FORMAL SEMANTICS OF NATURAL LANGUAGE

Papers from a colloquium sponsored by the King's College Research Centre, Cambridge

EDITED BY
EDWARD L. KEENAN
Senior Research Fellow, King's College, Cambridge
(now Associate Professor, Department of Linguistics, UCLA)

1975

CAMBRIDGE UNIVERSITY PRESS
CAMBRIDGE
LONDON . NEW YORK . MELBOURNE

On the relations between syntax and semantics

MAURICE GROSS

One of the main objectives of traditional grammarians was to relate form and meaning. This programme ran into many difficulties and was abandoned by structural linguists who found it much more fruitful to concentrate on the voluntarily limited study of the combinatorial properties of words.

Transformational linguists also exclude meaning from the grammar rules they build. However, the definition of a transformational rule (unlike the definition of a distributional rule) explicitly involves meaning, since transformationally related sentences must have identical meanings.[1]

There are important differences in the ways we just referred to the term 'meaning'. Traditional grammars classify forms into families, and attribute to these families absolute categories of meanings.[2] For example, the notion of phrase is a notion of form, so is the notion of *when*-phrase (i.e. adverbial phrase whose left-most word is *when*). Often, the semantic notion /time/ is associated with these forms (i.e. adverbs of time).

The modern formalized version of this activity is usually stated in the general framework of formal logic. On the one hand, the syntactic rules of some formal system[3] define a set of well-formed formulae (here sentence forms), on the other hand, a semantic model provides interpretation for each formula. As in mathematical logic, the question of setting up a dividing line between the syntactic theory and its model constantly arises.[4] In both

[1] The definition of distributional rules involves meaning implicitly. Meaning is then part of the global notion of acceptability. Transformationally related sentences may have systematic differences of meaning. For example, one may consider that the declarative sentence *John gave a book to this girl* and the corresponding cleft ones *It is John that gave a book to this girl, It is a book that John gave to this girl, It is to this girl that John gave a book* are transformationally related. Between the source sentence and each of the cleft ones, we observe the same difference: /emphasis/, /contrast/, or the like.

[2] The names of these categories will be written between strokes.

[3] This attitude is by no means the only possible one. As Chomsky has pointed out, performing syntactic descriptions in the framework of the formal systems of mathematical logic implies a particular hypothesis that may turn out to be empirically inadequate. In fact Z. S. Harris (1951:372–3) who first proposed it, has moved towards using algebraic systems, which, owing to their more abstract character, eliminate the possibility of raising certain questions which may not make any linguistic sense (e.g. zeroing of morphemes, directionality of a transformation).

[4] The problems raised by generative semantics relate closely to this question.

the traditional and the formal approach, absolute notions of meaning are needed to interpret the sentences.

Generative grammars provide numerous examples of this approach, and many empirical data, old or discovered within this framework, have been described from just a syntactic point of view. However, despite all kinds of efforts, the study of the semantics of natural languages remains an entirely open field. Many proposals of models of interpretation have been made, but none of the most basic questions has been answered yet. All examples are quite limited with respect to the range of semantic units that come immediately to mind and that seem relevant to semantic descriptions. In fact, they all seem to suffer from the same defect: the lack of empirical basis, and often, a not very careful study of the notions involved raises serious criticisms that may put in question the whole validity of this approach.

For example, most (if not all) traditional grammars associate the notion /time/ to *when*-phrases. But in sentences like *When John makes a mistake, he is unhappy* which are synonymous with *If John makes a mistake, he is unhappy* it is by no means clear why one should attach the concept of /time/ to the phrase, rather than any of the notions /implication/, /condition/, /concomitance/. This type of criticism is fairly general, and applies to all such associations of meaning and form.

Similarly, the semantic notions /true/ and /false/, so widely discussed and formalized in the context of the relations between logic and linguistics, do not seem to have an indisputable empirical basis. For example, it is widely assumed that the sentence

(1) I know that Max has arrived

'presupposes'[1] that the proposition: *Max has arrived* is /true/, while in the sentence

(2) I believe that Max has arrived

the same proposition can be either /true/ or /false/. This difference in 'presupposition' has been attributed to the main verb (i.e. *to know* vs. *to believe*). However, it is by no means clear that the difference that we just observed really holds. In a discourse like *Max's hat and boots are in the entrance, so I know that he has arrived* the subject *I* may have been mistaken by certain

[1] The terms 'to presuppose', 'presupposition', although not defined, appear to be used in a technical sense. They are simply to be interpreted as 'to mean partially', 'a part or component of meaning', respectively. For an evaluation of this notion see Kuroda (1973).

clues, and it could be the case that the proposition *Max has arrived* is /false/.

The standard view, namely the first observations we made on (1) and (2), may be correct, but presumably under quite complex linguistic conditions that have not been determined so far. Unless such conditions are clearly stated, the notions /true/ and /false/ may just be as inadequate as the notion /time/ was.

Thus, the absolute notions of meaning that are needed for interpretation, and that have been proposed so far (e.g. notions of time, space, and truth) all appear to be empirically inadequate. Moreover, it is far from obvious to imagine how one should proceed to determine some of them, and how they could be motivated on any empirical and theoretical ground.

The way the semantic notions currently discussed have been arrived at is quite clear. Grammarians and philosophers have performed observations on syntactic classes of sentences or of phrases. Their reading triggered intuitions in the mind of the investigators, and the intuitions were given names that were supposed to reproduce corresponding intuitions appearing in the mind of other students of the same forms. These names are of two main kinds, either words taken from the vocabulary of the language under study, or else they are abstractions whose meaning is technical and defined elsewhere (e.g. logical implication). These two naming activities correspond to two different theoretical attitudes, both easy to criticize: on the one hand one does not see why elementary semantic units should have an observable counterpart (i.e. words) in a natural language, on the other, one does not see why the semantic units should be the ones that constitute the basis of a technical language (e.g. logic) built for reasons that do not have much to do with the study of natural languages. Again, we are faced with the basic empirical problem of semantics: what is a semantic fact?

We already mentioned that there are in generative transformational grammars manipulations of meaning that are of a different kind from the one we just criticized. There, pairs of sentences[1] that are candidates for being related by a transformation are judged to be synonymous or not. Thus, meaning is only involved in comparisons, and *differences* in meaning are detected in this manner. In the physical sciences, it is well-known that *absolute* evaluations of a variable (e.g. temperature) lead always to rather crude results, when compared to *differential* evaluations of the same variable. The situation appears to be the same in linguistics with respect to meaning. Attributing absolute terms to forms is quite problematic, and anyway, has proved to be rather unsuccessful, while comparing the

[1] The status of these pairs should be that of minimal pairs in phonology.

meanings of similar forms may bring to light subtle differences that may be hard to detect directly. This situation has allowed transformational grammarians to handle certain aspects of meaning. But the question of providing interpretations for sentences in terms of units of meaning is not solved. These elements of meaning which have been extracted by differential tests still have to be given names that will make explicit the interpretation of sentences. Such units may have a good empirical adequacy, but the problems we mentioned about naming still remain, and it is hard to see any solution for them.

While attempting to construct a syntactic classification for a large set of simple sentence types (Gross (1975)) we encountered various correlations between forms and meanings which suggested that an empirical study of absolute notions of meaning, while more difficult and less precise than the study of differential units, might not be out of reach. In fact, the main criticisms that have been made about the use of absolute notions are essentially based on the fact that it is always very easy to find counter-examples to any statement involving absolute notions. One of the causes of this situation is that no systematic study of any syntactic phenomenon has ever been made for a natural language. All studies are quite fragmentary, and they only affect a small part of the lexicon, so that it is quite obvious that in no respect is it possible to base a statement on data that reasonably cover a natural language; whence the ease with which one can find counter-examples to statements that are always much too general in comparison with the few examples from which they are extracted. In certain areas of syntax, the study we made avoids this difficulty to a large extent, so that finding counter-examples to our statements will not be as easy as it usually is.

We took as a test case the distribution of about 150 syntactic properties over a lexicon of about 6000 French verbs. The properties we chose turned out to be such that they were relevant to about half of the verbs. This study deals essentially with French complementizers, namely with verbs accepting in at least one of their syntactic positions (i.e. subject or object(s)) at least one of the forms:

que P (*that S*, in English)
si P ou non (*whether S or not*, or *if S or not*)
VΩ (infinitive *VP*)

The 3000 corresponding verbs have been classified mainly according to their pattern of complement(s), which resulted in the definition of 19 classes each containing between 20 and 300 verbs.

Each class has been represented by a matrix. On each row there is a verb, and each column corresponds to a syntactic property. When a verb (i.e. its construction as defined by the class) has a given property, a plus sign is placed at the intersection of the corresponding row and column; a minus sign is placed in the opposite case.

Our syntactic properties are of two types: distributional properties and transformational ones. Some of the distributional properties are clearly semantic and their operational value is rather low. For example, the distinction /human/ vs. /non-human/ has so far turned out to be of little interest, since there are numerous verbs for which there is no sharp distinction (or no distinction at all) between these two terms. Thus /human/ nouns can sometimes be interpreted as /non-human/ subjects, like *brother* in the sentence *My brother functions well.* The reverse is quite frequent too, for example with nouns used as /containers/ of /human/, e.g. *street* in *She amused the whole street.* Again the distinction /human/ vs. /non-human/ does not seem to be relevant to the interpretation of the complement of *to look for.*

Other distributional properties lead to much sharper distinctions. Thus, the distribution of a phrase with a sentential modifier like *the fact that John did it* classified our verbs in a sharply reproducible way, presumably because the occurrence of this phrase is much less dependent on the meaning of its head noun (e.g. we have *I know the facts*, but **I know the fact that John did it*).

The transformational properties that were studied are the ones that are currently found in the literature. In most cases, these properties have been deduced from a small number of examples. The study of a large number of cases led us to revise most of them, introducing new conditions on them, and sometimes revising significantly their formulation. In fact, transformations are only indirectly represented in our tables. Each syntactic property is a structure that a verb may enter into or not. For example, the structure

(A) NP V NP

is such a property, and the corresponding structure

(P) NP est Vpp par NP

with interchanged NPs is another property. It is the pair [(A), (P)] that defines the passive transformation. This definition of the tables allowed us to represent non-transformational relations between different constructions of what ought to be considered as containing the same verb. For example, we observed the existence of pairs like

(3) a. Que Max soit venu a irrité Luc
(That Max came irritated Luke)

and

b. Luc s'est irrité auprès de Guy de ce que Max soit venu
(Luke told Guy that he was irritated that Max came)

Sentences (a) and (b) differ in meaning, but they share several syntactic features: the distributional and semantic nature of the subject in (a) is the same as the one of the *de*-complement in (b), the direct object of (a) is also identical to the subject of (b). Sentence (b) has three arguments,[1] where (a) has only two, and the extra complement *auprès de Guy* which cannot occur in (a)

(4) *Que Max soit venu a irrité Luc après de Guy

adds to the verb the meaning of /saying/. Thus, there does not seem to be any possibility of relating (a) and (b) by transformational means. We could consider that there exist two verbs: *irriter* and *s'irriter*, this position could perhaps be justified by the fact that there are constructions identical in form and similiar in meaning to (b), and involving verbs that do not have the (a) construction:

(5) Luc a protesté auprès de Guy de ce que Max soit venu
*Que Max soit venu a protesté Luc
(6) Luc s'est plaint auprès de Guy de ce que Max soit venu
*Que Max soit venu a plaint Luc

Also, there are verbs that have the (a) but not the (b) construction:

(7) Que Max soit venu a ennuyé Luc
*Luc s'est ennuyé auprès de Guy de ce que Max soit venu

However, through a systematic study of the lexicon, we observed that among the 500 verbs that we described by means of the property

(8) Que S V NP

about 40 also had the construction

(9) NP se V auprès de NP de ce que S

and this correspondence seems to be quite productive, namely it can be extended to other verbs used in (8) in a figurative meaning. On the other

[1] We call arguments (of the verb) the subject and the complement(s).

hand, verbs like *protester* that only have the (9) construction are not numerous, we observed fewer than 40 such examples, most of them obtained by a rather difficult extension of some other use of the verb. A typical case would be *soupirer* (*to sigh*) in the sentence

(10) Luc soupire auprès de Léa de ce qu'elle ne vienne plus chez lui

which, although easily understandable, could be rejected by many native speakers as unacceptable.

Thus the solution of two lexical entries does not seem to be justifiable, mainly since it does not capture the relation between (8) and (9). We choose to indicate the relation since our system of representation allows it in a natural way: in our matrix (8) and (9) will be independent properties, and individual verbs like *irriter* will have both.

Another typical example of a non-transformational relation involves pairs of constructions like:

(11) Paul a hurlé à Jean qu'il viendrait
(Paul shouted to Jean that he would come)
(12) Paul a poussé un hurlement
(Paul gave a shout)

The constructions (11) and (12) are related morphologically: *hurlement* is a nominal derived from the verb *hurler*; but there is a syntactic and semantic correspondence too, in both sentences: *Paul* is the subject. However, it seems hard to derive (11) from (12) since we observe that the constructions do not have the same complements:

(13) a. *Paul a poussé un hurlement à Jean
b. *Paul a poussé un hurlement qu'il viendrait
c. *Paul a poussé un hurlement à Jean qu'il viendrait

similarly, (12) cannot be derived from (11), i.e. from the substructure *Paul hurle* of (11), since *hurlement* can have determiners and modifiers that are not found with *hurler*; for example the source of *un grand nombre* in the sentence

(14) Paul a poussé un grand nombre de hurlements stridents

would be hard to justify.

There is however another observation that indicates the existence of a relation between (11) and (12). We have listed about 150 verbs of /saying/ that have the construction of *dire* (*to say*); among them, 40 have the associated construction (12). In (12), the nominalizing suffixes are highly

restricted: *-ment* in 38 cases, and zero in 2 cases. Moreover the pairing is practically nonexistent outside of the class of the verbs of /saying/. We consider this situation as strongly supporting the existence of the indicated relationship. We thus treat certain nominalizations as processes that relate two elementary sentences (which is quite different from the solution in Lees (1960)). The relation is not transformational in Chomsky's sense (Chomsky (1967)), but is considered as such by Harris (1964) who considers the verb *pousser* as an operator acting on a sentence.

Many such cases have led us to make more precise our notion of syntactic property and its relation to transformations. All properties appearing in the columns of our matrices are structures that a verb of a given row enters into or not. Thus, as mentioned, a transformation is a pair of columns. Such pairs are most of the time ordered in generative grammar. But we prefer to consider them as defining a relation,[1] that is, as non-ordered. The effect of the relation is to produce a classification of sentence types; between the classes further relations can be defined (Harris (1968)).

The choice of the syntactic properties is primarily determined by the operational quality of the tests that are used. Thus, the property for a verb *V* to enter into a passive form or not when it enters into the construction *NP V NP* provides sharp distinctions among verbs in a large number of cases. As a result of our choice of properties, we have a reasonable guarantee that the classes that we have defined are purely syntactic, taking into account the fact that all traditional attempts, like for example the attempt to relate the existence of passive forms to semantic properties of verbs have always failed.

In a number of cases, it came as a surprise that all the verbs of some of these syntactic classes triggered a common semantic intuition.

For example, we have defined a syntactic class by means of the following properties:

1. the verbs have a direct object, roughly speaking they enter into a structure (P_1): *NP V NP* (*Luc apprècie Max*) without entering into a larger structure like *NP V NP à NP* or *NP V NP de NP* where *à NP* and *de NP* are indirect objects;
2. the *NP* direct object can be the sentential complement *que S* where the main verb of *S* is in the subjunctive, i.e. the verbs enter into the structure (P_2): *NP V que S* (*Luc apprécie que Max soit venu*);
3. the verbs enter into the structure (P_3): *NP V NP de VP* (*Luc apprécie Max d'être venu*) related to the structure (P_2) by the raising (?)

[1] 'Relation' is to be taken in its technical sense.

relation *que* $S \rightarrow NP$ *de* VP; in (P_3) no nominal NP can be substituted for VP (i.e. **NP V NP de NP*).

These properties, when conjoined, isolate in the French lexicon a set of 60 verbs which all trigger the semantic intuition of /ethical judgement/. This situation should be surprising since in the general case the classes that can be defined by similar syntactic properties are semantically heterogeneous.

In our study, 4 classes out of 19 turned out to be semantically homogeneous; the 3 others are the following:

We defined the class of verbs V entering into the construction $NP\ V\ VP$, where NP is the subject of the infinitive VP, and where VP can be replaced by the interrogative pronoun *où*. All the members of this class are in some sense verbs of movement from one place to another. E.g.

(15) Guy (descend/court) voir Max

and:

(16) Question: Où Guy (descend/court)-il?
Answer: Voir Max

while for example, with the same structure:

(17) Guy (aime/doit) voir Max

we do not find the dialogue:

(18) Question: Où Guy (aime/doit)-il?
Answer: *Voir Max

The class we have described contains about 120 such verbs of movement.

A second class, related to the preceding one, is defined by the construction $NP\ V\ NP\ VP$, where the second NP is the subject of the infinitive VP, and where this VP can be replaced by the interrogative pronoun *où*. Most of these verbs[1] can be interpreted as /causative of movement/. We have

(19) Pierre envoie Guy voir Max
(20) Question: Où Pierre envoie-t-il Guy?
Answer: Voir Max

[1] *Accompagner* in *Luc accompagne Guy voir Eve* is in the class, but is paraphrasable by *Luc va avec Guy voir Eve*, hence it is not /causative/.

and this construction of *envoyer* is synonymous with the causative construction

(21) Pierre fait aller Guy voir Max

where *aller* belongs to the preceding class.

A final class that we have defined corresponds to the construction

(22) Que S V NP

where the *NP* has a /human/ head noun, and where the subject of *V* is semantically unrestricted, namely its head noun may belong to any semantic class. In particular we have

(23) Que Guy soit arrivé (amuse/ennuie) Max.

All these verbs are verbs of /sentiment/: the unrestricted subject triggers a feeling in the /human/ direct object.[1] It may be added that this semantic description can also be applied to the class of verbs of /ethical judgement/ already mentioned, but with reversed syntactic relations: the unrestricted direct object triggers a feeling in the /human/ subject.[2]

These examples are by no means accidents, and there are other cases of subclasses of verbs (with respect to our classification) such that a syntactic definition leads to a set of verbs which are all semantically related. For example we can define a class by the structural property:

(24) NP V que S à NP

all of its verbs must have a sentential direct object, and an indirect object with the preposition *à*. This class is quite heterogeneous, but a subclass of these verbs defined by the following properties is homogeneous:

1. the verbs have a sentential direct object *que S* in an indicative form that undergoes equi-*NP* deletion, when the subject of *S* is co-referential with the subject of *V*:

(25) a. Je dit à Max que je me suis évanoui

⇒

b. Je dit à Max m'être évanoui

[1] There are semantically analogous verbs which are syntactically different. For example *plaire* (*to please*) has an unrestricted subject too, but an indirect /human/ object with the preposition *à*.

[2] In certain associated constructions the order of the arguments is similar, though reversed here. We have for example: *Luc hait Eve*/*Luc a de la haine pour Eve* but *Eve dégoute Luc*/*Luc a du dégout pour Eve*.

2. the verbs also have a sentential direct object in subjunctive form that undergoes equi-NP deletion when the subject of S is co-referential with the indirect object of V:

(26) a. Je dit à Max qu'il s'en aille

⇒

b. Je dit à Max de s'en aller

The verbs that are defined by these properties are all verbs of /saying/, and, as mentioned, there are about 150 of them.

Although the semantic notions appearing in each class are 'absolute' notions, they are perceived in a remarkably consistent way by all native speakers. This remark is the basis of what we call semantic homogeneity. But these classes are remarkably homogeneous too from a purely syntactic point of view. The notion of syntactic homogeneity that we are attempting to define is based on observations made on our classification.

As we have already observed, our material can be viewed as a binary matrix of 3000 by 150. Each description of a verb (i.e. each row) has been transferred to a punched card, so that computer programs (Bely and Vasseux (1973)) can easily extract from the general matrix various types of subclasses. For example, we have constructed the set of classes which is such that each class contains only verbs that have exactly the same syntactic properties. For 3000 verbs, we obtain 2000 classes, and when we studied the classes containing more than one verb, we noticed that it was easy to find new syntactic properties that divided these classes into further subclasses containing only one verb. Thus, we can assert that in French there are no two verbs that have exactly the same syntactic properties. Examination of the columns leads to a similar observation: there are no two syntactic properties that have the same distribution over the lexicon. As a consequence all relations between sentences, whether transformational or not, have exceptions. This picture of the syntactic structures of a natural language indicates that a large number of irregularities are present. How to deal with them is not clear,[1] but it is fairly obvious that the irregularities are by no means randomly distributed. In order to separate 3000 verbs in 2000 classes, 12 properties are sufficient, but we had to use more than 150 of them to obtain this result. This indicates that a large

[1] In Gross (1975) we suggested an approach to this problem which arises under nearly the same conditions in phonology (Lightner (1972)). Other similar studies have been performed with similar results: Boons et al. (1975) have studied 4000 verbs without sentential arguments, Giry (1972), 1500 nominal constructions of the type *faire N* (e.g. *faire des compliments*, *faire la vaisselle*), and Picabia (1970), 1200 adjectives with sentential subject and/or sentential complement.

amount of redundancy is present in our matrix, and in particular that the contents of certain rows on the one hand, and of certain columns on the other, are related. Such relations appear to be more numerous for the semantically natural classes than for the other cases. Examples of these numerical differences are given in Table I where it is quite apparent that the semantic classes that we have defined have a high proportion of constant columns, i.e. of columns containing either +signs or −signs.

TABLE I

	No. of constant columns	Total no. of columns
*V*s of /ethical judgement/	17	30
*V*s of /movement/	15	28
*V*s /causative of movement/	20	28
*V*s of /sentiment/	4	19
*V*s in *NP V que P à N*	2	45
*V*s in *NP V que P*	0	36

The ratios of the number of constant columns to the total number of columns are respectively 0.5, 0.5, 0.7, and 0.25 for the semantic classes, but −0.05 and 0 for the syntactic classes. Such differences seem to be significant. The fact that the ratio is relatively low for the *V*s of /sentiment/ is due to the large number of *V*s in this class (i.e. 540), and to the presence of morphological rather than syntactic properties. It is important to remember that the semantic notions that we have deduced are not *characterized* by the syntactic properties we have been using. There are verbs which semantically are covered by our definitions, but that do not appear in our syntactic classes. For example: the verbs of /ethical judgement/ can also be tagged verbs of /sentiment/, and these two types do not have the same syntactic properties. Secondly, the verb *marcher* (*to walk*) is not in the syntactically defined class of verbs of /movement/, although semantically it does not seem very different from *courir* (*to run*). And thirdly, in the construction already mentioned

(3) b. Luc s'est irrité auprès de Guy de ce que Max soit venu

s'irriter is semantically a verb of /saying/. However, its syntactic structure: *NP V auprès de NP de ce que P* is rather different from the one of *dire* (*to say*) (indirect sentential complement, and a different preposition for the receiver of the message).

At any rate, our examples appear to be clear cases of relations between absolute semantic notion and syntactic properties. However, the way the

relationship should be described is by no means obvious. Our observations could be stated in the following general way:

Let S_1, S_2,..., S_i,...be absolute semantic notions. Let P_1, P_2,..., P_j,...be syntactic properties. The rules that relate meaning and shapes are of the form

(R) $\Sigma \rightarrow \Pi$

where Σ and Π are boolean combinations of the S_is and the P_js respectively.

A few remarks about such rules can already be made that may turn out to be quite general. In our examples, Σ was reduced to one S_i (e.g. /sentiment/) or to two (/causative/ of /movement/), that is, Σ was composed of a small number of very intuitive notions. On the other hand Π was made of a rather large number of syntactic properties (i.e. the columns that were constant in a class), and these properties are by no means intuitive: it is very hard (if at all possible for the moment) to imagine a process by which they could be learned by a child or an adult. If the rules (R) had further specific properties, they might be considered as the basis of an empirically adequate theory of learning.[1]

However, even the few examples that we have presented force us to look at the rules (R) from several different points of view, without there being, for the moment, any empirical way of deciding how to formulate them in a precise way.

Thus, each of the semantic classes that we defined seems to contain one verb which, in some sense, is semantically minimal: for the verbs of /ethical judgement/ *aimer* (*to like*) would be this element; for the verbs of /movement/ we would consider *aller* (*to go*) to be minimal; and for the verbs of /saying/ we would take *dire* (*to say* or *to tell*).

In all these cases, the other verbs of the class (or a large majority of them) would be interpreted with the meaning of the minimal element, together with some extra notions that would account for the difference of meaning. But there are many ways of expressing this situation.

Let us consider the verbs of /saying/. In French, practically any verb that can be interpreted as corresponding to an emission of sound or of

[1] Element S_i like /sentiment/ may have to be considered as combinations of simpler S's; the same often happens with syntactic properties. Thus the difference in size of the two members of a rule R may not be essential; what is crucial for a theory of learning is to explain how the very complex non-conscious properties P_j are acquired. We propose that the child starts from the conscious and intuitive, hence cultural S_i's, and that a universal process associates these with the syntactic properties.

light can be used syntactically and semantically as a verb of /saying/.[1] We have for example sentences like

(27) Luc (bégaie/ronronne) à Guy (de venir/qu'il est ici)
(Luc stammers/purrs) to Guy (to come/that he is here))

Such sentences can be paraphrased by:

(28) Luc dit à Guy (de venir/qu'il est ici) en (bégayant/ronronnant)

Moreover, the construction (28) appears to be acceptable only with *dire* (the minimal verb) as main verb:[2]

(29) *Luc bégaie à Guy (qu'il est ici/de venir) en ronronnant
(30) *Luc ronronne à Guy (qu'il est ici/de venir) en bégayant

One might describe these restrictions by means of a standard transformational solution: thus the syntactic properties that have been observed for the verbs of /saying/ would only be attributed to the verb *dire*. All other verbs of /saying/, namely all verbs that indicate an emission of sound or of light, would be considered as intransitive verbs. Independently, it can be observed that *dire* is the only verb which has obligatory complements (at least in declarative sentences), i.e. which cannot be considered as intransitive. Then a transformation would relate (27) and (28) by deleting the minimal element *dire* and by inserting the intransitive verb in the former position of *dire*. Morphological adjustments then take place. Such a solution accounts for the meaning and the syntactic properties of the set of described structures. Moreover it is morphologically simpler than an analysis along the lines of generative semantics (Joshi (1972), McCawley (1968), Postal (1970), Ross (1971)), aside from the grammatical elements (*en*, *-ant*), the only zeroed element is the root of *dire* which is recoverable since in our sense the minimal verb is lexically unique.

The same type of solution could be used for the verbs of /movement/:[3] all the verbs except (minimal) *aller* can be considered as intransitive and the same transformation as above could be used, which results in:

(31) Max va voir Guy en courant
⇒
(32) Max court voir Guy

[1] Also constructions that describe certain gestures can be used as verbs of /saying/: *Luc fait signe de la tête à Guy* (*qu'il est ici*/*de venir*) (*to nod*).

[2] Constructions of the type (28) might be acceptable with other verbs than *dire*, provided that some kind of semantic inclusion holds between the two verbs of /saying/.

[3] Notice that sentences like *Max* (*monte*/*descend*) *voir Guy en courant* are acceptable. This observation is to be related to the remark of n. 2, above.

The verbs of /ethical judgement/ cannot, however, receive a similar 73
treatment. We do not observe the same relations between the minimal verb *aimer* and the others. Most of them cannot be used intransitively, and if they could, they would not yield the paraphrase observed before:

(33) *Max aime Guy en adorant
(34) *Max aime Guy en haïssant
⇒
(35) Max hait Guy

Rather, these verbs seem to be compounded of *aimer* and some adverbial adjunction. This hypothesis is supported by the fact that practically all of them cannot be used with certain adverbs, like, for example, the comparative *mieux* (*better*):

(36) *Max adorerait mieux venir
(37) *Max haïrait mieux venir

while with *aimer*, we do have

(38) Max aimerait mieux venir

Also, under the same conditions, negations are difficult to accept at least without contrastive effect:

(39) ?*Max n'adore pas venir
(40) ?*Max ne hait pas venir

while *Max n'aime pas venir* is acceptable.

Thus, the verbs of this class could perhaps be viewed as composed of *aimer* and some /intensity/ adverbial and/or a negation. Such an analysis resembles the analyses proposed within the framework of generative semantics. However, we are not concerned with the theoretical problems raised in this context. We only want to point out that the first type of solution that we have proposed is not quite general, and that processes which involve factorization of words may have to be used also in order to describe the structure of the lexicon.

Thus, we seem to be advocating here an empirical approach to semantics that is largely based on syntax, but the separation between syntax and semantics has never been very sharp. It is clear that distributional and transformational grammars are all based on combinatorial processes acting on morphemes, while the rest of the study of language is called semantics.[1]

[1] It seems even more difficult to separate pragmatics from semantics.

This dichotomy, essentially based on various formal properties, seems very hard to justify, and linguists have noticed that many problems arise within such a framework, which, to a certain extent, do not seem to be significant. As a consequence, we have shifted our main interest to the empirical data, considering that the study of natural languages should comply with the rules which all experimental sciences obey. One of our main concerns is the reproducibility of data. Most (if not all) of the data in syntax originates from experiments that consist in building strings and checking whether they are acceptable or not. Such experiments turn out to be reproducible in a large number of cases. Other experiments are possible that use other types of intuitions than the intuition of acceptability. As already mentioned, most of the experiments that use (absolute) semantic intuitions have turned out to be non-reproducible, and this has led structural linguists to abandon them completely. But some of these experiments might be reproducible, and that is what our examples are meant to suggest.[1] Thus, it seems to us that a dichotomy should be made, but one based on the criterion of reproducibility: experiments that are reproducible whether syntactic or semantic, yield facts that constitute the subject matter of linguistics, while the rest of the intuitions should be dealt with in philosophy.

REFERENCES

Bely, N. and Vasseux, Ph. (1973). *Le système de gestion de la grammaire du L.A.D.L.* Paris: L.A.D.L.

Boons, J. P., Guillet, A. and Leclère, C. (1975). *La structure des phrases simples en français* (Constructions sans complétives). 3 volumes. Paris: L.A.D.L.

Chomsky, N. (1967). 'Remarks on nominalization'. In R. Jacobs and P. Rosenbaum, eds., *Readings in English Transformational Grammar.* Waltham, Mass.: Blaisdell.

Chomsky, N. and Miller, G. A. (1963). 'Introduction to the formal analysis of natural languages.' In *Handbook of Mathematical Psychology*, **2**, 269–322. New York: Wiley.

Culioli, A. (1971). 'A propos d'opérations intervenant dans le traitement formel des langues naturelles', *Mathématiques et Sciences Humaines*, 34, 7–15.

Giry, J. (1972). 'Analyse syntaxique des constructions du verbe "faire" '. Thèse. Paris: L.A.D.L.

Gross, M. (1975). *Méthodes en syntaxe.* Paris: Hermann.

Harris, Z. S. (1951). *Methods in Structural Linguistics.* Chicago: Chicago University Press.

[1] Culioli (1971) has proposed an interesting set of concepts of meaning that appear to be reproducible.

—— (1964). 'The elementary transformations'. In *Papers in Structural and Transformational Linguistics* (1970), pp. 482–532. Dordrecht: Reidel.
—— (1968). *Mathematical Structures of Language*. New York: Wiley.
Joshi, A. K. (1972). 'Factorization of verbs: an analysis of verbs of seeing.' Mimeographed.
Kuroda, S. Y. (1973). 'Geach and Katz on presupposition'. Mimeographed.
Lees, R. B. (1960). *The Grammar of English Nominalizations*. La Haye: Mouton.
Lightner, T. M. (1972). *Problems in the Theory of Phonology*. Edmonton: Linguistics Research, Inc.
McCawley, J. (1968). 'Lexical insertion in a transformational grammar without deep structure'. In *Papers from the Fourth Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society.
Picabia, L. (1970). 'Etudes transformationelles de constructions adjectivales en français'. Thèse. Paris: L.A.D.L.
Postal, P. (1970). 'On the surface verb "remind" ', *Linguistic Inquiry*, 1, 37–120.
Ross, J. R. (1971). 'Act.' In R. Jacobs and P. Rosenbaum, eds., *Readings in English Transformational Grammar*. Waltham, Mass.: Blaisdell.