THE STUDY OF LEXICON IN GENERATIVE GRAMMAR. SOME ASPECTS OF WORDS FORMATION

LEOCADIO MARTIN MINGORANCE

Our main concern in this paper will be the exposition of some controversial issues in the study of the lexical component within a TG framework: the processes of formation of new lexical units.

The main processes of word formation (compounding, derivation, conversion, etc.) are an excellent test field for the exploration (and also for an attempt at formalization) of the creative capacity of the "ideal speaker-listener". This leads us directly to the question of whether the native speaker of a language has in his internalized grammar a specific system of rules which he applies systematically in the creation of new lexical units from already existing ones.

In TG at least two models have been developed with an acceptable degree of completeness. I am referring to R.B. Lees' *The Grammar of English Nominalizations*[1] and to H.E. Brekle's *Generative Satzsemantik im System der englischen Nominalkomposition*[2]. Chomsky's article "Remarks on Nominalization"[3] is a proposal rather than a model.

The title of a very recent article by D. Kastovsky[4]: "Word-formation, or: At the Cross-roads of Morphology, Syntax, Semantics and the Lexicon" could be a good summary of the levels of analysis of this field of grammar.

This problem is still far from having been given a satisfactory solution.

The transformationalist hypothesis, as was first developed by Lees in 1960, starts from a quite old assumption: that there exists a syntactic relationship between the members of complex lexical units; the parallel between sentence structures and word formation processes.

Consequently, these processes should follow the rules of the syntactic component.

This idea is not so new as it might appear at first sight. Even two outstanding Indo-european scholars, J.F. Staal and P. Kiparsky[5], have found this conception in Pāṇini. We have also found this idea in Grimm's *Deutsche Grammatik*, v. 2, in Brugmann and Del-

brück's *Grundriss* and in Paul's *Prinzipien*. It was developed in a more systematic way by Bally[6] with the introduction of the concepts of "thème" and "propos", taken from the Prague School. It has been Marchand[7] who has applied them systematically to the analysis of English word formation processes.

Lees (1960) has applied this system of an ordered set of transformations to English nominal compounds. In his system lexical insertion takes place after several transformations have been applied. With regard to the nature of these units, Lees points out:

> "The nominals generated by the rules
> which we shall study herein are not
> themselves sentences, but rather noun-
> like versions of sentences" Op. cit., p. 54.

He takes kernel sentences as the starting point, introducing at the same time the concept of matrix and constituent sentence for the analysis of complex structures. He classifies all English nominal compounds into nine types of sentence structures, with subgroups, according to the functions of the elements of the compound words.

In accordance with the model developed in Chomsky's *The Logical Structure of Linguistic Theory* and *Syntactic Structures*, he leaves aside the semantic component, because at that time no formal apparatus to describe it had been developed[8].

Within the transformationalist hypothesis the major breakthrough after Lees' model was H. Brekle's *Generative Satzsemantik...* His model, following the generative semantics approach, especially Fillmore's "case grammar", takes propositional concepts as primitive units of description. These propositional concepts ("Satzbegriffe") are propositions without quantification, modality, etc.

Their formal representation is by means of well-formed formulas of predicate calculus, following mainly Reichenbach and Carnap[9]. The central relational constants of this model are more or less similar to Fillmore's deep cases. These formulas can also be represented as tree structures.

In the description of more complex structures he follows Lees' analysis of matrix and constituent sentences, which in his model means the insertion of a propositional concept into another by means of special symbols.

With the use of two operators, which he calls Q and "lambda", the process of selection of topic and comment, which will become, respectively, the determinatum and the determinans of the compound or derived word, is started.

Not surprisingly, Brekle's book, an "Habilitationschrift", had Marchand as advisor.

Although he also postulates a deep syntactic level and a transformational component, he

has only developed the semantic component.

As we can see, Brekle's is a very powerful model, which allows us the description of the system of the language, i.e., in the case of lexical units, actual and possible formations. A modified version of Brekle's model has been applied by W. Kürschner[10] to the study of German nominal compounds in 1974. Lees himself, in an article published in 1970[11], proposed the introduction of a semantic component.

On the other hand, Chomsky's 1967 lexicalist hypothesis, was the natural reaction against the problems arisen with the application of Lees' transformational model[12].

Chomsky postulates a strict separation between the syntactic component and the lexicon. Lexical insertion should take place before transformations. For Chomsky, only syntactic processes follow regular patterns of formation. Therefore, they are the only ones susceptible of undergoing transformations. On the other hand, lexical processes are idio-syncratic and their productivity is restricted.

Chomsky, following Bloomfield's[13] idea that "the lexicon is an appendix of the grammar, a list of basic irregularities", considers the lexicon as "the full set of linguistic irregularities"[14].

His most important and best-known innovation is perhaps the introduction of the "$\bar{X}$-convention" to account for the common elements in the internal structure of, for instance, a verb and its derived nominal. X is a variable which could stand for the syntactic categories N,V or Adj. For instance, "propose" and "proposal" would constitute a single lexical entry. Under the node N it would take the physical form "proposal" and under the node V, "propose".

A substantial modification of this proposal has been made by R.S. Jackendoff[15] in 1975. Instead of the $\bar{X}$-convention, he proposes "distinct but related lexical entries". In the case of compounds, for instance, each one would have a completely specified lexical entry with a morphological redundancy rule and a semantic one. To indicate the relation between the compound and its constituent parts, he uses an arrow pointing in two directions as a symbol.

The paraphrase is specified, indicating the semantic relation between the parts of the compound.

This proposal is very interesting, because of the specification of the semantic component. Jackendoff's paraphrases are in most cases like Lees' sentences. He applies the same scheme to derivation and other word formation processes. This proposal would be worth being explored in more detail.

As regards derivational morphology, we will leave aside in this paper M. Halle's 1973

article[16], D. Siegel's 1974 M.I.T. Ph. D. dissertation[17] and M. Aronoff's 1976 mono-
graph[18].

As a general evaluation of both hypotheses, the lexicalist and the transformationalist,
we could say that the latter tries to analyze the system of the language, allowing trans-
formations an extraordinary power. The lexicalist hypothesis is more concerned with the
norm. With regard to information cost, the lexicalist hypothesis restricts the trans-
formational component at the expense of the lexical component. Perhaps the most sound
argument of the defenders of the lexicalist hypothesis is that transformations must be
meaning preserving and the whole process must be recoverable. According to this hypo-
thesis, this does not happen in a transformational study of lexicon.

Nevertheless we should make a fundamental distinction: between new formations or coinages
and lexicalized complex words. The latter should be entered directly in the lexicon, be-
cause once a word has come into existence, it is like a person: it develops its own
personality and its own ways. But when a speaker of a language creates a new lexical unit,
"ein durchsichtiges Wort" as H.M. Gauger[19] would call it, i.e. a complex word which can
be understood just knowing the meaning of its constituent parts, in the overwhelming
majority of cases he does it by applying specific rules, semantic, syntactic, morpho-
logical.

We can see these processes more clearly in an analysis of the relations between syntagmas
and compounds and the passing from one type to another, with the progressive acquisition
of the phonological, morphological and graphemic characteristics of a single word.

We have collected some evidence of this process in our Ph. D. dissertation about
compounding processes in English and Spanish[20]. We have devised a model with a semantic
component akin to Berkle's, with propositional concepts, and a deep syntactic component,
with simple sentences in their so-called "logical form": S-V-O-Adv. Each element of the
deep syntactic component is subcategorized to allow for selection restrictions. The verb
is subcategorized following Fillmore's "case frame" proposal.

We have classified a corpus of around 4000 nominal compounds and syntagmas collected from
different sources: novel, theatre plays, poetry, comics, magazines and newspapers,
advertisements and technical leaflets.

We have found that these "transparent" syntagmas and compounds have the same semantic
and deep syntactic structure and the same syntactic behaviour. They can be formed using
a limited set of semantic and syntactic structures. We find the differences between these
formations in both languages in their phonological and morphological structure.

As Kastovsky's title points out, word formation processes are semantic, syntactic, morpho-
logical and also phonological.

It is only with the study of the interaction of these four levels that we can have a complete account of the structure of complex lexical units.

NOTES

1. Lees, R.B.: *The Grammar of English Nominalizations*. The Hague: Mouton, 1960.

2. Brekle, H.E.: *Generative Satzsemantik im System der englischen Nominalkomposition*. Munich: Fink, 1976 (1[st.] ed. 1970).

Chomsky, N.: "Reamrks on Nominalization", in R. Jacobs & P. Rosenbaum: *Readings in English Transformational Grammar*, Waltham, Mass., 1970. (The article was written in 1967).

4. Kastovsky, D.: "Word-formation, or: At the Crossroads of Morphology, Syntax, Semantics and the Lexicon", *Folia Linguistica*, X, 1/2, 1977.

5. a. Staal, J.F.: "Room at the Top in Sanscrit. Ancient and Modern Descriptions of Nominal Compounds", *Indoiranian Journal*, 9, 1966.

    b. Kiparsky, P. & J. Staal: "Syntactic and Semantic Relations in Pāṇini", *Foundations of Language*, 5, 1969.

6. Bally, Ch.: *Linguistique générale et linguistique française*. Bern, 1965.

7. Marchand, H.: *The Categories and Types of Present-day English Word-formation*. Munich: Beck, 1969 (1st ed. 1960).

8. This model, without development of the transformational component, has been applied to the study of French nominal compounds by C. Rohrer (8.1.). A. Giurescu (8.2.) has applied a simplified version of the transformational component to the generation of Romanian, Italian, French and Spanish compounds.

L. Gleitman (8.3.) has tried to evaluate, in an experiment with different types of speakers, the regularities in the paraphrases of different types of three-word compounds. She has applied Lees' model in the syntactic classification of these compounds. J. Praninskas (8.4.) also follows Lees' analysis in her 1968 study of the processes and patterns of trade name creation.

   8.1. Rohrer, C.: *Die Wortzusammentsetzung im modernen Französisch*. Ph. D. diss. Erlangen: 1967. (Advisor: E. Coseriu).

   8.2. Giurescu. A.: *Les mots composés dans les langues romanes*. The Hague: Mouton, 1975.

8.3. Gleitman, L.: *An Experiment Concerning the Use and Perception of Compound Nominals by English Speakers*. Ph. D. diss. Pennsylvania Univ., 1967. (Advisor: H.Hiż).

8.4. Praninskas, J.: *Trade Name Creation: Processes and Patterns*. The Hague: Mouton, 1968.

9. a. Reichenbach, H.: *Elements of Symbolic Logic*. N. York, 1947. b. Carnap, R.: *Introduction to Symbolic Logic and Its Applications*. N. York, 1958.

10. Kürschener, W.: *Zur syntaktischen Beschreibung deutscher Nominalkomposita*. Tubingen: Niemeyer, 1974.

11. Lees, R.B.: "Problems in the Grammatical Analysis of English Noun Compounds", in M. Bierwisch & K. Heidolph: *Progress in Linguistics*. The Hague: Mouton, 1970.

12. Before Chomsky's "Remarks on Nominalization" the first comprehensive model describing the structure of the lexicon which postulated a non-transformational solution was Gruber: a. Studies in Lexical Relations. Ph. D. Diss.: M.I.T., 1965. (Advisor: N. Chomsky). b. *The Function of the Lexicon in Formal Descriptive Grammar*. Technical Memorandum. Systems Development Corp. Santa Mónica, Calif.

13. Bloomfield, L.: *Language*. London: Allen & Unwin, 1973, p. 274.

14. Chomsky, N.: op. cit.

15. Jackendoff, R.S.: "Morphological and Semantic Regularities in the Lexicon" *Language*, 51.1975.

16. Halle, M.: "Prolegomena to a Theory of Word-formation", *Linguistic Inquiry*, 1973,4.

17. Siegel, D.: *Topics in English Morphology*. Ph. D. Diss.; M.I.T., 1974. (Advisor: N. Chomsky).

18. Aronoff, M.: *Word Formation in Generative Grammar*. Cambridge, Mass.: The M.I.T. Press, 1976. (A slightly revised version of his M.I.T. Ph. D. diss.: *Word-Structure*, 1974.

19. Gauger, H.M.: *Durchsichtige Wörter. Zur Theorie der Wortbildung*. Heidelberg: C. Winter, 1971.

20. Martín Mingorance, L.: *Estudio contrastivo del proceso de composición en inglés y en español*. Ph. D. diss. Granada, 1977. (Advisor: Dr. R. Fente).