LINGUISTICS

NOTES ON FUNCTIONAL PROPERTIES OF DEEP STRUCTURE CATEGORIES

KAZIMIERZ POLAŃSKI

Adam Mickiewicz University, Poznań

The notion 'deep structure' has been introduced into linguistics in order to better account for the speaker's intuition regarding utterances. What is directly open to linguistic analysis is surface structure. Deep structure is a set of hypotheses about our understanding of surface structure phenomena.

Although many papers have been published concerning the theory of deep structure, its very character still seems to be insufficiently explored. If deep structure is to account for the speaker's intuition with regard to his understanding of utterances, it has to be closely related to surface structure. Deep structure is an abstraction. In this respect every type of grammar in the past had to take it into account to some extent, even though linguists were not always aware of it. If a linguist said that such forms as *sing, sings, singing, sang, sung* in English represent one word it was an abstraction based on the fact that speakers assign these forms to one word. None of the forms cited separately could be considered as the word, as word is an abstraction.

Examples like this are common in traditional or structural linguistics. They cannot, however, be recognized as a conscious theory of deep structure. Chomsky has tried to discover elements of such a theory in the work of scholars of previous centuries (cf. Chomsky 1964, 1965: 198–200), but it is the linguistics of recent years that should be credited with its explicit formulation.

The deep structure of particular languages must be based on regularities which can be observed in utterances. Forms like *sing, sings, singing, sang, sung* can be considered as representing one word only because the relationship between them resembles the relationship between such forms as *ring, rings, ringing, rang, rung; drink, drinks, drinking, drank, drunk*, and further *like, likes, liking, liked, liked* etc. On the basis of such proportions forms of the type *go, goes, going, went, gone* can be accounted for because their functions and distribution resemble the functions and distribution of the other sets.

In my opinion, basing conclusions about deep structure on isolated examples is not justified. I mean here statements such as the one by J. F. Staal (1967: 66–87) to the effect that the deep structures of the sentences *He follows her* and *She precedes him* are identical. Such investigations may certainly have value for the logical analysis of sentences, but they cannot be considered as purely linguistic. Linguistic analysis should be based on language data. Paraphrases do not always justify conclusions about deep structure. Only constructions displaying regular relationships can be taken into account in looking for deep structure types. I do not, however, want to deny the value of logical analysis of language, whose importance for some linguistic areas has been proved. Structures determined by logical analysis are especially important for purposes of machine translation, as they have a more abstract character and are therefore more universal than structures singled out by purely linguistic analysis. But only a deep structure which is closely related to surface structure data will contribute to the understanding of a particular language system.

There is no one-to-one correspondence between given deep structures and surface structures from the point of view either of the former or of the latter. The same deep structure can be represented by several synonymous surface structures and, vice versa, several different deep structures may have their counterparts in an identical surface structure. Ambiguities of surface structures result from the fact that the meaning of the deep structure is preserved while it is being transformed into a surface structure. Deep structure simply reflects the intuition of the speaker, for whom there is no ambiguity or homonymity (cf. Jakobson 1963: 64). The speaker always refers a surface structure utterance to an appropriate unambiguous deep structure pattern, just as he always connects every phonetic word of his language with one unambiguous lexical unit. There is no general, ambiguous *rait* for English speakers but either *rait* = *write*, *rait* = *rite*, or *rait* = *right* etc. It is precisely for this reason that we say deep structure accounts for the speaker's intuition. What we mean by this is that the deep structure hypothesis implies that speakers relate various surface structure constructions to appropriate deep structure types.

I should like to emphasize here, however, that deep structure cannot be expected to reflect the speaker's thoughts. What linguists can examine is the speaker's utterances. A linguist is not entitled to ascribe meanings to sentences which are not expressed in some formal way, although logical or ontological analysis might allow such meanings. Grammar should be concerned with language synonymity but ought not to concern itself with equivalences which are not conditioned linguistically, e.g. equivalences of the type:

Irene came = Edward's daughter came. (1)

Such equivalences lie beyond the scope of the linguist's competence. In order to state that *Irene* = *Edward's daughter*, it is necessary to refer to extralinguistic experience. That is why we cannot require a grammar to account formally for all synonymous utterances often cited as paraphrases. A linguist can be expected to explain relationships between constructions which display regular interdependences (e.g. active — passive, subordinate clause — nominal construction, participial construction — infinitival construction). Isolated facts should be shifted to the lexicon (the lexicon may also be examined from the viewpoint of its deep structure, but this is another question).

As is well known, two types of relations, paradigmatic and syntagmatic, can be distinguished in language. Too little attention has as yet been given in transformational-generative grammar to the latter. The meaning of a sentence is made up both of the meanings of the morphemes contained in it and of the meanings expressed by its structure. This fact has long been recognized by structuralists. Take as an example the English sentence:

John kicked Harry. (2)

It informs us not only that reference is being made to 'a John', 'a Harry' and an action of 'kicking', but also that it is John who kicked Harry and not the other way around. The latter piece of information, which results from the structure of the sentence, concerns a relational meaning. Countless extralinguistic relations can be observed, and they are expressed in language to different degrees. A linguist should not omit this aspect of language.

In transformational-generative grammar these problems have been neglected until quite recently. Categories, i.e. paradigmatic relations, have been emphasized instead. This tendency was inherited from American descriptive linguistics, which concerned itself very little with problems of syntagmatic relations, at least at the sentence level. We look in vain for such concepts as 'subject', 'object', etc. in the works of classic descriptivists; nor were they mentioned in Chomsky's early work. But the notions of parts of speech (or lexical classes) and word-order do not fully specify syntactic functions. The overemphasis on sequence problems in transformational-generative grammar resulted from the fact that it was primarily developed on the basis of English material where the role of word-order is significant (though only in surface structure). It is only recently that generativists' attention has turned to concepts expressing relational properties in deep structure.

Chomsky's solution proposed in Aspects (1965: 69) to reduce syntagmatic relations to statements of the type 'NP's relation to S' or 'NP's relation to VP' cannot be considered satisfactory, as he does not take their semantics into account. Shortcomings of this kind in transformational-generative grammar can be reduced to the underestimation of sentence members or so-called syntagmatic positions (cf. Karolak 1963). The structure of a sentence can be specified only by labelling its syntagmatic positions and indicating which lexical classes can take these positions. Each such position can be occupied by several different categories. It is only the position that interrelates such

categories. Sentence structures may differ as to the syntagmatic positions they provide, so that various sentence types can be distinguished in particular languages (cf. Polański 1966: 83–99).

In papers by transformationalists categories like 'Time' and 'Manner' have started to appear. It is easy to see that such categories written in the same line with the categories 'NP', 'VP', etc. display striking inconsistency, since 'NP', for instance, can also be 'Time' or 'Manner'. In fact, these are function names, and categories and functions are thus mixed in one line.

Traditional grammarians and European structural linguists have made use here of the concepts 'subject', 'object', etc. Such notions are good only as qualifications for surface structure phenomena; they do not account for deep structure facts[1]. Traditional grammarians tried to cope with this deficiency by introducing the distinction between 'grammatical subject' and 'logical subject' etc. where logical subject corresponds to some deep structure category and grammatical subject is simply the surface structure subject. Even the notion 'psychological subject' has been introduced into sentence analysis. Thus surface structure phenomena have been mixed with deep structure facts on the one hand, and sentence structure problems with text structure problems on the other. Note that these misleading notions have been used by some representatives of transformational-generative grammar (cf. Chomsky 1965: 23, 70, 163, 221; Lyons 1968: 343–344; Lee 1968: 37–113.)

If a passive sentence is to be considered a transformation of an active one and derived from an identical deep structure, it is impossible to talk about 'subject', 'object', etc. in the deep structure, since 'subject' refers to one thing in active sentences and another in passive ones; and, after all, it is by means of the concept of deep structure that we want to give a semantic interpretation of surface structure data.

In European structural linguistics the notion 'subject' has been relatively clearly defined (cf. Trubetzkoy 1939, Gołąb 1958), with reference, of course, to surface structure. A question arises as to which notions should be introduced to denote relational properties or functions in the deep structure. Consider the following Polish sentences:

Janek pisze ten list bardzo długo 'John has been writing this letter for a very long time' (3)

List ten jest pisany przez Janka bardzo długo 'This letter has been being written by John for a very long time' (4)

Pisanie tego listu przez Janka trwa bardzo długo 'John's writing of this letter has lasted for a very long time' (5)

[1] I emphasized this point in my paper presented at the 27-th Congress of the Polish Linguistic Society (April 1968; cf. Polański 1970, cf. also Fillmore 1968).

Although the sentences cited vary in form, the semantic relationships between *Janek — pisać — list* remain the same. These are relational meanings. This resembles the behaviour of words which can take various inflectional forms in various contexts but do not change their lexical values, compare in Polish: *ręk-a, ręk-i, ręc-e* etc. The only difference between these two types of semantic identity is that the first is relational and the second is lexical.

Relational meanings expressed by sentences should be carefully examined and appropriate semantic elements distinguished as deep structure counterparts of such surface structure notions as 'subject', 'object', 'complement', 'adverbials'. This investigation should be carried out in every language with careful reference to the expression system of that language. Then such concepts come into play as the following: action, state, trait, action-author, action-bearer, action-involved, trait-bearer, beneficiary, attendant, cause, condition, concession, time, place, circumstance, manner[2]. Languages differ with respect to distinguishing particular functions structurally. Some languages distinguish functions which are irrelevant in others, e.g. feeling as separate from action and state, or rather as a subtype of action. For example, in Lesghian the function 'feeling' is expressed in the following way. In sentences containing verbs which denote feeling, action-authors are expressed by the dative, while in other sentences with transitive verbs the agentive case shows up in this position (cf. Gołąb 1958: 37–38). The Avar language even distinguishes two types of feeling — sensorial and emotional. In the former case the action-author (or rather feeling-author) shows up in the dative, in the latter case in the superessive (cf. Gołąb 1958: 37–38).

In Polish, on the other hand, such distinctions are not made. The sentences

Janek bije Staszka 'John beats Staszek' (6)

Janek widzi Staszka 'John sees Staszek' (7)

display the same structural characteristics, for example both can be passivized:

Staszek jest bity przez Janka 'Staszek is being beaten by John' (8)

Staszek jest widziany przez Janka 'Staszek is being seen by John' (9)

Which relational meanings are expressed by sentences depends mainly on predication. I share the opinion that the center and at the same time the constitutive member of a sentence is predication (cf. Kuryłowicz 1948). Predication, however, is a wider notion than verb or copula. Nominal sentences have predication even if they do not contain a copula in the surface structure, compare in Russian:

On dobryj čelovek 'He is a good man' (10)

On učitel' 'He is a teacher' (11)

[2] Cf. above footnote 1.

Predication may be identified in generative rules with verbal modality (cf. Polański 1969: 91–100). In the surface structure zero may correspond to predication. But predication is always attached to some category, to some morpheme class:

predication + verbal morphemes → verbal forms (12)
predication + adjectival morphemes → a copula||zero + adjectival predicative (13)
predication + NP → a copula||zero + NP predicative (14)
predication + Adv → a copula||zero + adverbial predicative (15)

What determines the restrictions regarding relational meanings, then, is not predication itself but its lexical component, i.e. verb, adjective, adverb, noun. In Polish, verbs of the type *bić* 'beat', *nieść* 'carry', *widzieć* 'see', *słyszeć* 'hear' admit two NP's — one in the function of action-author (in the nominative), the other in the function of action-bearer (in the accusative or — after negation — genitive), whereas intransitive verbs, such as *iść* 'go', *spać* 'sleep', *leżeć* 'lie' etc., do not allow the distinction between action-author and action-bearer. The function of NP connected with them is neutral from this point of view and can therefore be called action-involved.

But action-bearer can also be expressed by the nominative in Polish. This is so not only in passive sentences but also in sentences containing some verbs with *się*, compare for example:

Zabawka się zepsuła 'The toy got spoilt' (16)
Filiżanka się rozbiła 'The cup broke' (17)
Okno się otworzyło 'The window opened' (18)
Okno się zamknęło 'The window closed' (19)

In my opinion, one can talk about action-bearer only when there is an opposition between this and action-author. It is true that action-authors are not specified in the above sentences, but they are implicit. The structural means of expressing them is the pronominal form *się*. When *się* is removed from the sentence a position opens for an explicit action-author, e.g. in questions:

Kto zepsuł zabawkę? 'Who spoilt the toy?' (20)
Kto rozbił filiżankę? 'Who broke the cup?' (21)
Kto otworzył okno? 'Who opened the window?' (22)
Kto zamknął okno? 'Who closed the window?' (23)

It is noteworthy that English makes use of another grammatical device in such cases. There is no difference between the forms of the verb, but the structures of the sentences vary. Compare the English equivalents of sentences 17–19 and 21–23 (cf. Fillmore 1968).

Besides action-author and action-bearer there should also be a position instrument':

Janek uderzył Staszka kamieniem 'John hit Staszek with a/the stone' (24)

There are many utterances like the above, in which all three positions are explicitly expressed, but we also encounter sentences which are incompletely specified from this viewpoint:

Staszka uderzono 'Staszek was hit' (25)
Staszka uderzono kamieniem 'Staszek was hit with a stone' (26)
Kamień uderzył Staszka 'A stone hit Staszek' (27)
Janek uderzył Staszka 'John hit Staszek' (28)

Some authors are apt to interpret the deep structure function of all inanimate surface structure subjects as instrument (cf. Fillmore 1968, Lee 1968). The problem, however, does not seem that simple. Consider the following sentences in Polish:

Cegła uderzyła Janka 'A brick hit John' (29)
Skała zmiażdżyła sarnę 'A rock crushed a deer' (30)
Wiatr złamał drzewo 'Wind broke the tree' (31)
Staszek uderzył Janka 'Staszek hit John' (29a)
Staszek zmiażdżył sarnę 'Staszek crushed a deer' (30a)
Staszek złamał drzewo 'Staszek broke the tree' (31a)

There are two main differences between sentences 29, 30, 31 and 29a, 30a, 31a. First, the classes of nouns occupying the position of the surface structure subject are different (inanimate in the former examples and animate in the latter). Second, the action itself can be treated as either intentional or non-intentional. These two factors need not co-occur, since animate action-authors can also cause non-intentional actions on the one hand, and inanimate action-authors (e.g. wind, water, sea, river, computer) can cause actions considered by speakers as intentional. We cannot, however, be concerned with the logical or ontological point of view here. What we have to do is try to find out those meanings which are expressed linguistically. If we transform such sentences with inanimate surface structure subjects into passive constructions we shall easily see that they may be divided into two groups. In one of those groups the prepositional expression 'przez + NP' becomes the equivalent of the surface structure subject, in the other the instrumental case:

Janek został uderzony cegłą 'John has been hit with a brick' (29b)
Sarna została zmiażdżona przez skałę 'A deer/The deer has been crushed by a rock' (30b)
Drzewo zostało złamane przez wiatr 'The tree has been broken by wind' (31b)

It seems reasonable to claim that this formal surface-structure difference corresponds to the speaker's intuition regarding the deep structure distinction between action-authors and instruments. It is only in examples of the first type (i.e. 29) that one can talk about the position 'instrument'. This supposition is confirmed by the transformational relations. It is not difficult to con-

struct sentences with action-authors specified for the utterances of the first type:

Jakiś chłopak uderzył Janka cegłą 'A (certain) boy hit John with a brick' (32)

But it is impossible to formulate such counterparts for the utterances of the second type (i.e. 30, 31). Compare other examples of this kind:

Rzeka podmyła brzeg 'The river has undermined the bank' (33)

Morze wyrzuciło szczątki statku na mieliznę 'The sea has thrown a wreckage on the rocks' (34)

Piorun zabił krowę 'The thunderbolt killed (struck) a cow' (35)

Attendant position is provided by such verbs as *kłócić się* 'quarrel', *walczyć* 'fight', *bić się* 'beat each other', *zgadzać się* 'agree', which occur either with the preposition *z* 'with' or without it; in the latter case both attendants are expressed by the nominative and the verb shows up in plural, compare:

Janek bije się ze Staszkiem||Janek i Staszek biją się 'John and Staszek beat each other' (36)

Janek kłóci się ze Staszkiem||Janek i Staszek kłócą się 'John and Staszek quarrel' etc. (37)

Since the position of both attendants can also be filled by coordinated constituents, and since one of these positions can optionally be empty (unexpressed, unspecified), some ambiguities may arise:

Janek walczy na froncie 'John is fighting at the front' (38)

Janek i Staszek walczą na froncie 'John and Staszek are fighting at the front' (39)

Janek i Staszek walczą ze sobą 'John and Staszek are fighting|| ||fight with each other' (40)

Janek i Staszek walczą z nieprzyjacielem 'John and Staszek are fighting||fight with the enemy' (41)

Janek i Staszek walczą

a) 'John and Staszek fight (with each other)'

or

b) 'John and Staszek fight (with somebody else)' (42)

The coordinated Noun Phrases *Janek i Staszek* in sentences 39 and 41 represent the same syntagmatic position as the simple constituent in sentence 38, while in sentence 40 *Janek* and *Staszek* represent different syntagmatic positions. Given these distinctions, sentence 42 is ambiguous.

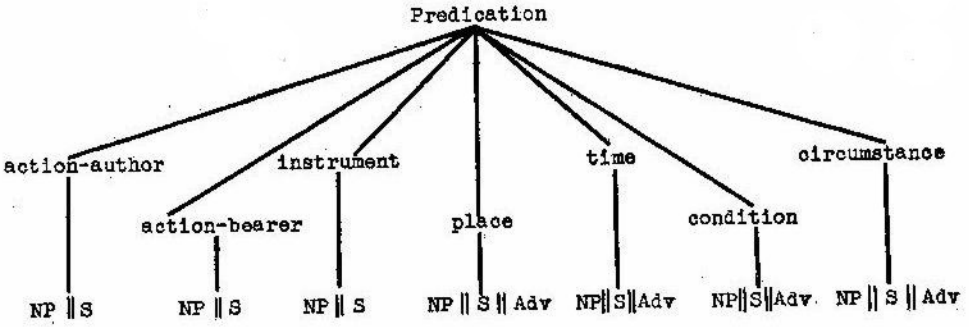
It has been shown (cf. Karolak 1963) that there can be only one specimen of any syntagmatic position in a simple sentence. From this point of view it is worth noticing the difference between such functions as 'circumstance' and 'time' or 'place'. Consider the following Polish sentence:

Janek szedł pewnego razu przez las wśród burzy 'John went once through a wood during a storm' (43)

The constituent *wśród burzy* has to be distinguished as a separate function distinct both from 'place' and 'time' (cf. Klemensiewicz 1963: 46) for the sentence has already constituents in these functions ('time': *pewnego razu*, 'place': *przez las*). This supposition is supported by the transformational relations in the sentence. 'Circumstance' can be taken away from the sentence and expressed by means of a coordinate sentence:

Janek szedł pewnego razu przez las i była burza 'John went once through a wood and there was storm' (43a)

Syntagmatic positions can be filled by various paradigmatic categories, first of all — if we abstract from predication — by Noun Phrase and Sentence (NP, S), then by Adjective, Adverb. Since it is predication which is the central member of the sentence, functional properties of deep structure categories may be represented schematically in terms of a dependency grammar as follows:



In principle, every such position can be taken by NP or S (i.e. an embedded sentence). Consider:

Janek często śpi po obiedzie 'John often sleeps in the afternoon' (44)

Kto jest zmęczony, ten dobrze śpi 'He who is tired sleeps well' (45)

Chłopiec czuje swoją samotność 'The boy feels his loneliness' (46)

Chłopiec czuje, że jest samotny 'The boy feels that he is lonely' etc. (47)

But in fact there is no consistent parallelism here (cf. Polański 1967: 43–45). There are positions which cannot be taken by an embedded sentence (e.g. 'action', 'state'), while other positions are exclusively or at least preferably filled by sentences, as the positions 'result' and 'condition' in Polish.

When a syntagmatic position in the deep structure admits a sentence (i.e. an embedded sentence), special problems arise. Since co-occurring sentences often refer to the same extralinguistic elements some constituents must be repeated:

Janek znalazł pióro. Okazało się, że pióro było własnością kolegi Janka. Kolega zgubił pióro wracając ze szkoły. Etc. 'John found

a pen. It turned out that the pen belonged to John's colleague.
The colleague lost the pen when returning from school' (48)

Real repetition of such constituents is only one of the possibilities language has. Others are ellipsis, pronominalization, and prosententialization (cf. Polański 1967: 64–69). Ellipsis is a very common device in inflectional languages and is closely connected with the functional properties of the elided categories. In Slavic languages, for instance, surface structure subjects are very often elided:

Janek wrócił do domu. Siadł i zaczął czytać 'John came back home. He sat down and started reading' (49)

In English, pronominalization occurs in such cases.

Prosententialization allows to avoid the repetition of whole sentences. It stands for the sentence as pronominalization stands for nominal and adverbial phrases. The means of prosententialization are some pronominal forms and several abstract nouns of general meaning like *fact* and *question*.

Prosentential exponents can occur only in positions where the special kinds of embedded sentences are admitted[3]. Once more we must refer to the lexical part of predication, since it is on this that such positions depend. In Polish such positions are allowed, for instance, by the verbs *wskazywać* 'show', *świadczyć* 'testify', *dowodzić* 'prove', the adjectives *słuszny* 'just', *dobry* 'good', the nouns *prawda* 'truth', *fałsz* 'falseness', etc. Some of them admit two such positions:

Na to, że tu było pierwotnie długie *e*, wskazuje fakt, że ...
'That there was previously a long *e* here shows the fact that ...' (50)

We can simplify this utterance by transforming it into several succeeding sentences and using prosententialization:

Tu było pierwotnie długie *e*. Na to wskazuje fakt, że ...
'There was previously a long *e* here. It is shown by the fact that ...' (50a)

It is clear that functions like action-author, action-bearer etc. are realized in such cases by whole embedded sentences. This is possible because the notions involved are abstract and really denote a meaning of the sort 'something that causes the action', 'something that was influenced by the action' etc.

The current emphasis on deep structure problems has caused considerable neglect of surface structure analysis. In my opinion a full analysis of a sentence should consist of both its deep structure interpretation and a careful inquiry into its surface structure phenomena, which do not always depend on the deep structure. Here such concepts as 'subject' and 'object' are very useful. Consider the formation of questions in English. Their structure differs according to whether or not the interrogative element which introduces them

[3] I call such embedded sentences intensional clauses (cf. Polański 1967: 93 - 139).

shows up in the function of subject. In the one case the surface structure of direct and dependent questions is identical, while in the other it differs:

Who teaches him mathematics? (51)
Do you know who teaches him mathematics? (51a)
What does he teach? (52)
Do you know what he teaches? (52a)

The linguist's task is to establish the deep structure on the basis of surface structure data and then to provide rules for transforming deep structures into surface structure utterances.

REFERENCES

Chomsky, N. 1964. "Current issues in linguistic theory". In J. A. Fodor and J. J. Katz (eds.), *The Structure of language*, 50 - 118. Englewood Cliffs, N. J.: Prentice-Hall, Inc.
— 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: M. I. T. Press.

Fillmore, J. 1968. "The case for case". In E. Bach and R. Harms (eds.), *Universals in linguistic theory*, 1 - 88. New York: Holt, Rinehart and Winston, Inc.

Gołąb, Z. 1958. "Problem podmiotu". *Biuletyn Polskiego Towarzystwa Językoznawczego 17*. 19 - 66.

Jakobson, R. 1963. *Essais de linguistique générale*, trad. Ruwet. Paris: Arguments 14.

Karolak, S. 1963. "Pojęcie pozycji syntagmatycznej a wyodrębnianie jednostek języka". *Biuletyn Polskiego Towarzystwa Językoznawczego 22*. 131 - 142.

Klemensiewicz, Z. 1963. *Zarys składni polskiej*, wyd. 4. Warszawa: Państwowe Wydawnictwo Naukowe.

Kuryłowicz, J. 1948. "Les structures fondamentales de la langue: groupe et proposition". *Studia Philosophica 3*. 203 - 209.

Lee, P. G. 1968. "Subjects and agents". In *Working Papers in Linguistics 3*, Technical Report 69 - 4, Computer and Information Science Research Center. 37 - 113. Columbus, Ohio: The Ohio State University.

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

Polański, K. 1966. "Główne typy struktur zdaniowych w języku polskim". *Zeszyty Naukowe WSP w Katowicach, Prace Językoznawcze 3*. 83 - 99.
— 1967. *Składnia zdania złożonego w języku górnołużyckim*. Wrocław — Warszawa — Kraków: Ossolineum.
— 1969. "Sentence modality and verbal modality in generative grammar". *Biuletyn Fonograficzny 10*. 91 - 100.
— 1970. "The place of generative grammar in present day linguistics". *Studia Anglica Posnaniensia 2*. 19 - 29.

Staal, J. F. 1967. "Some semantic relations between sentoids". *Foundations of Language 3*, fasc. 1. 66 - 87.

Trubetzkoy, N. S. 1939. "Le rapport entre le déterminé, le déterminant et le défini". In *Mélanges de linguistique offerts à Charles Bally*. 75 - 82.