

THE ROLE AND FORM OF ANALYSIS IN MACHINE TRANSLATION  
THE AUTOMATIC ANALYSIS OF FRENCH AT SAARBRUCKEN

J. Weissenborn

Sonderforschungsbereich 100

(Special Research Field 100)

University of the Saar

Saarbrücken

Abstract

The paper deals with a number of problems connected with the analysis of natural languages as it affects machine translation. It poses the question as to which of the possible ways of defining the analysis result at transfer level appears the most suitable for the practical objective of creating a translation process for large quantities of text and a broad linguistic spectrum, and reviews some of the difficulties of analysis created by the open-endedness and ambiguity of natural languages. It will be proposed that the analysis of natural languages within the context of machine translation be seen not as something different from the analysis of natural languages for other purposes, such as fact retrieval and question-answering systems, but that it be kept open so that it can be adapted to other fields, primarily with a view to creating polyvalent language processing systems. Finally, the approach used for the automatic analysis of French in project C of the special research field 100 'Electronic Linguistic Research', at the University of the Saar will be described.

## 1. INTRODUCTION

Problems affecting the analysis of natural languages form the focal point of work on machine translations. Much of what I am going to say has already been said in the same or a similar way, since anybody working in this field has come up against the same problems. The main objective is to make clear what the fundamental problems are which arise when an analysis procedure is being developed and what criteria can be used to select certain proposed solutions. It must be borne in mind in this regard that the problem of analysis is not a problem peculiar to machine translation, but also affects other processes which handle natural language data. This factor should influence the strategy of analysis with regard to machine translation.

I also assume that the linguistic component is separate from the algorithmic component in the analysis process. The latter will not be considered in this paper. I would merely like to point out that there are today a series of efficient parsers, we could mention the Augmented Transition Networks (W.A. Woods, 1970), the Q-systems (A. Colmerauer, 1971) and the tree transducers (J. Chauché, 1974) which can simulate the most important types of grammar, leaving the linguist relatively free to choose his principles for the development of analysis grammars and dictionaries.

## 2. THE FORM OF THE ANALYSIS RESULTS

The analysis of the source language text represents the first stage of the machine translation system. The other two stages are the transfer, during which the lexical units of the source language are replaced by those of the target language, and the synthesis which generates the target language text.

The analysis stage in machine translation thus corresponds

to the act of understanding the text to be translated when human translators are involved. Obviously, this stage is crucial in both human and machine translation. Provided that the text has been correctly understood, the actual translation no longer poses an insoluble problem. From this it follows that the quality of a process of machine translation is directly dependent on the efficiency of its analysis (we could almost say comprehension component). What we understand by 'comprehension' must remain unanswered. In the case of a technical mode speech which does not offer any explanations we could say that what is meant is that a sentence is allocated a semantic representation or in the case of ambiguity several semantic representations which produce an unambiguous version of the information content, the 'meaning' of the text. This allocation of semantic representations to the expressions in the source language represent the central problem of automatic analysis.

Wide agreement has already been reached on the need to include a component in the grammar of natural language which explains this semantic representation, but there has been no agreement on how this is to be formulated from case to case and here the how applies to both form and content. Theoretical and applied linguistics have produced a variety of proposals on which I cannot comment in depth. In the field of theoretical linguistics we could mention: Chomsky's Deep Structures (cf Chomsky, 1965), the base component of Fillmore's Case Grammar (1968), the  $\epsilon$ - $\lambda$ - context-free languages of v. Stechov (1974), cf also SFB 99 (1976), categorical languages (cf Cresswell (1973)), representations of a predicate calculus basis (cf Bartsch and Vennemann (1972)), Pusch and Schwarze (1974), etc. In the field of artificial intelligence and machine translation we could mention Schank's 'conceptual' dependency structures (1975) and Wilks' structures (1973), which both apply lexical decomposition, or the 'language pivot' of the CETA in Grenoble (cf Boitet on this subject (1976)).

These modes of representation are closely related, as is shown, for instance, by predicate calculus translations of conceptual dependency structures (cf Schubert 1976).

Which of these various modes of representation should be selected as the form for the analysis results in a system of machine translation? Since none of the modes of representation given, apart from the 'language pivot' of the CETA, has been used with large quantities of text and a not too restricted linguistic spectrum, it is not possible to make any precise statement on their effectiveness and it is difficult to estimate the time involved in developing an analysis grammar catering for a broader linguistic spectrum, together with the corresponding dictionaries.

I therefore base my comments on the assumption that all the proposed forms of representation are possible candidates for analysis output. As far as the semantic representations for machine translation are concerned, a relatively restricted concept of meaning will suffice for the time being. Thus it will not be necessary to define for these representations concepts of formal semantics, such as the truth of a sentence, logical equivalence, and relations of consequence between sentences, which presuppose a disambiguated language. For machine translation it is thus immaterial whether from the sentence

(1) Hans ist ein passionierter Tennisspieler

it can be deduced that Hans is a tennis player, or, from the sentence

(2) Hans ist der vermutliche Mörder des Herrn Meier

it is not possible to deduce that Hans is a murderer. This is the result of the differing nature of the adnominal elements in both sentences.

Such phenomena must be taken into account where a question-answering system is concerned. Moreover, the translatability of natural language expressions into expressions of a disambiguated language is still an unsolved problem to a large extent. Therefore such formal representations cannot, for the time being, be incorporated into a translation process which is to be available in the not too distant future and which is to have the broadest possible linguistic spectrum.

Another desirable feature of a system of machine translation which is to be of practical use is the speed of operation, which is synonymous with economy of operation. Obviously, any semantic representation which has a structure differing greatly from the surface structure of the sentence analysed, whether by change of the word and/or clause sequence, or by change of the lexical elements as a result of lexical decomposition, for example, or by both together, requires more analysis stages, in other words an analysis grammar with more rules than are necessary for semantic representations, which remain relatively close to the surface structure. Thus, for this purpose, we can term 'deep' any representation which aims at a completely or partially language-free notation, e.g. Schank's conceptual dependency structures (1975) or the CETA 'language pivot' (cf. Vauquois (1975)). This applies also to representations in predicate calculus notation or Chomsky's deep structures, in particular when the latter are required to represent extensionally synonymous sentences, e.g.

(3) Dicke Männer lachen gern

(4) Männer, die dick sind, lachen gern

as a single deep structure, which automatically involves loss of the surface structure differences. It is not clear why a reduction of this sort should be carried out for the purposes of machine translation.

The drawbacks arising from the selection of a deep semantic representation reappear at the synthesis stage. If we assume that a multilingual translation system will primarily have to handle European languages, it is clearly desirable to take full advantage of the wide range of common structures existing between them. A semantic representation which retains most of the surface structure of the source language will, following substitution of target language lexemes for source lexemes, allow a sentence in the target language to be generated with the aid of fewer rules than a representation which has abandoned most of the source language structures.

I should like at this point to put a question in parenthesis and ask whether under certain conditions it would not be advisable to adopt, in machine translation, an analysis programme which aims at an extended semantic representation, in other words a representation which, for instance, could also be used for question-answering systems. To my mind this question is all the more pertinent when we consider that algorithms for data processing have meanwhile been developed which enable systems to be designed to cope simultaneously with various tasks connected with the analysis of natural languages (machine translation, question-answering systems). The possibilities of such polyvalent systems should at least be kept in mind when analysis grammars and dictionaries for machine translation are being developed even if the ideal that the semantic sentence representations supplied by the analysis for machine translation should be immediately transformable into such representations as become necessary for other purposes, will not become reality for quite some time.

The assumption underlying such a conception, however, is that a much more refined linguistic analysis, for instance in the domain of intensional semantics, of the languages to be processed will be available.

Summing up what has been said so far, it can be seen that

the form of semantic representation should be close to the surface and yet sufficiently detailed to permit selection of a correct translation equivalent. At the same time it should be easy to process algorithmically and possibly even be able to absorb any further information required without changing its formal structure.

The requirement that they should be easily processable is fulfilled by inter alia tree structures. These also permit the close-to-the-surface organization of the sentence into complex syntactical units to be reproduced without difficulty. With a view to fast processing it is also desirable that the number of nodes be kept as small as possible. Supposing that a nominal group having an adverbial function were to be described, such as "à tombeau ouvert" in the following sentence:

(5) Il roule à tombeau ouvert

then a mode of representation which did not involve an additional non-terminal symbol for the labelling of this function would in this case be preferable to one that did.

This can be achieved by labelling the nodes of the tree diagram with labels which can absorb in the form of variables all the morphosyntactical and semantic/logical data necessary for the transfer. These labels have the added advantage of being capable at any time of further expansion by the addition of new values without any need to change the tree structure, meaning that one more of the demands mentioned above is fulfilled. A parser for labelled tree structures such as these has been developed at the GETA at the University of Grenoble on the basis of work by Chauché(1974) and has been implemented there.

A semantic representation in the context of this formal structure would have to lend a grammatical description to the word forms of the text, show their grouping in larger syntact-

ical units and specify the nature of the relations between the elements of these complex units and the complex units to one another. More specific data on this point can be found in Leibniz (1975).

### 3. THE PROBLEM OF ALLOCATING SEMANTIC REPRESENTATIONS TO UTTERANCES

The central problem in analysis is thus that of allocating a semantic representation to each sentence of the input text. The difficulties which arise in so doing can be attributed basically to two characteristic properties of natural languages: their fundamental open-endedness and their ambiguity (cf. Klein 1977).

I understand by open-endedness the fact that the number of lexical units (lexemes) of any one language has never been definitely established, which means that it can be extended according to requirements. The implication of this for machine analysis is that even when comprehensive analysis dictionaries are available the case can always arise where some lexemes of a text to be analysed are not contained in the dictionary. This means that only a partial semantic representation can be allocated to the sentence concerned, or even none at all. Attempts to allocate at least a partial grammatical description to the unknown word form with the aid of an inflectional or derivation-based morphological analysis are promising within limits. However, no target language equivalent can be allocated to this lexeme, the result being that the translation of the sentence concerned must remain incomplete.

The situation as regards ambiguities is not quite so hopeless. Possible solutions are offered for some types and in part they are operationally feasible. What I understand by possible solutions in this context is "possible solutions from the point of view of the computer". Human pre-editing of the



texts to be analysed could of course process the non-reducible cases of ambiguity to such an extent that they no longer present a problem for a machine.

I draw a distinction in my following comments between semantic, syntactical and pragmatic ambiguities. The latter, which basically concern context-dependent reference of deictic expressions such as "je", "tu", "hier", etc. will not be discussed further here since they are not a major problem for machine translation.

The semantic and syntactical ambiguities are divided into those which only affect a single word form and those affecting the complex units, in other words the syntactical groups. The latter type of ambiguity can be subdivided into primary ambiguities and secondary ambiguities; the secondary ambiguities are those which arise in the course of automatic analysis when rules are applied in a particular sequence. I shall not go into any more detail on the latter. As for the rest the distinction between semantic and syntactical ambiguity at syntagmatic group level is artificial by virtue of the fact that they always arise together.

The first question one should ask oneself with regard to disambiguation concerns the extent to which it is necessary for correct translation of a sentence. This depends to some extent on the target language involved. When translating the following construction into English or German the ambiguity as to 'genitivus objectivus' or 'subjectivus' does not need to be resolved:

(6) La critique de Chomsky

The following is an example of ambiguity which does not affect machine translation:

(7) Trois hommes ont vu deux filles

This sentence can be interpreted in a number of ways:

(8) Trois hommes ont vu chacun séparément deux filles  
différentes

(9) Trois hommes ont vu chacun séparément les deux mêmes  
filles

(10) Trois hommes ont vu ensemble deux filles

The interpretation of 'is' as expressing logical equivalence or class-membership, could also be mentioned here, as in the following sentences:

(11) Pierre est le maire de cette ville

(12) Le chat est un mammifère

or the distinction between restrictive and non-restrictive relative clauses:

(13) Die Abgeordneten, die für dieses Gesetz stimmten,  
besiegelten den Untergang der Universität

Whether all (non-restrictive) or only some (restrictive) of the delegates voted for the law is not important for the purposes of machine translation (the example is from J.M. Zemb, 1972).

In other words we need only consider ambiguities which imply a choice between different lexemes and/or syntactical structures in the target language as well as those ambiguities which lead to erroneous results (dead ends) during analysis of the source language. An example of the first type would be:

(14) Mutter von zwei Kindern brutal ermordet

where for the purposes of translation into French a choice

between

(15) Une mère assassinée sauvagement par deux enfants

(16) Une mère de deux enfants assassinée sauvagement

would have to be made.

The second type concerns such cases as

(17) Il ferme la porte

In this sentence several lexemes are ambiguous and only one interpretation would supply a correct result for the purposes of analysis grammar: thus we need to identify 'ferme' as a verb, 'la' as an article and 'porte' as a noun (a series of typical cases of ambiguity in French is discussed by H.L. Scheel (1976)).

What roads are open to us then to disambiguate such cases for the purposes of machine analysis? The question as to the stage of analysis at which the disambiguation should take place is also quite significant. In such cases as (17) the disambiguation is carried out immediately following the morphological analysis at an early stage in the syntactical analysis process. In such cases as

(18) Il monte l'escalier

(19) Il monte la valise

(20) Il monte un commerce

(21) Il monte une machine

to which the following German sentences correspond:

(22) Er geht die Treppe hinauf

(23) Er trägt den Koffer hinauf

(24) Er baut ein Geschäft auf

(25) Er montiert eine Maschine

the disambiguation does not take place until the actual transfer occurs (an attempt to differentiate between the various verbs 'monter' at the analysis stage would be unrealistic in so far as this distinction is partly irrelevant if the target language is, for example, Italian).

Solution of these ambiguities is at present only possible to a limited extent. The situation in cases of the kind presented in (17) is relatively easy since the categorial ambiguity of the lexemes is solved by analysis rules which only permit certain sequences of categories. Generally it can be said that ambiguities which would lead to a syntactically incorrect result can be removed by appropriate formulation of the analysis grammar. This applies above all to syntactical ambiguities in word forms. Semantic ambiguities in word forms are a different matter, however. Whether in the sentence

(26) Elle laissa tomber sa glace

we are dealing with ice-cream or a mirror could only be decided if during the analysis process it is possible to consult a data base which is already available or is being developed during the analysis of the text. The latter is admittedly not technically impossible, but would place a tremendous burden on the analysis operation. The only practical and to some extent promising procedure which offers itself here is to produce analysis dictionaries which are text-type specific.

In the case of (18) - (21) the situation is slightly different. The correct allocation of German equivalents could possibly be achieved by means of semantic characterization of the nominal elements. The latter presupposes of course that a very complex indexing of the analysis and transfer dictionary entries is carried out, but that would not guarantee unambiguity for, depending on context, (18) and (19) could also mean

(27) Er montiert die Treppe

(28) Er montiert den Koffer

Similar difficulties occur where the ambiguity arises from the differing structural interpretation of an expression, as in (14). Without knowledge of the event it is impossible to decide whether 'von zwei Kindern' refers to the 'Mutter' or is the subject of 'ermorden'. In this case both analyses and consequently both translations would have to be given.

Summing up it can be said that as far as the problem of ambiguity and the unambiguous allocation of a given semantic representation to a given linguistic expression are concerned, large-scale disambiguation which goes further than the disambiguation of the categorial ambiguity of individual word forms will only be possible by incorporating a data base and a deduction component into the analysis process. Even without this type of expansion the demands made on analysis dictionaries and analysis grammar in order to avoid erroneous analysis results are still exacting enough. The following outline will give you some idea of the analysis process applied to French at Saarbrücken.

#### 4. THE AUTOMATIC ANALYSIS OF FRENCH AT SAARBRUCKEN

The algorithmic side of the process will be largely ignored. The importance of its structure for the assessment of an analysis system as a whole is obvious. At the moment the analysis of French is being carried out on the basis of systems developed by GETA at Grenoble and in Project A of SFB 100, which were described in the papers by C. Boitet and H.D.Maas. It can be assumed that in principle the linguistic analysis of the source language must provide identical data, irrespective of the system used. The ease with which linguistic descriptions can become an effective analysis system depends to a large extent on the form of the algorithmic component, however. The follow-

ing comments, where not general, refer to the version of the grammar as produced within the GETA system. The state of the art in September 1976 can be read in Weissenborn (1976).

The two components of the analysis process are the dictionary and the grammar. Obviously the development of these components will depend on a detailed and adequate description of the expressions in the language to be analyzed, where 'adequate' is taken to imply no more than the usefulness of the linguistic analysis for the purposes of a multilingual translation system.

This description produces a number of grammatical variables with their respective variable values which are to be used for the description of the dictionary entries and the formulation of the grammar rules.

Unfortunately the linguist who would like to develop the dictionaries and grammars for the automatic analysis of a natural language cannot now simply fall back on the results of existing language descriptions in order to determine the variables and allocate them to linguistic expressions. This is particularly evident where lexemes are attributed to lexeme classes.

The mixing of functional, morphological, semantic and syntactical criteria led to classifications which were more of a hindrance to automatic analysis than an aid. To describe for example 'naturellement' simply as an adverb or 'pareil' as an adjective makes the allocation of translation equivalents more difficult. We have not only

(29) Il ment naturellement ('... in a natural way')

but also

(30) Naturellement, il ment ('of course')

Not only

(31) Pareille chose arrive rarement ('Something like that')

but also

(32) Les deux instruments sont pareils ('... are similar')

A definition of the elements concerned which differentiates on the basis of their function and identifies the 'naturellement' in (30) as a modalisator (cf. Zemb (1976)) and 'pareil' in (31) as a deictor ('article word') makes correct translation easier.

A functional analysis of this type leads to at least six classes within the traditional adjective and adverb group (cf. Belin (1976)).

- (a) determinant of a predicative verb (être fort)
- (b) determinant of a non-predicative verb (manger vite)
- (c) determinant of a noun (station balnéaire)
- (d) determinant of an element of classes (a) - (f) (très grand)
- (e) determinant of an element of the deictor class (environ tous les hommes...)
- (f) determinant of an element of the modalisator class (peut-être pas).

It can be seen that this classification is the result of an analysis which uses the syntactical relations of one expression to other expressions as its classification criteria. The relationship in question is the one between determinans to determinatum (operator to operand). The basis for the lexeme classification of the expressions examined in the syntactical category of the expression they could determine. An expression can belong to a number of classes.

This procedure is applied to all lexical units; the name given to the relationship should not create the impression that it can always be interpreted in the same way, as is shown by the syntactical behaviour of the complex units formed from the 'determiner' and the 'determined'. This behaviour must also be taken into account, leading, for instance, to the establish-

ment of a class of deictors ('article words').

With regard to the remark that analysis in machine translation should be so arranged that it can be used, if required, to generate structures for other purposes without a great deal of modification, it should be pointed out that on the basis of the lexeme description which has been outlined it may be possible to formulate a disambiguated categorical syntax for a fragment of French.

The description of dictionary entries by variables must make available all data required for morphosyntactical analysis. It is not possible to state definitely what data these are until the analysis grammar has been completed. This means that the dictionary and the system of variables used for indexing cannot be developed separately from the grammar. In addition, there is the question as to which lexical units should be indexed, in other words selected for the dictionary.

If syntactical analysis is preceded by a morphological analysis, as in the case of French, some thought could be given to the possibility of relating not only forms of an inflectional paradigm but also forms of a derivational paradigm to a unique base form, such as RECEPTION to the base form of RECEVOIR, viz. RECEV-. In view of the existence of such extensive paradigms as that for UTILE: UTILISER, UTILISATION, UTILISATEUR, UTILITE, UTILITAIRE, UTILITARISME, etc. an appreciable reduction of the number of dictionary entries could be obtained. The reduction to a unique basic form is only possible without complications in the case of the elements of an inflectional paradigm, since here morphosyntactical changes are not accompanied, with changes in the syntactical properties of the lexeme. With the elements of a derivational paradigm this is, however, often and idiosyncratically the case. The problem is further complicated by the fact that the relationship of the basic forms to the derived forms with a given affix is not standard.



## Compare

- (33) Il m'a parlé naturellement
- (34) Il m'a répondu faiblement
- (35) Naturellement, il m'a parlé
- (36) Faiblement, il m'a répondu

The function of "naturellement" must be interpreted differently from the function of "faiblement". These phenomena will have to be taken into account when indexing the basic forms and affixes and would correspondingly complicate it, which would also apply to the formulation of the morphological analysis grammar. The resulting slowing down of the analysis process would not be compensated by the reduction of basic elements in the dictionary. For this reason it was decided not to implement a reduction of derivational paradigms to one basic form. Nevertheless, it is useful to select a series of productive suffixes for the dictionary. Where unknown word forms occur in a sentence ascertain amount of grammatical information can be attributed to these forms which permit analysis of a sentence to such an extent that at least a partial translation of the sentence concerned is possible.

The grammatical part of the analysis comprises a morphological and a syntactical component. The morphological component has been completed; the syntactical component will shortly be able to handle the most important structures. The basic principle of the syntactical component is its modular structure; in other words it is divided up into a series of elementary grammars, each of which contains as few rules as possible. Each of these elementary grammars handles a certain type of structure, such as simple and complex nominal groups or adjectival groups.

The modular structure greatly facilitates the development of the grammar since it is thus possible to locate errors in the rules more simply. In addition, it is possible to efficiently monitor the running of the entire process of analysis so

that transfer level can be reached as quickly as possible.

Nevertheless, further detailed linguistic investigations remain a prerequisite for the efficiency and for any improvement to the process of analysis.

#### BIBLIOGRAPHY

- Bartsch, R. u Vennemann, Th. (1972), *Semantic structures. A study in the relation between semantics and syntax*, Frankfurt: Athenaum.
- Belin, M. (1976), *Zu einer Grammatik der A-Lexeme des Französischen: Adjektiv- und Adverbialbereich*, in: *Arbeiten zur automatischen Analyse des Französischen I*, Universität Saarbrücken, SFB 100
- Boitet, Ch. (1976). *Un essai de réponse à quelques questions théoriques et pratiques liées à la traduction automatique. Définition d'un système prototype. Thèse d'Etat*, Grenoble.
- Chomsky, N. (1965), *Aspects of the theory of syntax*, Cambridge (Mass): MIT-Press.
- Colmerauer, A. (1973), *Les systèmes-Q*, in *TAUM 71*, Université Montréal, S. 1-44
- Cresswell, M.J. (1973), *Logics and languages*, London, Methuen
- Fillmore, Ch.J. (1968), *the case for case*, in: *Bach/Harms (eds.), Universals in linguistic theory*, N.Y., S.1-88.
- Klein, W. (1976), *Sprachanalyse und Organisation des Wissens*, in: *IBM-Nachrichten*, Februar 1977
- LEIBNIZ (1975), *Vorschlag zur Darstellung von Satzstrukturen auf der Transferebene*, Papier zum Treffen der Gruppe LEIBNIZ, Lugano März 1975, vervielfältigt
- Pusch, L. u. Schwarze, Ch. (1974), *Probleme einer Semantiksprache für den Sprachvergleich*, vervielfältigt, L.A.U.T., Trier
- Schank, R.C. (1975), *Conceptual information processing*, North Holland: Amsterdam
- Scheel, H.-L. (1976), *Zur Problematik von "Ambiguitäten" des Französischen in der maschinellen Analyse*, in: *Preprints des Kolloquiums zur "Automatischen Lexikographie, Analyse und*

- Übersetzung", Universität Saarbrücken, SFB 100, S. 60-66
- Schubert, L.K. (1976), Extending the expressive power of semantic networks, in: Artificial intelligence, S. 163-198
- SFB 99 (1976), Sonderforschungsbereich 99, Linguistik (Universität Konstanz), Teilprojekt AZ: Automatische Übersetzung (Universität Heidelberg), Forschungsbericht 1.11.1973 - 31.3.1976, Teil 1: Das Übersetzungssystem SALAT
- v. Stechow, A., (1974), - kontextfreie Sprachen. Ein Beitrag zu einer natürlichen formalen Semantik, in: Linguistische Berichte 34, S. 1-33
- Vauquois, B. (1975), La traduction automatique à Grenoble, Paris, Dunod
- Wilks, Y. (1973), An artificial intelligence approach to machine translation, in: Schank, R.C. u. Colby, K.M., (eds.), Computer models of thought and language, San Francisco: Freeman, S. 114-151
- Weissenborn, J. (1976), Elemente einer automatischen morphologischen und syntaktischen Analyse des Französischen, in: Arbeiten zur automatischen Analyse des Französischen I, Universität Saarbrücken, SFB 100
- Woods, W. A. (1970), Transition network grammars for natural language analysis, in: CACM 13, S. 591 - 606
- Zemb, J.-M. (1972), Le même et l'autre, Les deux sources de la traduction, in: Languages 28, S. 85 - 101
- Zemb, J.-M. (1976), L'analyse de la proposition et le calcul des prédicats, in: David, J. u. Martin, R. (eds.) Modèles logiques et niveaux d'analyse linguistique, Paris, Klincksieck, S. 165 - 174.