of free sentences (Dubois 1969):

Max has no money. = *Max is without money.*
Mas has no courage. = *Max is without courage.*
This remark has no importance. = *This remark is without importance.*

The members of the pairs are synonymous, but it is hard to assume that a *syntactic* relation links them; the change of verb, and the presence of the negation (a clue indicating that we are not dealing with

elementary sentences)[24] are obstacles to such a hypothesis, in principle. However, there are frozen pairs such as:

Max has neither hearth nor home.
– *Max is without hearth or home.*

This remark has neither rhyme nor reason.
– *This remark is without rhyme or reason.*

similar to the preceding free ones. The unique combination of nouns they contain are found only in these two contexts (to within standard transformational variations).[25] We now face an alternative: either the frozen part occurs in these positions by mere accident, or else there is a relation between the two sentence forms. We consider that the number and variety (arguments (2) and (3) of section 1.2) of French pairs force us to establish a relation between the members of such a pair. Once established, the relation applies to the free forms above as well. This type of reasoning has occasionally been applied: to the deletion of TO HAVE in the analysis of constructions of TO WANT (McCawley 1979), to the movements of the form HEADWAY in the discussion of relative clauses (e.g. Vergnaud 1985), etc. However, we consider it a major tool, its generality being revealed by a systematic study of French.

Needless to say, compiling, filing, sorting, and classifying 20,000 frozen expressions raises serious practical problems. The present results could not have been obtained without the use of computer technology. For example, an expression such as *to take the bull by the horns* should be indexed according to the three words *take*, *bull*, and *horn* and according to syntactic criteria as well. Also, drawing the line between frozen adverbs and frozen sentences has proved quite difficult in certain situations; the automatic construction of indexes has helped us to avoid duplications beween the two categories. A special system for managing this database had to be built (Vasseux 1982); data can be corrected, updated, and expanded; access to all the information has been made convenient.

3.3 *Support verbs and operator verbs*

In section 3.1 we discussed verbs that could be roughly described as co-occurring with semantically selected sets of nouns. For these verbs,

[24] Notice that the obligatory negation involved in the relation should be no more surprising than it is with certain passive forms, such as: *Such an attitude was unheard of until now.*

[25] There are also forms such as:

There is neither rhyme nor reason to this remark.

which are associated with sentences containing the support verb TO HAVE, cf. § 3.3.1, or TO BE WITHOUT, and with obligatory negation.

the selections of the subjects and the object(s) are generally considered to be independent of each other (e.g. the range of subject nouns for the verb TO EAT is roughly independent of the range of its direct object). There are however simple sentences which cannot be analysed at all in terms of selectional restrictions. They constitute an important part of the lexicon, and they involve special structures and grammar rules. We now give examples of these elements.

3.3.1 *Support verbs*

Consider the following sentences:

14. (*a*) *Max signed this agreement with Jo.*
    (*b*) *Max criticized this agreement with Jo.*

They have identical analyses: the subjects and objects are the same and the complex object is decomposed into a head noun: *agreement* and a prepositional noun complement *with Jo*. There is however an important structural difference that appears when clefting is applied to the objects of (14*a*–*b*):

It is with Jo that Max signed this agreement.
It is this agreement that Max signed with Jo.
It is this agreement with Jo that Max signed.
**It is with Jo that Max criticized this agreement.*
**It is this agreement that Max criticized with Jo.*
It is this agreement with Jo that Max criticized.

There is also an important semantic difference between (14*a*) and (14*b*): in (14*a*) *Max* is a partner of *Jo* in the *agreement*, whereas this is not the case in (14*b*). This difference in meaning reflects a difference in form, as seen from:

**Max signed your agreement with Jo.*
Max criticized your agreement with Jo.

In some sense, the main predicate of (14*a*) is *agreement*: it selects the subject *Max* and the complement *with Jo*. The main predicate is *criticize* in (14*b*). Again, this difference has syntactic implications: (14*a*) is a symmetrical construction, that is, a construction where the subject and the *with* complement can be conjoined:

14. (*a*) = *Max and Jo signed this agreement.*

This is not the case with (14*b*):

14. (*b*) ≠ *Max and Jo criticized this agreement.*

If we compare (14*a*) and:

14. (*c*) *Max has this agreement with Jo.*

the difference of meaning is not substantial. We consider this difference to be mainly aspectual: TO SIGN indicates the beginning of the *agreement*, whereas the duration is not bounded in (14*c*). In (14*c*), the verb TO HAVE cannot be considered a main predicate, at least from the selectional point of view. Sentence (14*c*) is symmetrical in the same way as (14*a*):

14. (*c*) — *Max and Jo have this agreement*.

We call TO HAVE and TO SIGN support verbs in such constructions and note them Vsup. More generally, support verbs are verbs which do not present selectional restrictions with respect to their subject and complement. Instead, a selectional relation holds between subject and complement.

Thus, with respect to the phenomenon described, the lexicon-grammar of verbs must be subdivided into:

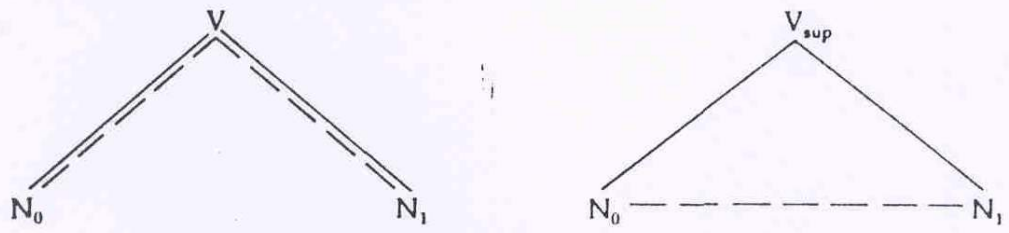
- ordinary verbs such as TO CRITICIZE, TO BREAK, TO FIGHT, TO APPROVE that select certain semantically defined subjects and objects;
- and support verbs such as TO HAVE in (14*c*), TO SIGN, TO CONCLUDE, TO POLISH, etc., that are in complementary distributions with TO HAVE (Vivès 1983).

Intuitively, the corresponding structures differ as in Fig. 8.1, where forms N_0 V (Prep) N_1 have been represented in dependency terms. The phenomenon is actually more complex, since the supported noun (here AGREEMENT) plays a crucial role. We have exactly the same facts with nouns such as CONTRACT and TREATY, whereas with other nouns the facts may be different. Consider the noun *cheque* in:

15. (*a*) *Max signed a cheque for Jo*.
    (*b*) *Max lost a cheque for Jo*, etc.

Clefting applies exactly as it does to sentence (14):

It is for Jo that Max signed a cheque.
**It is for Jo that Max lost a cheque*.



Note: Sectional constraints are noted with dotted lines, grammatical constraints with solid lines

Fig. 8.1 Distributional verbs versus support verbs

and the meaning difference is analogous to that of (14*a*)–(14*b*): in (15*a*) *Max* is the author of the *cheque*, but not in (15*b*). However, the set of support verbs is no longer the same: TO CONCLUDE is not accepted in forms such as (15*a*), whereas TO MAKE and TO WRITE are. Thus, TO SIGN is a support verb in its combination with *cheque*, and we will say that the variant support verbs are not the same as those with AGREEMENT.

Variability of support verbs is quite common and conspicuous in the context of nominalization and adjectivizations, as can be seen from the following examples:

16 (*a*) *This result contradicts your assertion.*
(*b*) = *This result is in contradiction with your assertion.*

17. (*a*) *Max loves Eva.*
(*b*) = *Max is in love with Eva.* (Danlos 1980; Négroni-Peyre 1978)

18. (*a*) *This wall is high.*
(*b*) = *This wall has a certain height.*
(*c*) = *This wall is of a certain height.*

19. (*a*) *My son is sensitive.*
(*b*) = *My son has a certain sensitivity.*
(*c*) = *My son is of a certain sensitivity.* (Dubois 1969; Meunier 1977)

20. (*a*) *Your proposal interests the Dean of Students.*
(*b*) = *Your proposal has a certain interest for the Dean of Students.*
(*c*) = *Your proposal is of some interest to the Dean of Students.* (Labelle 1974).

21. (*a*) *Bob sleepwalks*
(*b*) = *Bob is a sleepwalker.*

22. (*a*) *Bob skin-dives.*
(*b*) = *Bob does skin-diving.*
(*c*) = *Bob is a skin-diver.*

All these relations are described as involving a support verb that induces morphological changes in the verbs and adjectives associated with it. The support verbs used here are TO HAVE, TO BE IN, TO BE OF, TO BE, and TO DO. In general, support verbs preserve constraints present in the 'initial' sentences, very much as auxiliaries or modal verbs do. For example, in (16*b*), the subject *result* is still the semantic subject of *contradiction*. All these relations can be justified by means of arguments of the kind outlined in section 1:

- the qualitative argument applies obviously to the distributions of nouns, once the supported noun has been recognized. This

recognition is not trivial, as will be seen on further examples, which may explain why the impact of such relations has not been recognized earlier;

- the quantitative argument has an important role: where the lexicon has been investigated, thousands of verbs and adjectives were found to enter into one or several of the relations (16) to (22);
- the argument of variety applies too; for example, the relations in (18), (19), and (20) involve nouns that are semantically quite different: *concrete*, *human*, *abstract*;
- moreover, in all cases, we have been able to find pairs involving frozen expressions, which allowed us to use the argument of idiomatic invariance:

This drink rots your guts
– *This drink is a rot gut.*

Her words warmed Bob's heart
– *Her words were heartwarming to Bob.*

Although various questions are still open, we consider that the present coverage of the lexicon leads us to a new position with respect to derivational morphology. Traditional (derivational) morphology is the level of description of the relations between words. The way we deal with nominalizations, adjectivizations, and adverbializations (Harris 1976, 1982) eliminates this level, which becomes part of the description of simple sentences. Thus, the status of the relations presented is not different from the status of, say, passive or extraposition.

Support verbs have a crucial function with respect to a lexicon-grammar. In most of the preceding examples (i.e. from (16) on), they appear as a member of a syntactic relation. But we consider that they also occur in sentences that stand alone, as in (14*c*) and (15*a*), and in:

Max is off the hook.
Max is at odds with Bob. – *They are at odds.*
Max is in touch with Bob.
This drug has a certain effect on Bob.
The burden of the proof is on Bob.

The reasons for considering that these sentences contain a support verb are the following:

- it is difficult to conceive the verbs TO BE and TO HAVE as involved in selectional restrictions with the nouns,
- the verbs are exactly those found in the nominalization relations studied over the lexicon-grammar of verbs,

- the syntactic properties of the complements of the supported noun are exactly the same (e.g. clefting, symmetry).

From our theoretical point of view, we are thus able to incorporate nouns in the lexicon-grammar, using the same treatment as was used for other parts of speech. We will call stand-alone nouns those nouns that do not enter into any (syntactic) relation of derivation. So far, thousands of stand-alone nouns have been found in combination with support verbs. Current work consists in separating the meanings of nouns on the syntactic basis provided by the Vsups. The number of such expected stand-alone nouns is expected to grow into the hundred thousands.

3.3.2 *Operator verbs*

There are sentences which possess most of the features observed in the support constructions, but which are more complex. Whereas support constructions have the general form:

N_0 Vsup (Prep) N_1 —: *Max got into trouble.*

with a relation holding between N_0 and N_1, these constructions have the general type:

(*a*) N_0 Vop N_1 Prep N_2 —: *This remark got Max into trouble.*

and a relation between the two complements N_1 and N_2. In general, this relation coincides with a support relation, that is, one can always find a Vop and a Vsup such that (*a*) and (*b*) hold:

(*b*) N_1 Vsup Prep N_2 or N_2 Vsup Prep N_1

that is, such that these two structures have the same nominal content in N_1 and N_2. These complex support verbs will be called operator verbs: Vop. Thus, we propose to analyse (*a*) by an application to (*b*) of the operator N_0 Vop:

- TO PUT, TO GET, and TO SET (in French METTRE) are causative operators on sentences with Vsup —: *to be* Prep:

23. *The accident put this strange idea into Bob's mind.*
24. *Max got Bob off the hook.*
25. *I put Max in touch with Bob.*
26. *This theory (puts + sets) this result in contradiction with your assertion.*
27. *I put the burden of the proof on Bob.*
28. *I put the blame on Bob.*

- TO GIVE (donner) is an operator that mostly applies to TO HAVE, TO GET (avoir) as in the idiomatic pairs:

Bob has a finger in the pie
– *The shake-up in our office gave Bob a finger in the pie.*

Bob got the sack.
– *Max gave Bob the sack.*

Bob (got + has) the creeps.
– *The fire gave Bob the creeps.*

- more restricted causative forms are found in the pairs:

This remark set Max at odds with Bob.
– *This remark set them at odds.*

Bob is on the go.
– *That will keep Bob on the go.*

The crumb went down the wrong way.
– *Bob gulped the crumb down the wrong way.*

A shiver went down Bob's spine.
– *The crash sent a shiver down Bob's spine.*

The accident drove Max out of his mind.
24. – *Max let Bob off the hook.*

In French, operator verbs have been found to be of great generality. Their range is such that many verbs with two complements can be analysed as a causative operator on support verbs. Moreover, operator verbs have equivalents, very much as Vsups do (e.g. LET is an equivalent of GET in the last example). Equivalents of Vops have syntactic properties similar to those of Vsups (e.g. retaining obligatory binding of possessive adjectives, as discussed in Gross 1981). A different type of Vop is found in the pairs:

The solution is at Bob's fingertips.
– *Bob has the solution at his fingertips.*

Bob's job is at stake.
– *Bob has his job at stake.*

The effect of the operator N_0 *have* is to bind obligatorily to its N_0 a noun phrase of the Vsup-sentence. Here, the sentences under the operator are sentences with support verb TO BE AT, and the bound noun phrase is in a noun complement position. This Vop is different from the causative operators in the sense that it does not introduce a new argument in the Vop-sentence. The previous examples of Vop all had an agentive or causative subject (N_0) that was new with respect to the subject of the Vsup-sentence. Binding operators are quite general (Gross 1972), as seen from examples such as:

Bob has a strange habit: he eats one cigarette a day.
?**Bob has a strange habit: Max eats one cigarette a day.*

Such binding operators[26] also apply to derived sentences. Consider the utterances:

29. *Bob crossed his fingers.*

[Passive]: 30. *Bob's fingers are crossed.*

The possessive *his* must refer to *Bob*, otherwise the idiomatic sense dealt with here is lost, as it is lost for example in:

Bob crossed Eva's fingers.

Note that the source of *his* in the active form (29) reappears in this passive form (30) with its idiomatic sense. The passive form is a *be*-sentence, to which the binding operator N_0 *have* can be applied, retaining the idiomatic sense:

[Vop N_0 have]: 31. *Bob has his fingers crossed.*

Further equivalent aspectual verbs are also observed here:

Bob kept his fingers crossed.

and finally we observe sentences such as:

Bob waited for the result, with his fingers crossed.

where the *with* phrase has the idiomatic meaning. We are then led to consider *with his fingers crossed* a transform of (31), and to generalize this solution to free forms, as for example to those studied by Ruwet (1978). The motivations for introducing operator verbs are the same as for support verbs:

- impossibility of attaching selectional restrictions to them;
- the argument of variety, and its application to frozen forms (argument of idiomatic invariance); as a matter of fact, most of the examples we have given are idiomatic, which should not hide the fact that: (1) the quantitative argument applies to numerous free forms, and (2) from a theoretical point of view, the use of operators constitutes a general method for representing dependencies between the arguments of numerous verbs, dependencies that are never accounted for in grammars.

3.3.3 *Remarks on the place of Vsups and Vops in grammars*

Support and operator verbs are known devices, in a certain sense. In traditional grammar Vsups are dealt with under the name of verbal periphrases (e.g. Boström 1957). In German grammar they belong to an important chapter of traditional grammar. In English and in French

[26] These are interesting by-products of systematic studies. For example, the binding problem raised by N_0 *have* led us to focus on other binding difficulties that occur in examples such as:

Bob's eyes failed him.

The idea of killing himself did not occur to Bob.

grammar there is no such consensus, but one finds many isolated proposals.[27] For example, the problem raised by the so-called picture nouns constructions are linked to properties of Vsups (Cattell 1984). Harris (1964) introduced nominalizations as transformational relations, that is, as relations between *two sentences*:

Bob walked. — *Bob took a walk.*

whereas Chomsky (1969), criticizing the transformational treatment of nominalizations, assumed that they are relations between *one sentence* and *one noun phrase*, as in:

John is self-indulgent —> *John's self-indulgence.*
Bob walked. —> *Bob's walk.*

We immediately see that, in this type of generative formulation, there is no room for the support verb. In both cases, one has to state rules for the combinations of verbs with derived noun phrases. This question is left open in generative grammar, but with Harris's proposal two fundamental types of combinations are distinguished:

- one type is the combination of nominal forms with support verbs which is accounted for by the nominalization relations;
- the other case, observed in (7*b*) and (8*b*) (section 3.3.1), raises many difficulties, but Harris's model of 1982 accounts for many of them. Roughly, sentences such as (7*b*) and (8*b*) have to be considered as complex forms that result from the combination of a sentence with an ordinary verb and a sentence with a support verb that has been reduced to a noun-phrase. We will thus have such derivations as:

  Max criticized the agreement that you had with Jo.
  — *Max criticized your agreement with Jo.*

  Max lost a cheque that someone (signed + wrote) for Jo.
  — *Max lost a cheque that was (signed + written) for Jo.*
  — *Max lost a cheque (signed + written) for Jo.*
  — *Max lost a cheque for Jo.*

In these relations, the support verbs are erased. Notice that the zeroing operation is not an issue here, for only Vsups, which are narrowly constrained by the supported Ns, have to be reconstructed. Furthermore, not all Vsups are zeroable, so that not all sentences with Vsup yield an NP. For example, we have the relation:

Bob accused Ida. — *Bob levelled an accusation at Ida.*

and the last sentence can be nominalized as in the example:

[27] Lafaye (1841) produced a French dictionary where the support verb FAIRE was used as a systematic device for relating nominal entries to verbal ones.

The accusation that Bob levelled at Ida appalled Jo.

but we do not have reduction of the Vsup:

**Bob's accusation at Ida appalled Jo.*

Also, the zeroability of a given Vsup depends on the N it supports (cf. Giry-Schneider 1978, 1987; Gross 1981).

4. SEPARATING THE ENTRIES OF THE LEXICON-GRAMMAR

The problem of separating the different uses of a given verb is fundamental, for its solution in part determines the shape of the grammar. Consider a commonly encountered situation where a verb is said to have two uses, one proper, the other one figurative or metaphorical. We have been able to separate isolated (section 4.1) from productive (section 4.3) cases, only because we have operated in the framework of a lexicon-grammar, that is, in a framework where a methodological enumeration of facts is required. Actually, a lexicon-grammar representation can be viewed as a verification of the fact that the various uses of the dictionary entries of a language are satisfactorily described in terms of a given grammar, namely, that no relation between sentences which are obviously related to each other (i.e. by arguments (1) to (4) of section 1) is left unaccounted for by the rules of the grammar. We alluded to this activity of separation of entries in section 3.1.1, as a result of which Vsups and Vops had to be distinguished because their properties could not be fitted into the general description of the lexicon-grammar. Here, we will discuss further difficulties we have met and, in some cases, the solutions we were able to propose.

4.1 *Isolated examples*

4.1.1

Consider the two constructions:

32. *I wish I could help Bob.*
33. *I wish Bob more success.*

The question arises whether they correspond to a single entry, since intuitively the meanings of TO WISH could be considered as similar in both sentences. However, the structures are different: (32) has a single sentential complement; (33) has two complements. We could hypothesize a relation between (32) and (33):

I wished Bob had more success. — *I wished Bob more success.*

which would be similar to the object-raising proposed for TO BELIEVE in:

I believed Bob was sick. — *I believed Bob sick.*

However, instead of deleting TO BE, one would have to delete TO HAVE. But, even so, it is not clear how it is that (33) can be interpreted as a communication process between *I* and *Bob* (e.g. TO TELL, TO WRITE), whereas (32) cannot have this meaning. Moreover, the proposed deletion and the change of meaning seem limited to the single verb TO WISH, making it difficult to support the existence of the relation by means of a quantitative argument or an argument of variety. Under such dubious conditions, rather than using a raising operation, we decided in favour of the solution of two independent entries.

4.1.2

Consider now the constructions:

34. *Jo shared a dish with Bob.*
35. *Jo shared a dish between Max and Bob.*

Again, the intuition is that *to share* has about the same meaning in (34) and (35). However, in (34) the subject *Jo* gets a part of the dish, but not in (35). One could consider relating (34) and (35) by means of an operation that would change the sequence *between himself and* to the preposition *with*. But this operation does not have much plausibility, since it does not seem to apply to any other verbs than TO SHARE and TO SPLIT. Also, the change from structure (34) to structure (35) is too important, in the sense that no similar operation has been observed in other contexts. We are then led to conclude that we are dealing with two different entries TO SHARE, a solution supported by the observation of differences in the distribution of the object:

Jo shared a room with Bob.
?Jo shared a room between Bob and Max.

Notice that we could raise a similar problem about:

36. *Jo shared Bob's dish.*

Although (34) and (36) are closely related in meaning, the operation that would change the second complement *with* N_2 into a noun complement is not a standard one. But here we could use the sentence:

37. *Jo shared Bob's dish with him.*

with *him* referring to *Bob*. In (37) deletion of the redundant form *with him* is a likely solution. The *share* of (36) can thus be reduced to that of (34).

4.1.3

The same questions have to be raised about *proper* and *figurative* meaning of verbs, as seen for example in the parallel pairs:

(prop):
This soap irritates my skin.
– *This soap is irritating to my skin.*
(fig):
This problem irritates Max.
– *This problem is irritating to Max.*

But here, the pattern of adjectivization depends on the verb:

(prop):
The box's falling crushed Max's foot.
– **The box's falling was crushing to Max's foot.*
(fig):
The loss of his son crushed Max.
– *The loss of his son was crushing to Max.*

(prop):
These rags absorb oil.
– *These rags are absorbent.*
(fig):
These thoughts absorb Max.
– *These thoughts are absorbing to Max.*

The dependence between what is usually called *proper* or *figurative* meaning and syntactic properties is found with other pairs of constructions (Boons 1971). Thus, consider the following examples which involve passive: with TO REQUIRE, both meanings have a passive form:

(prop):
Max required this book.
– *This book was required by Max.*
(fig):
This question required all our attention.
– *All our attention was required by this question.*

but with TO DESERVE, only one meaning has a passive form:

(prop):
Max deserved this book.
– *This book was deserved by Max.*
(fig):
This question deserved all our attention.
– **All our attention was deserved by this question.*

The two uses of TO REQUIRE appear to differ only by the distributions of nouns in N_0 and N_1. The terms *proper* and *figurative* each refer to a separate distribution of nouns. The same is true for TO DESERVE, but, in addition, the two uses of this verb differ transformationally. In the latter case, the solution appears to be complete separation between uses which will correspond to two independent entries, i.e. two different verbs TO DESERVE. The process by which a given verb (e.g. TO REQUIRE) comes to acquire a new set of co-occurring N_is is seen so far to be largely accidental. In fact, many examples we have encountered and which could be termed figurative or metaphorical had to be described as frozen sentences (e.g. *Fortune smiled on Max*). This is an indication of the anecdotal character of the formation. But we have also observed situations where whole blocks of the lexicon underwent the same change of distribution (cf. section 4.3). However, at the point of the description that we have now reached, i.e. when no further productive relation can be readily detected, it is safer to assume that cases analogous to TO REQUIRE also correspond to two separate verbs.

4.2 *Dual subjects*

The following example is particularly important because it affects a large number of verbs, for which it involves the causative and agentive meanings of the grammatical subject. The difference to be discussed has not been observed in complement positions of verbs or in subject positions of adjectives. Consider the verb TO AMUSE in:

38. *Max amused Ida.*
39. *His face amused Ida.*

The semantic functions of the subject are different in (38) and (39). *His face* can be said to be *causative* or *non-active*,[28] whereas *Max* is ambiguous: in one sense, it can be *active* or *agentive*, in the other, *non-active* (e.g. synonymous with *Max's presence*). Parallel to (38) and (39), we have associated adjectivized sentences:

Max is amusing to Ida.
His faces are amusing to Ida.

where the subjects are only *non-active*. It is then legitimate to ask whether one is dealing here with one or two verbs TO AMUSE. Since there are more than 1,000 verbs about which exactly the same question can be raised, it may be worthwhile to hypothesize a regular relation between the two functions of the subject rather than duplicate the entries. Such a systematic difference is an important argument of the

[28] The semantic opposition *active* vs. *non-active* has turned out to be reproducible in these syntactic positions.

quantitative type in favour of a relation between the two forms. An argument of variety also applies, since the systematic difference is observed under varied semantic conditions. Moreover, the argument of idiomatic invariance also applies, for example, with dual subjects such as:

Max broke the ice by saying hello.
This simple remark broke the ice.

Max cut the ground from under Bob's feet.
This remark cut the ground from under Bob's feet.

Here it is difficult to argue that each interpretation constitutes a separate entry of the lexicon-grammar. Hence we have decided in favour of the relation, although we are unable to analyse it convincingly. Also, practical reasons of compactness of the lexicon-grammar led us to make this choice.

4.3 *Fusion*

Consider the following pairs of sentences:

He eats too much.
He will eat his family out of business.

He drinks expensive wines.
He will drink his family out of business.

He gambles on horses.
He will gamble his family out of business.

The first members correspond to the *proper* use of the verbs, the second members bear some resemblance to the first members, but they cannot be distributionally related to them. Do we consider for each verb two separate entries in the lexicon, or do we introduce a relation between the two constructions? In this case, we have chosen to introduce a relation based on an operation called Fusion (Gross 1975). We consider the basic sentences:

$$\textit{He will put his family out of business by} \begin{pmatrix} \textit{eating} \\ \textit{drinking} \\ \textit{gambling} \end{pmatrix}$$

Fusion is a transformational relation that replaces *put* by the verb root found in the *by* complement. Since the verb has its proper use in this complement, we have reduced the derived meaning to an operation on the entries of the proper meaning.[29] One argument in favour of this analysis is the unacceptability of such forms as:

[29] We analyse the basic sentence starting from a frozen expression: N_0 isPrep X. An utterance such as Prep X =: *out of business* is frozen. It should be limited in the range of forms where it

**He will gamble his parents out of business by drinking.*

where two verbs in complementary distributions are made to co-occur. In the same way, sentences such as:

They (*dug* / *tunnelled* / *gunned* / *blackmailed* / *cheated*) *their way* (*into the bank* + *out of prison*).

will be derived by Fusion from a sentence such as:

They made their way out of prison by (*digging* + . . . + *cheating*).

Presumably, sentences such as:

I (*argued* / *drove* / *talked* / *fooled* / *bribed* / *lured* / *trapped*) *him* (*into*) *leaving the place*,

where TO TALK, TO BRIBE, have different constructions from those of their proper use, will be analysed in a similar way. Fusion rules appear to be an explanation of productivity in cases where certain verbs acquire a new construction that cannot be derived distributionally from their already existing forms (Salkoff 1983). In French, introducing fusion rules in the lexicon-grammar has immediately reduced more than 1,000 complex forms to simple entries.[30]

5. LEARNING

5.1 *The role of meaning and the lexicon-grammar*

We have already mentioned that the rows of the matrix representing the lexicon-grammar of French are all different (section 3.1.2 (6)), as are the columns. Pairs of columns can be viewed as relations between

occurs. Several of the devices discussed above extend this range: support verbs, operators such as the causative TO PUT, and Fusion.

[30] It should be noted that lexical decomposition by Fusion rules has nothing to do with decomposition in generative semantics. Thus, in our first examples, we did not analyse *He will drink his family out of business* into simpler predicates, that is, in terms of predicates with two arguments instead of three. This sentence has been analysed into two sentences, one with two arguments and the other with three. In generative semantics such a reduction must be performed in terms of sentences with at most two arguments, preferably one.

two sentence forms. In certain cases, these relations are transformations between sentences. Since no two columns have identical sign contents, all syntactic relations, hence all transformations, have exceptions. Consider for example the complement *auprès de* N in the reflexive construction discussed in section 3.1.1 (example 1). This complement is permitted for some verbs in their reflexive construction N_1 *se* V *de* N_0, and forbidden for others in the same construction. The distribution of the '+' and '−' signs in the corresponding columns does not appear to be predictable on a semantic basis. This irregularity of syntactic relations raises new questions about language acquisition. The problem of learning the simple verbal structures can be stated in a suggestive way: how can a lexicon-grammar of a given size be learned in a fixed amount of time, knowing the type of data and of corrections to which a native speaker is exposed? It is difficult to conceive that the unacceptabilities we have seen so far can be learned *directly*. Thus, some indirect process must be at work that would derive unconscious unacceptabilities (and acceptabilities as well) from the limited examples that are daily heard and read. A well-known proposal can be used, that of the existence of a relationship between form and meaning. The hypothesis is that certain *intuitive* notions of meaning, namely, notions easy to perceive and to isolate for a child, are associated with certain syntactic forms. The key to the explanation is then in the nature of this association. But there are many questions that should be raised, and reasonable answers proposed for them, before the existence of the associations hinted at in traditional and generative grammar can even be made plausible. Various types of difficulties arise:

1. There are sentences which are closely synonymous such as:

Cette affaire concerne Max.
Cette affaire regarde Max.
(This affair concerns Max.)

but only one of them has a passive form. It is difficult to formulate a difference of meaning between the two sentences that accounts for the formal difference. Thus, it seems rather far-fetched to attribute to some vanishingly small difference of meaning the sharp difference of syntactic behaviour:

Max est concerné par cette affaire.
**Max est regardé par cette affaire.*

The following pair is somewhat similar to this case, but with a more noticeable difference in meaning:

Bob left the house.
**The house was left by Bob.*

Bob abandoned the house.
The house was abandoned by Bob.

Such pairs are never discussed in handbooks or in class, but the corresponding mistakes do not seem to be heard; although the facts are unknown outside of the profession, they are consistently learned by native speakers.

2. Another way of determining the relationship between form and meaning is given by the following example: a number of verbs (close to 1,000 in French) can be intuitively described as psychological verbs. More precisely, these verbs can be considered as functions that link a stimulus *s* and a person *h* who experiences the feeling *P* triggered by *s*. The sentences:

Max hates cakes and *Cakes disgust Max*

can be given this description: $P(s, h)$. Both sentences have the form N_0 V N_1 but different word orders. Rules of interpretation state the correspondence between form and meaning:

TO HATE: $N_0 - h$; $N_1 - s$
TO DISGUST: $N_0 - s$; $N_1 - h$

We asked the following question for French: given the two parameters of form, word order and prepositional form of the object, how many types of verbs can be found with the meaning $P(s, h)$? There are a priori six possible forms (two word orders combined with the three types of object: 'zero', *à*, *de*). We have already given two forms N_0 V N_1 *haïr* (to hate) and *dégoûter* (to disgust). Inspection of the lexicon-grammar immediately provides others:

- with Prep —: *à*:
  Les gâteaux plaisent à Max: $N_0 - s$; $N_1 - h$
  (Cakes please Max.)
  Max tient aux gâteaux: $N_0 - h$; $N_1 - s$
  (Max is keen on cakes.)
- with Prep —: *de*:
  Max rêve de gâteaux: $N_0 - h$; $N_1 - s$
  (Max dreams of cakes.)

The case *Les gâteaux* V *de Max* does not occur for any V. Thus, five cases are found out of six. With other analogous predicates, proportions are comparable. Such figures mean that the correspondence is close to arbitrary: to express a given meaning, any syntactic form a priori possible can be used.

3. Various syntactic categories are often thought to convey systematically certain semantic notions. Such associations would

simplify the learning of these notions. For example, the notions of singular and plural are usually localized in determiners. However, in examples such as:

Guests abounded in the hall.

the notion of plural is inherent in the meaning of the verb, and one could argue that here the verb plays the role of the determiner. The following relation with support verb *there is* makes this intuition explicit:

There is an abundance of guests in the hall.

Here the derived nominal occupies the syntactic position of a determiner, the same position is filled by a determiner in the quasi-synonymous sentences:

There is a large number of guests in the hall.
There are many guests in the hall.

5.2 *An abstract model of the learner*

How the first language is learned is the central question in generative linguistics. There, the problem is presented as arising from the fact that a native speaker, who has been exposed in his childhood to only a finite and perhaps even small number of utterances, has a potential for understanding an infinite number of sentences. This way of stating the problem hinges on certain assumptions that we will now examine in the light of our observations on various lexicon-grammars.

5.2.1 *Productivity*

The first question is about the infinite productivity of sentences. The opposition between the finite amounts of learnable terms and the infinite quantities that are understandable has been dramatized by emphasizing the way the number of sentences grows if their length is counted in number of words. To illustrate this point, consider the formula:

$$\begin{pmatrix}\textit{My little sister}\\ \textit{The man here}\end{pmatrix} \textit{ bought } \begin{pmatrix}\textit{large pots}\\ \textit{some books}\\ \textit{three hats}\\ \textit{three lamps}\end{pmatrix}.$$

The combinations correspond to $2 \times 4 = 8$ sentences, each of length six words. We increase its length to nine, by adding one complement, yielding:

$$\begin{pmatrix}\textit{My little sister}\\ \textit{The man here}\end{pmatrix} \textit{bought} \begin{pmatrix}\textit{large pots}\\ \textit{some books}\\ \textit{these lamps}\\ \textit{three hats}\end{pmatrix} \begin{pmatrix}\textit{for the sailor}\\ \textit{for poor Max}\\ \textit{for our father}\\ \textit{for the boys}\end{pmatrix}$$

The formula now includes $2 \times 4 \times 4 = 32$ sentences of length nine each. Roughly speaking, the number of sentences increases exponentially with the length. One has to conclude from this observation (i.e. combinatorial explosion) that sentences cannot be learned by rote.

There is however another way of looking at this situation, which seems more realistic and which attributes a clear role to the size of the lexicon, a factor not taken into account in the generative discussions of learning. We now illustrate this process with an example. Consider the following set of related frozen sentences:

40. $$\textit{Max will} \begin{pmatrix}\textit{beat}\\ \textit{whale}\\ \textit{lick}\end{pmatrix} \begin{pmatrix}\textit{the hell}\\ \textit{the shit}\\ \textit{the living daylights}\\ \textit{the daylights}\\ \textit{the tar}\end{pmatrix} \textit{out of Bob.}$$

They are all synonymous. Substituting one noun for another in the object position will not modify the meaning as it would, for example in:

41. $$\textit{Max smelled} \begin{pmatrix}\textit{the shit}\\ \textit{the hell}\\ \textit{the tar}\end{pmatrix}.$$

In other words, the substitution set in (40) is not a distribution, i.e. it is not constrained by selectional restrictions as is the set in (41). Substituting *shit* for *tar* in (40) introduces values such as *popular*, *slang*, or *obscene* English, but otherwise the sentences remain precise synonyms. The same is true when one verb is substituted for another in (40).

Notice that formula (40) contains fifteen sentences that have to be learned. However, since (40) has only one meaning associated to one syntactic form, it seems plausible to assume that only one unit has been learned and that this unit is independent of the meaning of the words it contains. This assumption is supported by the symmetrical role played by the words inside each pair of brackets. The temporal sequence for learning (40) will privilege one arbitrary sentence of (40), depending on circumstances; then the rest of the sentences can be

learned in any order, by adding new *words* in the corresponding pairs of brackets. With this *linear* way of counting, in order to learn (40), that is fifteen sentences, one has to learn:

1 unit (structure and meaning) + 8 words

The number of sentences is immaterial to the problem; for example, if we could add 2 verbs + 2 nouns to (40), the number of sentences would grow to thirty-five, a numerical increase that has not much to do with the amount of new information added.

We consider that the process presented is general, and that it also applies to free sentences (and to the free positions of frozen sentences). But for this to be so, one has to make the following hypotheses:

- free sentences are derived from frozen basic forms, by 'thawing' syntactic positions. For example, the sentence:[31]

  42. *Max sang Frère Jacques.*

  comes from a frozen form such as:

  (*A person* + *a singer*) *sings a song.*

  with a cognate or internal subject and object;
- variable nouns are introduced by means of the classificatory sentences that describe the external universe:

$$\begin{pmatrix} My\ sister \\ Max \\ A\ boy \end{pmatrix} is\ a\ person.$$

  Frère Jacques is a song.[32]
  This noise is a motet.

- classificatory sentences are combined with thawed sentences by the syntactic processes of relativization and redundancy reduction proposed by Harris (1968):

  The person who is Max sings the song which is Frère Jacques.
  – *The person Max sings the song Frère Jacques.*
  – *Max sings the song Frère Jacques.*
  – *Max sings Frère Jacques.*[33]

[31] Such abstract forms could be arrived at by induction from actual examples.

[32] The sentences with indefinite determiner *a* are more basic in the sense that they do not introduce the same type of information as the sentences with definite determiners, which require further analysis.

[33] The first two steps of this derivation are abstract intermediaries. For the reduction occurring in the last two steps, see also Gouet (1976).

5.2.2 *Stages in learning*

Various observations have been made concerning the existence of a crucial period in the development of language in an individual (as distinct from historical development). Returning to the study of the lexicon-grammar, we can study the vocabulary (i.e. the rows) and the syntactic forms (i.e. the columns) from the point of view of their acquisition in time.

The list of verbs includes 'elementary' verbs such as TO CRY, TO EAT, TO WANT, and more sophisticated ones, such as TO APPROXIMATE, TO IDENTIFY, TO INFER. Whereas it is clear that elementary verbs have been acquired and used in early childhood, it is unlikely that the others were learned as early as that. Investigation of the classification of verbs does not hint in any way at some dividing line that would separate the two types. On the contrary, there seems to be a continuum on a scale that would range from an intuitively primitive to a very sophisticated vocabulary.

When we examine the syntactic structures in the same way we are able to point at early structures such as those of *Bob slept*, or *Give me a cake*, but also at types such as *I asked that they be on time* that have presumably been learned later. But the situation is not as clear for syntactic structures as it is for lexical items. To the extent that there exists in this case a consensus among linguists as to which forms (intransitive, imperative, subjunctive, etc.) are legitimate syntactic structures, our choice of examples is well founded. But questions may well be raised about sentences such as:

A family without a grandmother is not a family,
A meal without red wine is not a meal,

which have the following syntactic structure:

43. N_0 *without* N_1 *is not* N_0

The two occurrences of N_0 represent a repetition of nouns; there is a semantic relation between N_0 and N_1. Some linguists may object that the proverbial flavour attached to these forms excludes them from the range of syntax: they would belong to rhetoric. But there is no formal basis for such a distinction. No serious method for the separation of structures has ever been proposed that leads to denying (43) the status of a syntactic structure, whereas there would be no question, for example, about the syntactic legitimacy of:

Max went from N_1 *to* N_1 —:
Max went from city to city.
Max went from store to store.

In the same way, consider the structure:

44. *It is* N_0*'s turn to* V W —: *It is Bob's turn to go to the market.*

where N_0 V W —: *Bob goes to the market* is a full sentence. This structure carries a special use of *turn* and a special meaning, which is not found in any transformed sentences.[34] Again, one could call this form a rhetorical process. However, the way the sentence form N_0 V W is split into two parts is analogous to the action of the general clefting form *It is . . .* WH-, as applied in:

It is Bob who goes to the market.

There is no reason why (44) should be a rhetorical device, hence excluded from syntax, whereas cleft sentences are to be full-fledged syntactic forms.

There seems to be a rather large number of templates like (43) and (44) that cannot be reduced (e.g. transformationally) to well-known simple structures. They are thus fully autonomous, both syntactically and semantically. We can then with some plausibility assume that they are learned one at a time; some, like (44), are learned in early childhood, others, more literary, such as (43), are learned much later. It becomes possible to view all syntactic forms as learned piecemeal and progressively. We are led to adopt this theoretical position by the fact that we have failed to observe any discontinuity among structures which would reflect separate layers in the language. Hence, we have no evidence internal to syntax that would suggest the existence of successive stages of development of grammar in children.

When we said that some lexical item or some structure is learned early or late, it should be clear that we meant earlier or later than the crucial periods of exposure that have been put into evidence in various animals, and that have been suggested by Chomsky and by Piaget to be related to the acquisition of language and other intellectual skills.

Thus, learning sentence forms and words, one by one, appears to be a process that goes on fairly regularly during one's entire lifetime. Furthermore, it is hard to believe that the items (words and structures) that have been suggested to be early or late acquisitions cannot be learned in the reverse order by children.

The external evidence in favour of discontinuous learning stages for language is quite limited; among the facts involved are the behaviour of a few 'wild children' and the vague and general observation that learning a second language appears to involve a large amount of conscious work, which would not be the case for the first language. It seems to us that the application to both arguments of the intuitive concept of motivation might be sufficient to account for the

[34] It is hard to see how (44) could be related synchronically to *They took turns going to the market*, where, intuitively, *turns* seem to be the same lexical item as *turn*.

corresponding linguistic limitation. Though it may be technically difficult to objectivize common-sense reflection, the theoretical proposals we have made reinforce the plausibility of our main thesis, namely that rote learning of whole sentences is an essential factor in the acquisition of the native languages.

5.2.3 *Learning frozen sentences*

In section 5.1 we discussed a possible association beween verbs and semantic predicates. Since the association is based on the meaning of each verb, the discussion of section 5.1 has practically no relevance for frozen sentences: their verbs have no meaning. One is then led to consider that they are learned one by one, or in sets, as indicated in section 5.2.1.

Many questions can be raised about the stability of the form and its meaning. Consider the example:

Max took the bull by the horns.

It translates word for word into a frozen sentence with the same meaning in the following languages: French, Italian, Portuguese, Spanish, Danish, Swedish, German, Polish, and Finnish, and presumably in other languages as well. The expression is old enough to make us wonder why a complex association between such a specific choice of words and such a specific meaning has been so well preserved through generations and across independent languages. Certain diachronic problems could be given new life in this context. For example, the fact that frozen expressions are well formed today indicates that expressions going back to Old French or to Latin have evolved syntactically: they have acquired determiners, adopted a new tense system, etc. Since frozen items are earmarked in a certain sense, they could be followed in time more easily than variable items.

Sentences with support verbs are analogous to frozen sentences with respect to the association of form and meaning discussed in section 5.1. Although it may be claimed here that the supported nouns provide a basis for semantic interpretation, this is not true for the support verbs themselves, which do not introduce any selectional restrictions. Again, we consider that they are learned one by one, or in series constituted by equivalent support verbs, that is, as in section 5.2.1.

5.3 *The function of syntax*

The importance of the lexicon of frozen forms forces us to reconsider the function of syntax in language. In generative grammar syntax is a

device used for semantic interpretation on the one hand, and for phonological interpretation on the other. In frozen expressions, syntax is at work, building correct utterances, but without contributing to their semantic interpretation. Instead, we must have a direct association between the lexical clues and some abstract units of meaning. To the extent that syntax, although regular with frozen expressions, does not play any role in their interpretation, we feel that it is legitimate to question the role of syntax in the interpretation of free sentences.

Also, one may ask what the function of frozen forms is as a component of language. In a sense, frozen forms are a burden: each of them often has a nearly synonymous free form that is shorter in number of words, hence easier to memorize. That frozen expressions are highly redundant in the lexicon is even more striking with frozen adverbs: they are made up of several words and have a highly restricted range of meanings (section 3.2); several hundred of them can readily be replaced by fewer than ten single words.

One may then hypothesize that the principal function of syntax is prosodic. When one speaks, or writes a first draft, one often has the feeling that sentences about to be produced are off balance for some rhythmic reason, and that they should be modified. A reflex seems to introduce more material in such sentences: modifiers, predeterminers such as (*a* + *the*) *type of*, adverbs that emphasize a point, etc. One way of lengthening sentences for this purpose might be by using frozen expressions.[35] The main function of frozen expressions would then be to adjust sentences to a certain satisfactory combination of structure and length, a purely prosodic function regulated through synonymous substitutions and syntactic changes.

In view of these possibilities, we can only speculate about a semantic function for syntax. For indeed, syntax does have a recognized rhythmic function; it introduces a periodic structure in utterances, and this prosodic organization of sentences may be crucial for the memorization of the lexicon-grammar. The role of prosody in memorization has long been recognized in poetry and as a tool in oral literatures. We might attribute a somewhat similar role to syntax, though certainly not as a conscious invention of man.

[35] In this respect, Harris (1976) has shown that prefixes and suffixes range over a small set of meanings and that they have an abbreviating function. Such a device could also be used for balancing the length of sentences.

6. CONCLUSION

Linguists who look at generative linguistics and at its offsprings as a scientific activity have often confused the hypothetico-deductive activity of theory building with the general pursuit of knowledge. Science establishes facts and deals with established facts. Theorizing a body of facts may take many forms, depending on the fields and on the facts. But increasing the range of established facts is a universal activity largely independent of theories. In every science accumulating data is by far the commonest activity. Thus, astronomers compile spectrographic measurements, from as many regions of the sky as they can explore and for as many ranges of frequencies they can instrument, looking for unusual effects for which theories will be built or adjusted a posteriori. Data collected by both Voyagers, for example about the rings of Saturn, directly provided new facts that are valued for their own sake: they increase our knowledge about Saturn and the form of the solar system; they will probably not modify theories of gravitation, and will presumably be explained within Newton's theory. Molecular biologists analyse as many nucleic acids as their time, methods, and equipment allow. No formal theory is involved here: the problem consists in determining the functions of proteins, in tracing their source to genes, with the hope that accumulating this type of knowledge will result in an understanding of some aspects of life. When a theory is proposed in the framework of molecular biology, it is so concrete that no linguist could recognize it as such.

Linguists commonly believe that operating within a formal theory is a condition *sine qua non* for achieving a scientific status. This attitude is also common among philosophers of science, who take fundamental physics as a model for the relation between empirical observations and theories or explanations. Linguists have adopted this paradigm, equating the exceptional success of theoretical physics with the generality of scientific procedures. In contrast, we think that linguistic theories should have more modest and immediate goals, than constructing a universal grammar that explains learning and generalizes the neural sciences.

When we construct lexicon-grammars, we consider that we are following the example of such undisputed sciences as biology, chemistry, or the earth sciences. In these fields formal mathematical theories are still the exception, and current work consists in accumulating empirical knowledge that leads to theories of a *concrete* nature. A well-known example is Wegener's theory; it took about thirty years for scientists to accept the theory that continents move

apart from each other. The theory became accepted only because a large number of measurements forced earth scientists into adopting this solution. In this case, the theory is highly concrete and can be reduced to a simple idea that does not need any mathematical or formal expression. And this theory is remarkably explanatory.[36] Needless to say, the fact that a theory is explanatory and the fact that it is abstract, formal, or mathematical are totally independent characteristics; whether quantum theory, which is a model of abstraction and formality, is explanatory is a moot question. For example, there have been abstract discussions about memory and the process of learning, but no one has attempted to evaluate the size of the syntactic and lexical information contained in the memory for a whole language, as if this parameter could not affect the possible theories of learning.

One of the reasons why no such attempts have been made is a belief in a strange philosophical a priorism: it is claimed that counting lexical items and their syntactic properties has no effect on syntactic theory, or no explanatory value, hence is of no interest. As a matter of fact, the very term 'explanatory theory' does not seem to be used in any other domain than in linguistics. The term is quite confusing, and largely redundant, especially when one looks at domains like physics or chemistry where a new explanation is seldom a new theory; in general, explaining a phenomenon consists in accommodating it within an already existing theory.

The arguments used by linguists in favour of 'explanatory theories' and against taxonomic linguistics would not be understood by any active scientist, simply because, for such a scientist, enumerating facts systematically is an important method for discovering new phenomena. Moreover, if enumeration of facts is possible and easy, no theoretician would take the risk of proposing a theory if accessible data had not already been gathered. Methods exist for describing large parts of the lexicon and of the grammar of several languages, but they are not applied; abstract speculation is the pre-eminent linguistic activity. We think that this overwhelming emphasis on theoretical activity is out of place when not one grammar is as yet available of a language as much studied as English.

REFERENCES

Benson, M., Benson, E., and Ilson, R. (1986), *The BBI Combinatory Dictionary*, Amsterdam-Philadelphia: J. Benjamins.

Bonnard, H., Leisinger, H., and Traub, W. (1970), *Grammatisches Woerterbuch-Franzoesisch*, Dortmund: Lensing.

[36] Cf. *Science*, 13/4,503 (3 July 1981): *Solid Earth*.

Boons, Jean-Paul (1971), 'Métaphore et baisse de la redondance', *Langue française*, 11/15: 16.
— (1974), 'Acceptabilité, interprétation et connaissance du monde', in *Actes du colloque franco-allemand de grammaire transformationnelle*, vol. ii, Tübingen: Max Niemeyer (Eng. trans. in *Cognition*, 2/2: 183–211).
— Guillet, Alain, and Leclère, Christian (1976*a*), *La Structure des phrases simples en français*, i: *Constructions intransitives*, Geneva: Droz.
— — — (1976*b*), *La Structure des phrases simples en français*, ii: *Constructions transitives*, Paris: Rapport de recherches du LADL, no. 6.
Boström, I. (1957), *Les Noms abstraits accompagnés d'un infinitif et combinés avec* avoir, Lund: C. W. K. Gleerup.
Caput, J., and Caput, J.-P. (1969), *Dictionnaire des verbes français*, Paris: Larousse.
Cattell, Ray (1984), *Composite Predicates in English*, Syntax and Semantics 17, New York: Academic Press.
Chomsky, Noam (1969), 'Remarks on Nominalization', in Jacobs and Rosenbaum (eds.), *Readings in Transformational Grammar*, Waltham, Mass.: Blaisdell.
— Schützenberger, Marcel-Paul (1963), 'The Algebraic Theory of Context-Free Languages', in Braffort and Hirschberg (eds.), *Computer Programming and Formal Systems*, Amsterdam: North Holland Pub. Co.
d'Agostino, Emilio (1983), *Lessico e sintassi delle costruzioni locative*, Naples: Liguori.
Danlos, Laurence (1980), 'Représentation d'informations linguistiques: Constructions *N être Prép X*', doctoral thesis, Université Paris VII.
Dubois, Jean (1969), *Grammaire structurale du français*, Paris: Larousse.
Elia, Annibale (1984), *Le Verbe italien: Les Complétives dans les phrases à un complément*, Bari: Schena-Nizet.
Giry-Schneider, Jacqueline (1978), *Les Nominalisations en français: L'Opérateur FAIRE dans le lexique*, Geneva: Droz.
— (1987), *Étude de prédicats nominaux en français: Les Constructions* faire N, Geneva: Droz.
Gouet, Michel (1971), 'Lexical Problems Raised by Some of the *Foutre* Constructions', in *Studies in Left out Fields*, Edmonton: Linguistic Research Inc.: 79–85.
Gross, Gaston (1989), *Les Constructions converses du français*, Geneva: Droz.
Gross, Maurice (1968), *Grammaire transformationnelle du français*, i: *Syntaxe du verbe*, Paris: Larousse, Cantilène.
— (1972), 'On Grammatical Reference', in Bierwisch, Kiefer, and Ruwet (eds.), *Generative Grammar in Europe*, Dordrecht: Reidel.
— (1975), *Méthodes en syntaxe*, Paris: Hermann.
— (1979), 'On the Failure of Generative Grammar', *Language*, 55/4: 859–85.
— (1981), 'Les Bases empiriques de a notion de prédicat sémantique', in A. Guillet and C. Leclère (éds.), *Formes syntaxiques et prédicats sémantiques*, Languages 63, Paris: Larousse: 7–52.
— (1982), 'Une classification des phrases figées du français', *Revue québécoise de linguistique*, 11/2: 151–85.

Gross, Maurice (1990), *Grammaire transformationnelle du français*, iii: *Syntaxe de l'adverbe*, Paris: ASSTRIL.

Guillet, Alain, and Leclère, Christian (1991), *La Structure des phrases simples en français*, ii: *Constructions transitives à compléments locatifs*, Geneva: Droz.

Harris, Zellig S. (1946), 'From Morpheme to Utterance', *Language*, 22: 161–83.

— (1951), *Method in Structural Linguistics*, Chicago: University of Chicago Press.

— (1952), 'Discourse Analysis', *Language*, 28/1: 1–30.

— (1964), *Elementary Transformations*, TDAP 54, Philadelphia: University of Pennsylvania (repr. in *Papers in Structural Transformational Linguistics*, Dordrecht: Reidel, 1970: 482–532).

— (1968), *Mathematical Structures of Language*, New York: Wiley (Fr. trans. by C. Fuchs, *Structures mathématiques du langage*, Paris: Dunod, 1971).

— (1976), *Notes du cours de syntaxe*, Paris: Le Seuil.

— (1982), *A Grammar of English on Mathematical Principles*, New York: Wiley-Interscience.

Hornby, A. S., and Cowie, A. P. (1989), *The Advanced Learner's Dictionary of Current English*, London: Oxford University Press.

Jespersen, Otto (1961–5), *A Modern English Grammar*, London: Allen & Unwin.

Labelle, Jacques (1974), 'Étude de constructions avec opérateur AVOIR (nominalisations et extensions)', doctoral thesis, Université Paris VIII.

Lafaye, Bernard (1841), *Synonymes du français*, Paris: Hachette.

Leclère, Christian, and Subirats, Carlos (1991), *A Bibliography of Studies on Lexicon-Grammar*, *Lingvisticae Investigationes* 25/1, Amsterdam-Philadelphia: J. Benjamins BV: 347–409.

Lightner, Theodore M. (1983), *Introduction to English Derivational Morphology*, *Lingvisticae investigationes supplementa* 6, Amsterdam-Philadelphia: J. Benjamins BV.

McCawley, James (1979), *Adverbs, Vowels and Other Wonders*, Chicago: University of Chicago Press.

Macedo Oliveira, Maria Elisa de (1979), *Syntaxe des verbes psychologiques du portugais*, Lisbon: Instituto Nacional de Investigaçaõ Cientifica.

Martinelli, Maurizio (1984), 'Sur la distinction entre compléments de verbes et de phrase', *Lingvisticae investigationes*, 8/1: 195–9.

Meunier, Annie (1977), 'Sur les bases syntaxiques de la morphologie dérivationnelle', *Lingvisticae investigationes*, 2: 287–332.

— (1981), 'Nominalisation d'adjectifs par verbes supports', doctoral thesis, Université Paris VII.

Molinier, Christian (1982), 'Étude syntaxique et sémantique des adverbes en *-ment*, doctoral thesis, University of Toulouse-le-Mirail.

Négroni-Peyre, Dominique de (1978), 'Nominalisation par *être en* et réflexivation (*admiration, opposition, révolte et rage*)', *Lingvisticae investigationes*, 2/1: 127–64.

Ranchhod, Elisabete (1990), *Sintaxe dos predicados nominais com* estar, Lisbon: Instituto Nacional de Investigaçaõ Cientifica.

Rosenbaum, Peter S. (1967), *The Grammar of English Predicate Constructions*, Cambridge, Mass.: MIT Press.

Ruwet, Nicolas (1978), 'Une construction absolue du français', *Lingvisticae investigationes*, 2/2: 165–210.

Salkoff, Morris (1983), 'Bees are Swarming in the Garden: A Synchronic Study of Productivity', *Language*, 59/2: 288–316.

Subirats, Carlos (1986), *Sentential Complementation in Spanish*, *Lingvisticae investigationes supplementa* 14, Amsterdam-Philadelphia: J. Benjamins BV.

Summers, D., *et al.* (1987), *Longman Dictionary of Contemporary English*, London: Longman.

Vasseux, Philippe (1982), 'Le Système LEXSYN de gestion des donnés lexico-syntaxiques du LADL', Rapport de recherche du LADL, Paris: LADL.

Vergnaud, Jean-Roger (1985), *Dépendances et niveaux de représentation en syntaxe*, *Lingvisticae investigationes supplementa* 13, Amsterdam-Philadelphia: J. Benjamins BV.

Vivès, Robert (1983), 'Avoir, prendre, perdre: Constructions à verbe support et extensions aspectuelles', doctoral thesis, Université Paris VIII.