

MAURICE GROSS

OBSERVATIONS ON SEMANTIC THEORIES

A general framework for the relations between syntactic forms and semantic elements is discussed. Examples are taken from traditional and transformational grammar. A more detailed analysis is given of a family of verbs (verbs of transfer: *to give*). Interpretation rules are proposed. The notion of quantifier is extended. The problem of the nature of semantic elements is raised in connection with generative semantics.

1. Introduction

I consider that the issue under discussion brings together several areas of intellectual activity which ought to be kept disjoint and that the present situation, which is a direct result of this, has been determined by factors external to linguistics. It is essentially a search for novelty which is at stake, a novelty that has been confused with new discovery or progress.

Specialists of mathematical logic have found in natural languages motivations for extending their formal activities normally devoted to the description of mathematical theories.

As linguists lost interest in descriptive work, theorization became their activity, and so, they were bound to be attracted by logical formalisms.

And philosophers of knowledge could see in the intersection of these two fields a domain where they might base some of their theories.

It is in this zone of converging interests that problems appear to arise. I believe that definite barriers still exist between these three fields and that their recent merging has introduced pseudo-questions, in a domain where clear-cut discussions had already separated facts and methods of quite different natures.

Meanwhile, the methodological situation has not basically changed in the last thirty years (i. e. since Harris 1946 introduced systematic formalization). A brief résumé of this situation should indicate where serious doubts arise as to the existence of any relation between natural language and formal logic, and where the problems of representing knowl-

edge are, when one attempts to extract it from natural language forms. We will discuss these questions within the framework of the universal model for the syntax and semantics of languages, either natural, or artificial as in mathematical logic and in computer science. On the one hand, we have a set F of syntactic forms¹, on the other a set M of elements of meaning. A correspondence between F and M is a set of interpretation rules for the formulas of F in terms of the elements of M (cf. figure 1).

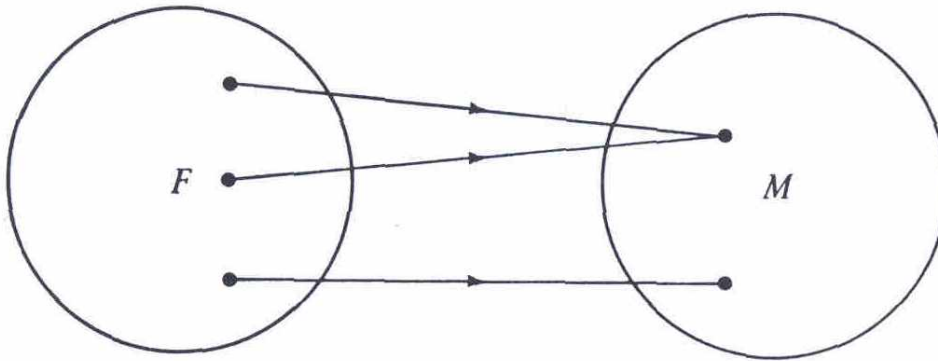


Figure 1

The set F is generated by a grammar. The set of the generators of F contains elementary forms such as words, parts of speech and types of phrases. The generators are combined by combinatorial rules (Chomsky 1957). Combinatorial processes consist of three basic types of operations: insertion, erasure and permutation. The reason for the success of formal grammar as compared to traditional grammar should be clear: concepts of form (i.e. categories and rules of construction) are operational², and lead to reproducible judgments of acceptability over a wide range of phenomena. Hence, among syntacticians, the agreement on the empirical adequacy of the combinatorial approach is almost complete.

But in semantics, the situation is quite different and the set M has virtually no status. There has not yet been any proposal about the content of M that has not proven controversial. Many authors propose abstract elements for M , and empirical data are in general limited to syntactically trivial examples (e.g. *Max saw a unicorn*, *Bachelors are unmarried men*, *The king of France is bald*, etc.³) Lack of agreement between semanticists

¹ A syntactic form is, as usual, a sentence together with some amount of abstract structure.

² This is less true for certain abstract parts of speech and phrases.

³ The problems that have been raised about such examples may not be trivial from a semantic point of view.

is the direct consequence of the non reproducibility of the concepts proposed *a priori* for the interpretation of the elements of *F*.

The main semantic theories admitted so far can be sketched as indicated in figure 1:

- traditional grammars have as elements of *M* terms like “time”, “place”, “cause”, “action” that are associated with syntactic forms such as nouns, (prepositional)phrases, verbs, etc. Case grammars are special types of traditional grammars in which the elements of *M* are “genitive”, “dative”, etc.; they are used to interpret subjects and complements to which are attached special morphemes (e.g. case endings, prepositions). Traditional grammars were intended to provide for straightforward intuitive resonance, that is, for a certain amount of satisfaction for the reader faced with grammatical statements. This is not true with cases, which are left undefined and are so abstract that no reproducible meaning can be triggered by them;

- we discuss below generative semantics;

- logics are also used for the interpretation of sentences. The elements of *M* are logical formulas, which may be true or false; many attempts to introduce modalities have failed. The correspondence between *F* and *M* consists in a translation procedure of sentences into more or less standard formulas of predicate calculus.

We will now discuss in a precise way certain features of these different treatments of meaning and will provide detailed examples of the schema of figure 1. It is hoped that this discussion will clarify the so-called issue of “logically-based vs autonomous” syntax.

2. Formalizing the traditional approach

Semantic notions have often been attached to verbs. For example in French, the subjunctive mood in sentential complements (e.g. in the substructure *V que S*) is linked to verbs of will (e.g. *vouloir* (to want)), of doubt (e.g. *douter* (to doubt)), of fear (e.g. *craindre* (to fear)). The notions “will”, “doubt”, “fear” are elements of *M*, and the first question to be asked is how they have been selected. As a matter of fact, the arbitrariness of analogous notions should appear in the following examples.

We first consider the semantic notion “transfer of an object from one person to another”, which incidentally involves the term “dative complement”. The sentence

Max gave a bed to Lea

typically involves this element of meaning which will be noted $\mathcal{T}(p, q, o)$: $\mathcal{T}(p, q, o) \in M$, p is the sender, q is the receiver and o the object. We will study the syntactic forms of French (i.e. the elements of F) that may be associated with this element of meaning. We choose French for the reason that it is the only language for which systematic data are available. A reasonably complete taxonomy exists with respect to the criteria of form involved here (Boons, Guillet, Leclère 1976a, 1976b; Gross 1975).

The sentences

- (1) *Max a donné un lit à Léa*
(Max gave a bed to Lea)
- (2) *Max a pris un lit à Léa*
(Max took a bed from Lea)

have identical forms: same preposition(s) and same three syntactic positions; we write these as $NP_0 V NP_1 \text{ à } NP_2$, in which only the verb changes. With respect to $\mathcal{T}(p, q, o)$, 1 and 2 are different, hence we use the following interpretation rules:

- for (1), $V = \text{donner}$: $NP_0 = p$; $NP_1 = o$; $NP_2 = q$,
- for (2), $V = \text{prendre}$: $NP_0 = q$; $NP_1 = o$; $NP_2 = p$.

We can now present the problem in a somewhat more precise fashion.

The semantic intuition $\mathcal{T}(p, q, o)$ is reasonably well reproducible over the French lexicon of verbs. Its relevance is thus established.

Parameters of form are the following:

- (a) verbs have 2 object complements,
- (b) prepositions are limited to “zero”, *à* and *de*,
- (c) the left-to-right order of the noun phrase varies, but the subject is kept “human”.

The number of a priori possible cases is then 8: NP_0 can be p or q , NP_1 can be either p or q or o , NP_2 can take either *à* or *de*.

We have already given 2 cases, and four more can be found:

Max a reçu un lit de Léa
(Max received a bed from Lea)

where $NP_0 = q$; $NP_1 = o$; $NP_2 = p$,

Max a abonné Léa à cette revue
(Max subscribed to this journal for Lea)

where $NP_0 = p$; $NP_1 = q$; $NP_2 = o$,

Max a doté Léa d'un lit
(Max endowed Lea with a bed)

where $NP_0 = p$; $NP_1 = q$; $NP_2 = o$,

Max a escroqué Léa d'un dollar
(Max cheated Lea of a dollar)

where $NP_0 = q$; $NP_1 = p$; $NP_2 = o$.

We have thus observed 6 cases out of 8. Two positions can be taken with respect to the 2 gaps:

(i) the gaps are significant. The corresponding redundancy must then be expressed in the rules (of grammar or of interpretation);

(ii) the gaps are accidental. Then, the situation is not different from what it would have been, had a correspondence been observed between $\mathcal{T}(p, q, o)$ and the 8 a priori possible cases.

We will take the stand (ii), since we find no reason to believe that principles exist that can exclude the missing forms. The consequences of this attitude are straightforward: the parameters of form we have chosen for study (a, b, c above) are independent of meaning.

This remark is not based on the single example just presented; many others have been studied in a systematic way, and they lead to the same conclusion. For example, let us consider what may be called a "psychological" predicate $\Psi(s, b)$, where s is a stimulus that results in a response of a human being b ; we will restrict our examples to "positive" responses (i.e. s triggers something "good" for b) in order to make our problem more specific, that is, to diminish the number of cases possibly involved. It is natural to investigate the syntactic forms

$NP_0 V(Prep) NP_1$

and perhaps only these. A priori, s and b can be found in 2 positions: NP_0 and NP_1 ; once more, we let $Prep$ vary over the 3 main cases: "zero", *à* and *de*. Hence, 6 forms may possibly be related to the element of meaning $\Psi(s, b) \in M$. Examining the lexicon of French verbs, we find that 5 out of the 6 cases are attested. With $s = ceci$ (this) and $b = Max$, examples are

Ceci (intéresse + réjouit) Max
(This (interests + delights) Max)

Max (aime + apprécie) ceci
 (Max (likes + appreciates) this)
Ceci plait à Max
 (This pleases Max)
Max (tient + penche) à ceci
 (Max (clings + inclines) to this)
Max (exulte + jubile) de ceci
 (Max rejoices at this)

Other examples also suggest, at least as a first approximation, that the correspondence between syntactic forms (F) and meaning (M) is arbitrary, an arbitrariness that can be viewed as an extension of the principle of the arbitrariness of the sign, commonly accepted at the level of the correspondence between the morphophonemic constitution of words and their meaning.

However, one can speculate over this failure to associate form and meaning in an interesting (i.e. non trivial) way:

– It can be argued that this negative result is due to improper choice of the elements of meaning M . This may be the case, but then the content of M would be a total mystery, to a degree where the term semantics would lose all meaning. It seems to us that sentences have to be interpreted in terms of INTUITIVE concepts, scenarios or the like. If the elements of M are not $\mathcal{T}(p, q, o)$, $\Psi(s, h)$, etc., one will nevertheless have to come close to them. The main traditional attempts used more abstract elements, and they failed. One of the reasons is that a special terminology always had to be devised for M ; then, the reproducibility of the association form-meaning was invariably restricted to the author of the system. Also, with respect to the question of language acquisition, it seems reasonable to assume that the semantic units available to the child are of the above naive type.

– It should be noted that the statistical distribution of the elements of F over the French lexicon is asymmetrical with respect to their semantics. For example, in the case of $\mathcal{T}(p, q, o)$, the form $NP_0 V NP_1 \dot{a} NP_2$, when restricted to $NP_1 = o$, corresponds to the majority (80%) of the verbs of transfer; the two subcases $NP_0 = p$ (i.e. “give”) and $NP_0 = q$ (i.e. “take”) are about equally distributed. The situation is similar with “psychological” verbs. The interpretation of the asymmetry cannot be settled simply. An important part of it could be due to a purely random effect (Schützenberger, personal communication). But some notion of productivity is surely involved. For example, are we to consider that the most frequent

form in the lexicon (i.e. $NP_0 V NP_1 \dot{\rightarrow} NP_2$) is the general or regular case⁴ and that all other forms found are accidents, such as remnants of old forms or marginal literary experiments. It will be difficult to answer this question. First, forms such as *recevoir* (to receive) although rare in the lexicon belong to the current vocabulary, they are used quite commonly. How is their importance to be weighed? Second, in the construction of *escroquer* (to cheat) above, the verb can be replaced by many other slang verbs which are regularly invented. How is the importance or generality of these verbs to be evaluated? Under such conditions, decisions about the extension of the gaps will be hard to make.

We can also ignore such statistical arguments, as in traditional and generative grammar. We have then to consider that all forms are of equal importance for the expression of semantic predicates such as $\mathcal{T}(p, q, o)$ and $\Psi(s, b)$: any meaning predicate can be associated with any sentence form.

Remark on the interpretation rules

In both cases presented, the rules were simple, because the number of "syntactic arguments" of the verbs was identical with the number of arguments of the semantic predicates. The correspondence between elements of F and of M could then be stated in the form of sets of equations. However, this type of association will be hard to maintain in certain cases:

French sentences can be represented in the following general form:

$NP V NP^n$ with n presumably finite

We take as the definition of syntactic argument the property for an NP to be cleftable into *C'est . . . Qu* (It is . . . that). Various corrections have to be added for the sentential complements and for certain noun complements (Gross 1976). The syntactic arguments are then easy to detect and to count. But there are examples such as

Le lit passe de Max à Luc
(The bed passes from Max to Luke)

⁴ If the answer is yes, then again the difference of meaning (i.e. the sense of the transfer of o) between these two types cannot be connected to a form which is unique.

which may have to be interpreted as predicates $\mathcal{T}(p, q, o)$. Clefting applies here to the whole complement sequence *de Max à Luc* and not to each of the two NP's included in it:

- C'est de Max à Luc que le lit passe*
 (It is from Max to Luke that the bed passes)
 **C'est de Max que le lit passe à Luc*
 (*It is from Max that the bed passes to Luke)
 **C'est à Luc que le lit passe de Max*
 (*It is to Luke that the bed passes from Max)

It seems that such cases have to be treated in an *ad hoc* way.

There are also examples of the type

- Max habille Léa = NP V NP*
 (Max dresses Lea)

which are synonymous with forms that have three arguments, according to our definitions:

- $\equiv$ *Max (donne + met) des habits à Léa = NP V NP à NP*
 [Max (gives + puts) (clothes + dresses) (to + on) Lea]

But there, a syntactic relation (i.e. a relation in *F*: Harris' operators 1964) may relate the two types of forms.

3. Formalizing generative semantics

The idea of generative semantics is to decompose complex predicates into simpler ones. Consider the verb *donner* (to give). It is felt as synonymous with some composition of causative and of possessive: to give something to someone is close to causing someone to own something. Now, in French, we observe a syntactic operation with causative meaning, as in

- La soif a fait: Max boit du vin*
 (Thirst caused: Max drinks wine)
La soif a fait boire du vin à Max
 (Thirst caused Max to drink wine)

We can apply the operation with a verb of possession in the second sentence:

Léa a fait: Léo a un lit

(Lea caused: Lea has a bed)

Léa a fait avoir un lit à Léo

(Lea caused Leo to have a bed)

The resulting sentence has the meaning of

Léa a procuré un lit à Léo

(Lea provided Leo with a bed)

a type of meaning already encountered: $\mathcal{T}(p, q, o)$.

We thus have on the one hand a (causative) form (*faire* $\in F$) to which can be attached a semantic predicate $\mathcal{C}(p, x)^5$ ($\in M$) where p causes x . On the other hand, there are sentence forms with a meaning of possession: $\mathcal{P}(q, o)$, where the person q owns the object o . Various forms are in correspondence with $\mathcal{P}(q, o)$, in the sense of the preceding paragraph. We find for example⁶

Max a un lit (Max has a bed)

Ce lit appartient à Max (This bed belongs to Max)

Max dispose d'un lit (Max has a bed at his disposal)

Syntactic composition occurs in F and is paralleled in M , as in figure 2. The situation is then the following: sentences with *avoir* (to own) have $\mathcal{P}(q, o)$ for image in M and those with *faire* (to cause) are associated with $\mathcal{C}(p, x)$; $\mathcal{T}(p, q, o)$ must then be dependent on these two predicates and we must therefore define a composition operation in M and write, for example, $\mathcal{T} = \mathcal{C} \oplus \mathcal{P}$. But the empirical status of the basic predicates compared to

⁵ When such semantic predicates are compounded, some of the variables may have to be identified. Here, we even have to identify the variable x with the predicate $\mathcal{P}$. Thus, binding problems also arise in non-logical frameworks. Binding techniques can be borrowed from fields foreign to linguistics (i. e. mathematical logic, computer science), and then be used in a natural way, but specific problems affect natural languages only (e.g. coreference: Fauconnier 1974, Postal 1971).

⁶ There are other examples which, though somewhat removed from this meaning, include some notion of possession:

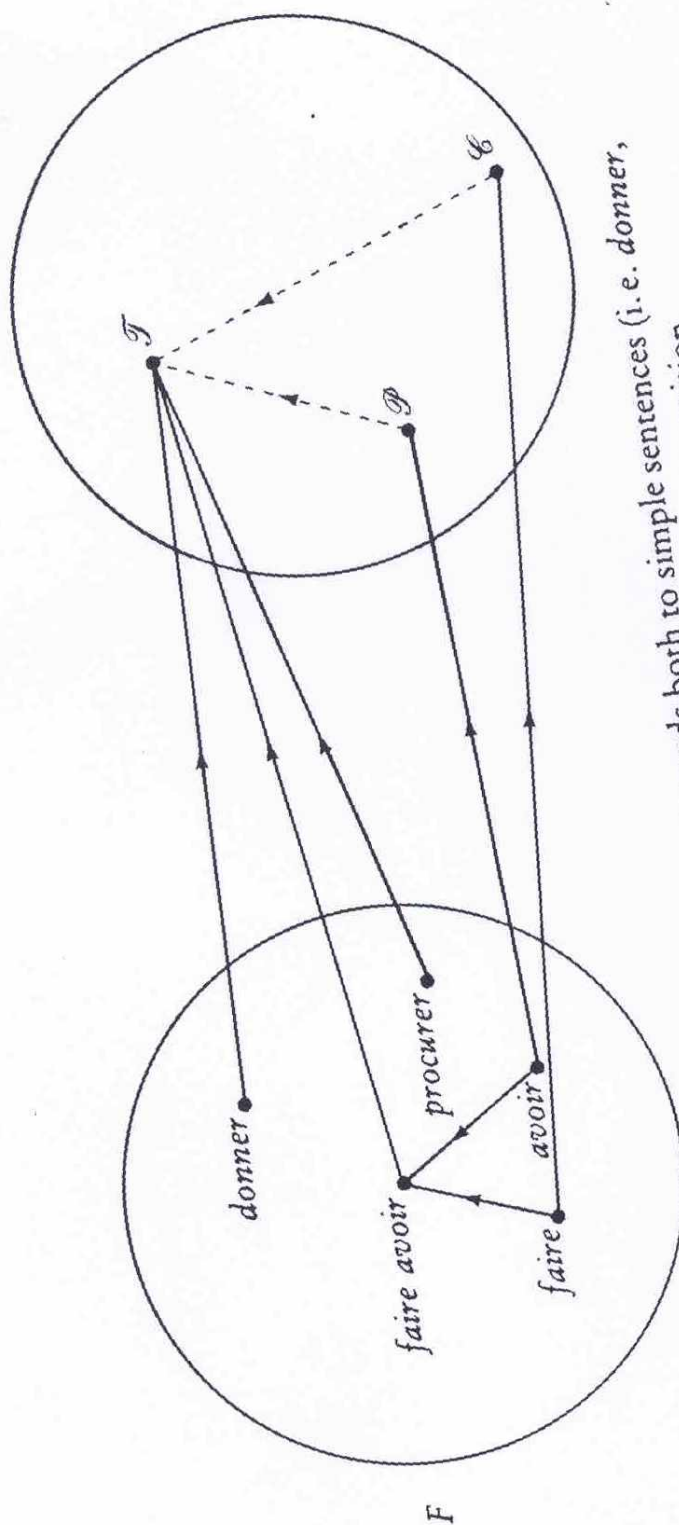
Max prétend à ce lit

(Max claims this bed)

Ce lit dépend de Max

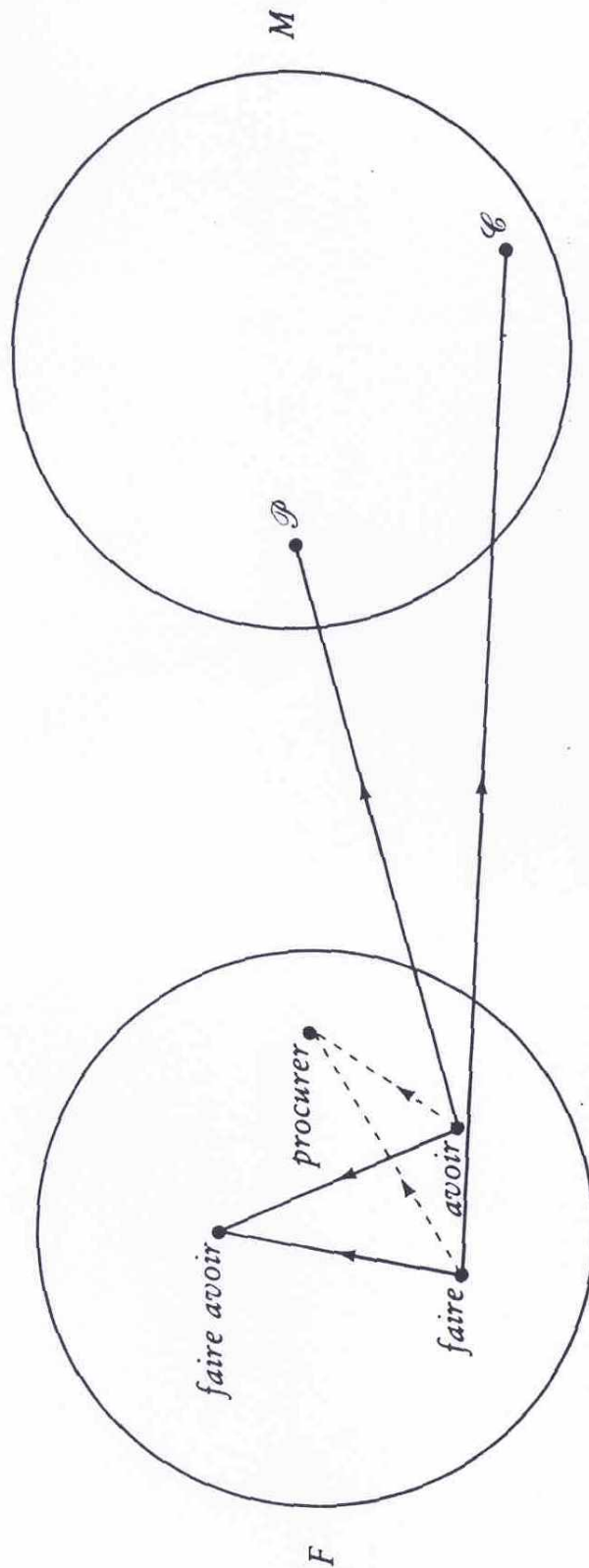
(This bed depends on Max)

If they are related to $\mathcal{P}(p, o)$ 5 forms out of 6 are found. The only gap would be something close to *Ce lit affecte Max* (This bed affects Max).



The predicate $\mathcal{T}(p, q, o)$ corresponds both to simple sentences (i.e. *donner*, *procurer*) and to sentences obtained by syntactic composition.

Figure 2



The predicate $\mathcal{T}(p, q, o)$ is no longer used for interpretation.

Figure 3

the compound ones is undetermined: the only basis for separating them is a fuzzy intuition probably based on the syntactic properties of the corresponding verbs. We are led to duplicate in M the composition operation of F . A possible solution may be generative semantics: the composition process would be limited to F , and M would contain only basic predicates of meaning, as exemplified in figure 3.

The discussions of generative semantics present an ambiguity that has never been eliminated. A general issue is in fact involved: Should the elements of M be abstract entities or actual words, both of which are basic in some sense and thus restricted in number and complexity? In the few examples treated in generative semantics, terms such as BELIEVE, CAUSE, NOT are used to describe sentences. They are considered, it seems, as abstract terms. But they have concrete (i.e. combinatorial) properties: exactly those of the corresponding words. Moreover, their names convey "concrete" meanings (i.e. they are directly translated into the corresponding words). Thus, on the one hand, there are reasons to consider these terms as abstract elements used to generate F and at the same time as elements of M used to interpret sentences.

Remarks

1) In our evaluations of meaning, we have ignored many intuitions. For example, the meaning of *procurer, faire avoir* (to provide) excludes that $NP_0 = p$ owns $NP_1 = o$, this is not the case with *donner* (to give). One should ask on what grounds such differences have been eliminated. As a matter of fact, the difference between *to kill* and *to cause someone to become not alive* has been used as a criticism of generative semantics. One can argue in favour of simplification on the ground that we have to limit ourselves to reproducible meaning. In certain instances, there exist nuances of meaning which are such that the requirement of reproducibility is hard to satisfy.

2) Our wording and the drawing of figures 1–3 suggest that we are constructing a mapping from F to M . However, we could just as well have constructed the inverse mapping from M to F (or to the set of the parts of F). We also said that F was generated independently of M . All these formal remarks can be related to the opposition "logically-based" vs "autonomous" syntax. It should be clear that the schema presented is neutral with respect to the discussion. Moreover, it is difficult to imagine any kind of empirical data that could enable us to discriminate among such general proposals.

4. Formalizing logical approaches

The association of logical predicates with syntactic forms can be placed in the same general framework.

For example, relativization has attracted the attention of logicians. Many papers deal with translation procedures for sentences such as

The men who are tall are here

A set is attached to the phrase *The men*, another set of men to the phrase *who are tall*, and their intersection represents the meaning of the subject. Or else, the same operation is performed in terms of logical formulas with quantifiers. This activity contains in itself the seeds of innumerable variants of calculus and the sterile controversies that go with them: there are as many logical systems as there are logicians, maybe more. Why then choose one system rather than any other? Notice that not only does this kind of choice have no bearing on the determination of *M*, but at the same time, fundamental linguistic problems raised by relative clauses and quantifiers are overlooked.

Compare for example the two sentences

- (1) *Max met the man who is here*
- (2) *Max met the man who is tall*

Sentence 1 is normal in the sense it can be interpreted without any context; sentence 2, although syntactically identical, can be considered as unacceptable with respect to 1. Its interpretation requires at the very least a contrastive situation that will have a logical translation different from that of 1. It is significant that a phenomenon of this size has been missed by logicians and linguists busy with formalizing relative clauses⁷. It is no surprise then that the logical approach does not meet with much success.

Another example of gross empirical inadequacy is the treatment of quantifiers. Consider the pairs

- (3)
 - (a) *Max drinks large amounts of beer*
 - (b) *Max drinks beer in large amounts*
- (4)
 - (a) *Vast quantities of gold have been sold*
 - (b) *Gold has been sold in vast quantities*

⁷ Kuroda has mentioned the sources of 1 and 2.

They have a wide distribution and can be observed for practically all verbs and in practically all their arguments. Their properties present important regularities, but they are never discussed. These pairs involve a large family of quantifiers that are in fact a subset of a much more general family of modifiers. We observe for example

- (5) *Eva's faces amused Lea*
Eva amused Lea with her faces
- (6) *Max liked the taste of this food*
Max liked this food for its taste

There is here a formal difference between these pairs, as compared with the case of quantifiers: the complement *Prep NP* contains a possessive adjective obligatorily coreferent in some sense to the subject or to a complement of the verb:

**I amused Lea with her faces*

**Max liked this food for my taste*

With quantifiers, this effect is present but not marked. All these pairs present slight variations, depending on the verbs and nouns (Gross 1975, Leclère 1976), but they share many features such as adverbial mobility, clefting and behaviour with respect to passive. These properties point to the unicity of a wide-spread but unstudied phenomenon. In contrast with these data, the efforts of linguists and logicians have been concentrated on the four words *all*, *each*, *some*, *any* and on their interaction with negation and operations such as Passive. It has not occurred to any specialist that these words might have a teratological behaviour. They have always been singled out in grammars because of their exceptional morphology. When they are compared to the quantifiers of 3 and 4, it appears that their syntax is exceptional too. The properties so far observed of these four exceptional items have no reason whatsoever to be significant for language structure.

Our position should now be clear: syntactic methods are far from having exhausted the range of observable phenomena. Considering the paucity of present empirical data, discussions of abstract levels appear to be devoid of empirical content. This situation where no syntactic control can be placed upon the choice of abstract semantic elements has sometimes

resulted in surprising proposals⁸. For example, nothing prevents the use of autobiographical data: when an author is deeply convinced from his own psychic experience that three or four concepts are enough to describe the Universe, it may be hard to adduce any argument tending to demonstrate that his view is erroneous. Only detailed empirical studies can indicate the inexactitudes of such proposals.

BIBLIOGRAPHY

- BOONS, J.-P., GUILLET, A. and CH. LECLÈRE (1976a), *La structure des phrases simples en français, I – Constructions intransitives*, Droz, Geneva.
- BOONS, J.-P., GUILLET, A. and CH. LECLÈRE (1976b), *La structure des phrases simples en français. II – Classes de verbes transitifs*, Travaux de recherches du L. A. D. L., Université Paris-7, Paris.
- CHOMSKY, N. (1957), *Syntactic Structure*. Mouton, La Haye.
- FAUCONNIER, G. (1974), *La coréférence: Syntaxe ou Sémantique?* Le Seuil, Paris.
- GROSS, M. (1975), *Méthodes en syntaxe*. Hermann, Paris.
- GROSS, M. (1976), Sur quelques groupes nomiaux complexes. In *Méthodes en grammaire française*, J.-C. Chevalier et M. Gross eds, Klincksieck, Paris, pp. 97–120.
- HARRIS, Z. S. (1946), From Morpheme to Utterance. *Language* 22, 161–82.
- HARRIS, Z. S. (1964), The Elementary Transformations. T. D. A. P., Univ. of Pennsylvania. In *Papers in Structural and Transformational Linguistics*, Reidel, Dordrecht, 1970.
- KURODA, S.-Y. (1968), English Relativization and Certain Related Problems. *Language* 44, 2–1, The Waverly Press, Baltimore.
- LECLÈRE, CH. (1976), Datifs syntaxiques et datifs éthiques. In *Méthodes en grammaire française*, J.-C. Chevalier et M. Gross eds, Klincksieck, Paris, pp. 73–96.
- POSTAL, P. (1971), *Cross-over Phenomena*. Holt, Rinehart and Wilson, New York.

⁸ Notice that the introduction of pragmatics does not shed much light on the nature of *M*. We even think that it will blur the possibility of determining a semantic level based on empirical data and which could thus be investigated systematically.