LINGUISTICS

SOME REMARKS ON HEADS

ROBERT D. BORSLEY

Adam Mickiewicz University, Poznań

The notion 'head' figured quite prominently in structuralist discussions of syntax. In classical transformational grammar, however, it was largely ignored, in spite of Postal's (1964:65) recognition that notions like head, modifier and endocentricity are "ones which we should certainly require that general linguistic theory characterize correctly". Since 1968, two developments have brought the notion back into syntactic discussion: on the one hand, the so-called X-bar theory of Chomsky (1970) and Jackendoff (1977), on the other, the dependency proposals of Robinson (1970) and Anderson (1971, 1977). Both approaches claim to provide a natural formalization of the head-modifier distinction. In this paper, I want to take a critical look at these claims. I will look first at the structuralist conception of the distinction. I will then argue that neither X-bar theory nor dependency theory represents an advance on structuralist views in this area. Then I will consider a related theoretical framework — categorial grammar. This is a more viable framework. It is questionable, however, whether it is preferable to a conventional phrase structure framework.

We can begin with the discussion of heads in Bloomfield (1933). Bloomfield's discussion is fairly representative of structuralist views, and, although quite brief, it is distinctly clearer than some recent discussions. For Bloomfield, the head of a construction is an immediate constituent with the same syntactic function as the whole construction.[1] It follows, of course, that only certain constructions have heads. These constructions are termed 'endocentric'. Other constructions, which lack heads, are termed 'exocentric'.[2] As an

[1] For Bloomfield, the syntactic function of an expression is defined in terms of its surface distribution. Modern grammarians will, of course, assume a more abstract notion.

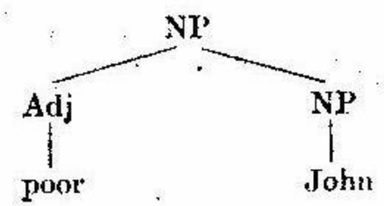
[2] This is a slight oversimplification. Phrases like *Bill and John*, where two constituents have the same syntactic function as the whole phrase, are regarded as head-less but endocentric.

example of an endocentric construction, Bloomfield gives the noun phrase *poor John*. *John* here has the same syntactic function as the whole phrase and is thus its head. As an example of an exocentric construction, he gives the sentence ('actor-action' construction) *John ran*. Bloomfield assumes that the head of a phrase may itself be a phrase. He suggests that the noun phrase *this fresh milk* has *fresh milk* as its head, and that this in turn has *milk* as its head. In such a phrase, the lexical item that is head of the most deeply embedded phrase is said to be the 'centre' of the whole phrase. Here, then, *milk* is the centre of *this fresh milk*.

A further strand in Bloomfield's discussion is the suggestion that constructions which lack heads generally have characteristic constituents which differentiate them from other constructions. He suggests, for example, that finite verbs are characteristic of sentences and that prepositions are characteristic of prepositional phrases ('relation-axis' constructions). We will see later that the notion of a characteristic constituent plays a central role in dependency proposals.

Bloomfield appears to assume, as others have since, that a noun is the head of a noun phrase. It is worth looking at this assumption in some detail. The obvious motivation for it comes from noun phrases like *poor John*. Here, we certainly seem to have a noun that is the head of a noun phrase. We cannot say, however, that a noun is always the head of a noun phrase. Many nouns do not have the same syntactic function as noun phrases. Consider, for example, the noun *man*. This combines with various determiners to form noun phrases like *the man*, *a man* and *that man*. These noun phrases cannot themselves combine with determiners. *the the man*, *a a man* and *that that man* are clearly nonsensical. It is clear, then, that an unqualified statement that a noun is the head of a noun phrase is inadequate. One response to this problem is to suggest that nouns fall into two classes: one class, exemplified by *John*, having the same function as NP's, the other, exemplified by *man*, having a different function. There is a simpler solution, however. Since *John* has the same function as a NP, we can say that it simply is a NP and assign *poor John* the following structure:

(1)



We can then reserve the label N for common nouns like *man*, which combine with determiners to form NP's. Given such an analysis a N is never the head of a NP. Only a NP can be the head of a NP.

One might also think that a verb is the head of a verb phrase. It seems, however, that this view is untenable too. Most verbs combine with various other constituents to form verb phrases. The resulting verb phrases cannot themselves combine with the same constituents. Consider, for example, the verb *kill*. This combines with a NP to form verb phrases like *killed Sam*. Such verb phrases cannot themselves combine with a NP. *killed Sam Mary* is clearly nonsense. Consider also *give*. This combines with a NP and a prepositional phrase containing *to* to form verb phrases like *gave a book to Jim*. Such verb phrases cannot combine with the same constituents. *gave a book to Jim a record to Jane* is obviously meaningless. It seems, then, that verbs are not the heads of verb phrases.

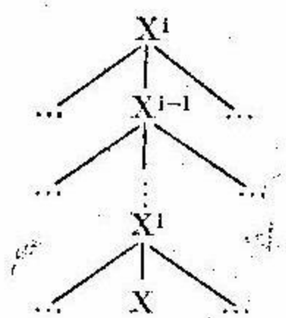
Having looked at a typical structuralist discussion and considered some of its implications, we can turn now to recent developments. I will look first at the X-bar theory. The theory is first and foremost an attempt to provide a specification of what is a possible phrase structure rule. It also claims, however, to provide a natural formalization of the head-modifier distinction. We will see that it does not provide an adequate formalization of the distinction as it is formulated in structuralist discussion. No alternative conception of the distinction is offered. This claim, then, must be rejected.

As it is developed in Jackendoff (1977), the X-bar theory claims that phrase structure rules are of three kinds, the main kind having the following form, where X is a complex of features and i some integer:[3]

(2) $X^i \rightarrow \ldots X^{i-1} \ldots$

Such rules generate phrase structure configurations of the following form:

(3)



Jackendoff suggests that either the X^{i-1} or the X at the bottom of the entire configuration can be defined as the head of X^i. The former reflects Bloomfield's usage of the term. The latter recalls Bloomfield's term centre.

---

[3] The other two rule types are rules for coordination and "category switching" rules like the following:

(i) $N'' \rightarrow ing - V''$

This is assumed to be involved in gerunds like *Noam's inventing a new theory*.

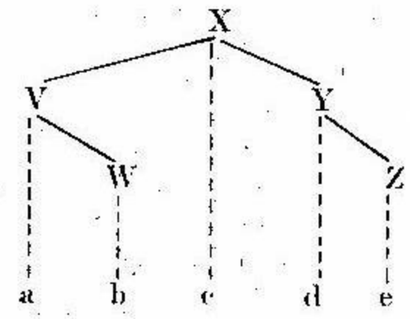
Jackendoff assumes that noun phrases, verb phrases, prepositional phrases and adjective phrases involve configurations like (3). It seems, then, that he is claiming that nouns, verbs, prepositions and adjectives are the heads of noun phrases, verb phrases, prepositional phrases and adjective phrases, respectively. We have already seen that the view that nouns and verbs are the heads of noun phrases and verb phrases is untenable. We can show that this is also true of the view that prepositions and adjectives are the heads of prepositional phrases and adjective phrases. Prepositions combine with NP's to form PP's. The resulting PP's cannot combine with NP's. *in*, for example, combines with a NP to form PP's like *in the house*. This cannot comb ne with a NP. *in the house the garden* is nonsense. Similarly, adjectives combine with various constituents to form AP's, but the AP's cannot combine with the same constituents. *afraid*, for example, combines with a PP containing *of* to form AP's like *afraid of spiders*. This cannot combine with a similar constituent. *afraid of spiders of snakes* is again nonsense. It seems, then, that neither prepositions nor adjectives are the heads of their phrases.

Thus, it seems that the X-bar theory identifies various constituents as heads which are not heads in the structuralist sense. Clearly, then, the theory does not provide an adequate formalization of the head-modifier distinction understood in structuralist discussion. It is possible that proponents of the theory have a different conception of heads in mind. The central claim of the X-bar theory is that constituents labelled with different feature complexes but the same superscripts show similar syntactic behaviour. Neither Jackendoff nor Chomsky claims that constituents labelled with the same feature complex but different superscripts show the same syntactic behaviour. It is possible, then, that they are assuming a different conception of heads. Neither, however, says what he understands by a head. We must conclude, then, that as far as the head-modifier distinction is concerned, the X-bar theory is no advance on structuralist discussion.

We can turn now to the dependency proposals of Robinson and Anderson. Whereas Chomsky and Jackendoff are concerned to specify what is a possible phrase structure rule, Robinson and Anderson argue that phrase structure representations should be replaced by dependency representations. Dependency theory has its origins in the work of Tesnière (e.g. Tesnière 1959). Hays (1964) first suggested that dependency representations might be preferable to phrase structure representations in a transformational grammar. Robinson and Anderson argue for dependency representations on a number of grounds. Particularly important in the present context is the claim that dependency representations provide a natural formalization of the head-modifier distinction. As we will see, the conception of heads that is assumed here differs from that of the structuralists. I will suggest that this conception is untenable.

Dependency representations are based on the transitive, asymmetric and irreflexive relation of dependency.[1] This is defined on terminal elements. Hence only terminal categories are employed. In a dependency representation, one element is independent and all other elements depend directly on just one element. The following is a typical dependency representation.

(4)



Capitals represent terminal categories and small letters terminal elements. c is independent. a and d depend directly on c, b and e depend directly on a and d, respectively, and indirectly on c. In such a representation, a phrase is an element plus all its direct and indirect dependents. In every phrase, then, there is one element on which the others directly or indirectly depend. This element is the head of the phrase and its immediate dependents are modifiers.

In a dependency representation, then, every constituent, whether endocentric or exocentric, has a head. Clearly, therefore, dependency representations do not formalize the structuralist notion of head. Robinson and Anderson are well aware of this. They propose a broader notion of head of which the structuralist notion is a special case. Echoing Bloomfield, they suggest that the head of a construction is a characteristic constituent, a constituent being characteristic of a construction just in case (a) it is an immediate constituent of every instance of the construction, and (b) it is not an immediate constituent of any other construction. In an endocentric construction, this constituent is the only obligatory constituent. Here, then, it is the head in structuralist terms as well as in the broader sense.

A dependency framework claims, then, that every construction contains a characteristic constituent. This is an interesting claim. Surprisingly, however, relatively little effort has been devoted to showing that it is tenable. I will suggest that it is not.

Particularly important here are noun phrases. An advocate of dependency representations must say that nouns are characteristic of noun phrases and hence that every noun phrase contains a noun as an immediate constituent. I suggested earlier that the label N should be reserved for common nouns. An advocate of dependency representations will have to use it more widely. Specifically,

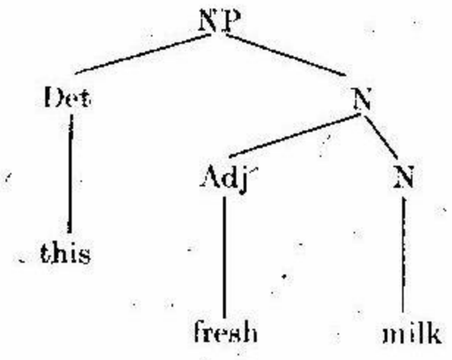
[1] This characterization is based on Robinson's. Anderson's characterization differs slightly. In the present context, however, the differences are of no importance.

both proper names and pronouns will have to be labelled N. If one assumes such a labelling, one is immediately faced with the problem of ruling out expressions like *the John* and *the he*. One way to do this is to mark common nouns as [+det] and proper names and pronouns as [−det]. Notice, however, that one is now saying that every NP contains either a [+det] N or a [−det] N. It is not at all obvious that this is any different from saying that every NP contains one of two distinct categories.

that clauses pose a further problem. It is generally agreed that at least some such clauses are NP's. On the face of it, then, these are NP's that do not contain a N as an immediate constituent. Anderson treats *that* clauses as dependents of an empty N. He offers no independent motivation for such an analysis, however. In the absence of such motivation, this can only be considered as an ad hoc move to preserve the claim that every NP contains a N as an immediate constituent. I think, then, that this claim is very dubious.

Equally dubious, I think, is the associated claim that N's are only immediate constituents of NP's. Recall here Bloomfield's phrase *this fresh milk*. As Bloomfield notes, *milk* and *fresh milk* have the same syntactic function. The former is a N. It seems, then, that the latter should be too. The whole phrase, then, might have the following structure:

(5)



Here, we have a N which is an immediate constituent of another N. I think, then, that the claim that N's are only immediate constituents of NP's is untenable.

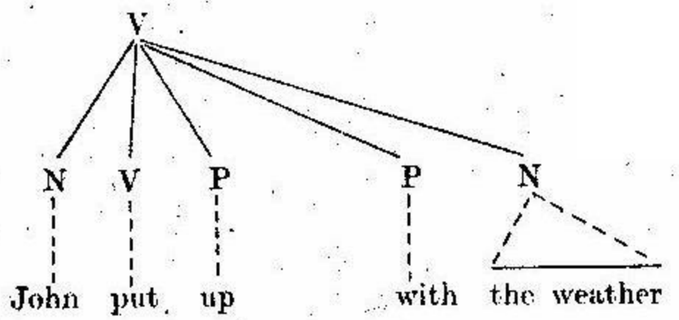
A number of considerations suggest, then, that the claim that N's are characteristic of NP's is false, and hence that not every construction contains a characteristic constituent. Thus, it looks as if the conception of heads that is assumed in a dependency framework is untenable.

Even if every constituent did contain a characteristic constituent, dependency representations would still face problems. An important fact about dependency representations is that the head of a phrase cannot itself be a phrase. This poses problems. Within a dependency framework, it is natural to assume that the head of a sentence is the main verb. It is generally assumed

that expressions like *put up with* and *take advantage of* are complex verbs. Within a dependency framework, however, it is impossible to treat them as such. A sentence like (6) would have to be assigned a structure like (7):

(6) John put up with the weather.

(7)



Presumably, no-one would regard this as an adequate structure. Thus, even if the basic assumption of dependency theory were correct, it would still face problems.

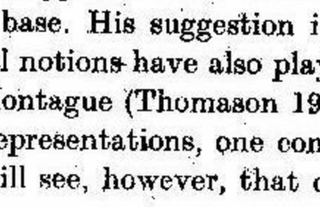
There is one further element in Anderson's discussion of dependency representations that requires consideration. This is the suggestion that dependency representations provide a natural formalization of Bartsch and Vennemann's operator-operand relation (Bartsch and Vennemann 1972; Vennemann 1973). Bartsch and Vennemann argue that various constructions consist of an operator and an operand and suggest that natural languages tend to serialize all operator-operand pairs in the same direction. Anderson suggests (1975, 1977) that the operator-operand relation can be equated with the modifier-head relation of dependency representations. It seems likely that a grammar must recognize something like Bartsch and Vennemann's operator-operand relation. It is not at all obvious, however, that this provides any support for dependency representations. The crucial point is that a construction can only contain a single operator since no expression can simultaneously be the operand of two operators. In a dependency representation, however, a construction can contain any number of modifiers. It seems, then, that the operator-operand relation cannot be equated with the modifier-head relation and hence that dependency representations cannot draw any support from the observations of Bartsch and Vennemann.

I think, then, that Robinson and Anderson fail to show that dependency representations are a viable alternative to phrase structure representations. In the present context, the crucial point is that the conception of heads that is assumed in a dependency framework appears to be untenable. The notion of a characteristic constituent is an important one, but the claim that every construction contains a characteristic constituent seems to be false. It seems, then, that, where the head-modifier distinction is concerned, dependency theory, like X-bar theory, is no advance on structuralist conceptions.

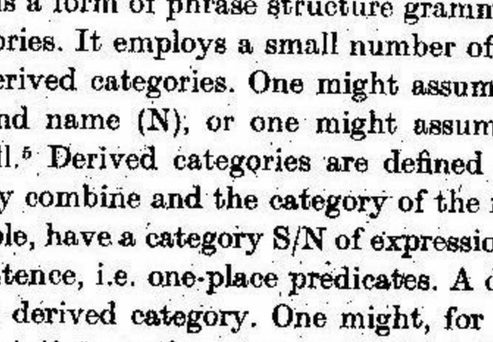
I want now to take a look at categorial grammar. Categorial grammar has its origins in the work of the Polish logicians Leśniewski and Ajdukiewicz. Lyons (1966, 1968) first suggested that transformational grammars should incorporate a categorial base. His suggestion is developed in Lewis (1972). In recent years, categorial notions have also played a central role in the work of Cresswell (1973) and Montague (Thomason 1974). In categorial representations, as in dependency representations, one constituent has a special role in every construction. We will see, however, that categorial representations are more viable.

A categorial grammar is a form of phrase structure grammar distinguished by the nature of its categories. It employs a small number of basic categories and a larger number of derived categories. One might assume just two basic categories, sentence (S) and name (N), or one might assume a category of common nouns (C) as well.[5] Derived categories are defined in terms of the categories with which they combine and the category of the resulting expression. One might, for example, have a category S/N of expressions that combine with a name to form a sentence, i.e. one-place predicates. A derived category may itself combine with a derived category. One might, for example, assign definite and indefinite descriptions to the category S/(S/N) of expressions that combine with a one-place predicate to form a sentence. Assuming these categories and a category (S/N)/N of expressions that take a name to form a one place predicate, i.e. transitive verbs, and a category (S/(S/N))/C of expressions that take a common noun to form a definite or indefinite description, i.e. determiners, we might have representations like the following:

(8)

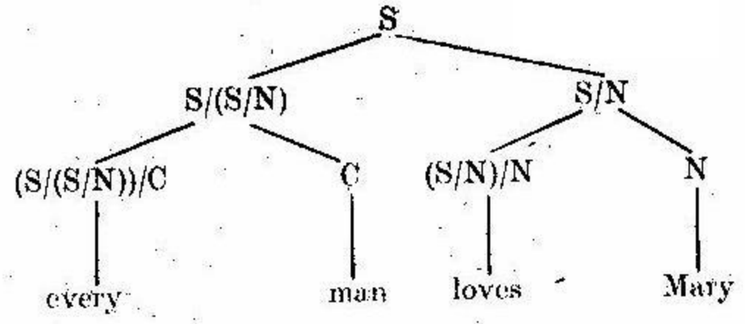


(9)



[5] For some discussion of alternative possibilities, see Lewis (1972).

(10)



As it stands, this system of categories does not allow for definite and indefinite descriptions in non-subject position. For some discussion of how they might be accomodated, see Lewis (1972).

A central fact about categorial representations is that every phrase contains a constituent whose category label contains on the left of the main slash the category label of the whole phrase and on the right the category label(s) of the other constituent(s). For reasons which will be indicated shortly, we can call this constituent the functor of the construction. In (8) and (9), then, S/N is the functor of S, while in (10) S/(S/N) is. (S/N)/N is the functor of S/N in (9) and (10). Finally, in (10), (S/(S/N))/C is the functor of S/(S/N).

It should be stressed that claiming that every construction contains a functor is quite different from claiming that every construction contains a characteristic constituent. The fact that a certain category is the functor of a construction does not imply that it is an immediate constituent of every instance of the construction. S/(S/N) is the functor of S in (10), but does not appear in (8) or (9). Nor does it imply that it never appears as an immediate constituent of any other construction, only that it cannot be the functor of any other construction. It follows, then, that categorial grammar faces none of the problems associated with the claim that every construction contains a characteristic constituent.

There are two further points to be made about functors. Firstly, notice that the functor of a construction is never its head in the structuralist sense. For a functor to be the head of its construction it would have to be an expression which combines with certain constituents to form the same expression. It is fairly clear that such an expression could have no finite specification. Hence, there can be no such expression. Secondly, notice that the functor of a construction has a distinctive character if the construction is endocentric. Specifically, the category label on the left of the main slash appears on the right as well. To illustrate this point we can return again to Bloomfield's example *this fresh milk*. As we have observed, *milk* and *fresh milk* have the same function. *fresh milk*, then, is an endocentric construction. In the present framework, both *milk* and *fresh milk* will be labelled C. The adjective, then, will be labelled C/C.

Lyons (1968:235) suggests that the fact that it is immediately apparent from the functor of a construction whether it is endocentric or exocentric is a reason for preferring a categorial base to a conventional phrase structure base. It is not at all obvious, however, that this is so. As long as expressions with the same syntactic function have the same label, it will be perfectly clear in a conventional phrase structure representation whether a construction is endocentric or exocentric. A construction will be endocentric just in case the first node that exhaustively dominates it directly dominates a node with the same label.[6] In (1), the first node that exhaustively dominates *poor John* is the upper NP. This directly dominates another NP. Thus, the construction is correctly identified as endocentric. Similarly, in (5), the first node that exhaustively dominates *fresh milk* is the upper N. This directly dominates another N. *fresh milk*, then, is identified as endocentric. I don't think, then, that the endocentric-exocentric distinction provides any motivation for preferring a categorial base to a conventional phrase structure base.

The principal reason for recent interest in categorial representations is that they provide a natural basis for explicit semantic interpretation. We have seen that every constituent contains a distinguished constituent which we have called the functor. The reason for adopting this term is that we can interpret this constituent as a function operating on the intensions (meanings) of its sister constituents to give the intension of the whole construction. First, of course, we need intensions for the basic categories. For sentences we can suggest functions from contexts and possible worlds to truth values,[7] for names functions from contexts to individuals, and for common nouns functions from contexts to sets. We can then interpret a S/N as a function from N intensions to S intensions, a S/(S/N) as a function from S/N intensions to S intensions, a (S/N)/N as a function from N intensions to S/N intensions, and a (S/(S/N))/C as a function from C intensions to S/(S/N) intensions. More generally, we can say that the intension of a $c/c_1 \ldots c_n$, where c, $c_1, \ldots$, and c_n are any categories, basic or derived, is an n-place function from c_1 intensions, ..., and c_n intensions to c intensions. It is fairly clear, then, that categorial representations fit naturally into an explicit semantics.

It looks, then, as if semantic considerations favour categorial representations over conventional phrase structure representations. This is not necessarily the case, however. It is fairly clear that any explicit semantics must analyze sentences as sets of function-argument relations. It is not so clear, however, that these relations must be explicitly represented by configurations of category labels. In a grammar employing conventional phrase structure representations, it would be possible to specify for each construction which constituent is the functor. There are moreover certain disadvantages to the view that function-argument relations should be represented by configurations of category labels. We have seen that transitive and intransitive verbs must be assigned to different categories. Clearly, ditransitive verbs like *give* must be assigned to a further category. All verbs, however, have the same morphology. If they are assigned to a variety of different categories, this fact will be an accident. It is not at all obvious, then, that categorial representations are preferable to conventional phrase structure representations.

[6] It is worth noting that the X-bar theory, as developed by Jackendoff, explicitly rules out such configurations, since it allows no phrase structure rules in which the category on the left also appears on the right. Quite why Jackendoff wishes to prohibit such rules is unclear.

[7] Alternatively, one might take the intension of a sentence to be a function from contexts to propositions, the latter being functions from possible worlds to truth values.

REFERENCES

Anderson, J. M. 1971. *The grammar of case: Towards a localistic theory*. London: Cambridge University Press.

Anderson, J. M. 1975. "Noun phrases as adjectives: Serialisation in seven parts". Linguistic Agency, University of Trier.

Anderson, J. M. 1977. *On case grammar*. London: Croom Helm.

Bartsch, R. and T. Vennemann. 1972. *Semantic structures*. Frankfurt am Main: Athenäum.

Bloomfield, L. 1933. *Language*. New York: Holt, Rinehart and Winston.

Chomsky, N. A. 1970. "Remarks on nominalization". In Jacobs, R. A. and P. S. Rosenbaum. (eds). 1970. 184—221.

Cresswell, M. J. 1973. *Logics and languages*. London: Methuen.

Culicover, P., Wasow, T. and A. Akmajian. (eds). 1977. *Formal syntax*. New York: Academic Press.

Davidson, D. and G. Harman. (eds). 1972. *Semantics of natural language*. Dordrecht: Reidel.

Hays, D. G. 1964. "Dependency theory: A formalism and some observations". *Language 40*. 511—25.

Jackendoff, R. S. 1977. "Constraints on phrase structure rules". In Culicover, P., Wasow, T. and A. Akmajian. (eds). 1977. 249—83.

Jacobs, R. A. and P. S. Rosenbaum. (eds). 1970. *Readings in English transformational grammar*. Waltham, Mass.: Ginn.

Kimball, J. P. (ed.). 1973. *Syntax and semantics*. vol. 2. New York: Academic Press.

Lewis, D. 1972. "General semantics". In Davidson, D. and G. Harman. (eds). 1972. 169—218.

Lyons, J. 1966. "Towards a 'notional' theory of the 'parts of speech'". *Journal of linguistics 2*. 209—36.

Lyons, J. 1968. *Introduction to theoretical linguistics*. London: Cambridge University Press.

Postal, P. M. 1964. *Constituent structure: A study of contemporary models of syntactic description*. Bloomington: Indiana University Press.

Robinson, J. J. 1970. "Dependency structures and transformational rules". *Language 46*. 259—85.

Tesnière, L. 1959. *Elements de syntaxe structurale*. Paris: Librairie C. Klincksieck.

Thomason, R. H. (ed.). 1974. *Formal philosophy: Selected papers of Richard Montague*. New Haven, Conn.: Yale University Press.

Wennemann, T. 1973. "Explanation in syntax". In Kimball, J. P. (ed.). 1973. 1—50.