Witold Marciszewski

SYNTACTIC DESCRIPTION OF REPORTED SPEECH IN CATEGORIAL GRAMMAR

Originally published as "Opis syntaktyczny mowy zależnej na gruncie gramatyki kategorialnej," Studia Semiotyczne 7 (1977), 205–226. Translated by Magdalena Tomaszewska.

1. The main ideas of categorial grammar

1.1. The word "grammar" shall be understood here as a synonym to "syntax," thus more narrowly than it is understood by some linguists (Levelt 1974: 2-3). Categorial grammar is one of the alternative grammars which are within the interest of linguists; It comes from K. Ajdukiewicz (1935) who referred to Husserl's ideas and some results of Leśniewski and Łukaszewicz. Husserl originated the conception of syntactic categories, Leśniewski applied categorial grammar to languages of logic, and Łukaszewicz created the positional language which is used in logic under the name Łukaszewicz's notation or Polish notation.

What turned out to be an important achievement was Ajdukiewicz's algorithm which allows us to settle for any expression of a given language (for which categorial grammar is given) if this expression is built correctly, that is if it is grammatical, that is — as Ajdukiewicz put it — syntactically coherent (Ajdukiewicz 1935). This conception was a quarter of a century ahead of the currently vivid idea of language as a decidable system (Bar-Hillel 1970).

The result mentioned above was obtained before the Second World War and gained interest among philosophers and logicians of that time. The merit of popularizing the results among linguists after the war and supplementing the results with new ones belongs to Y. Bar-Hillel. Had it not been for his activity, such a development of categorial grammar, observable in recent years, most probably would not have been taking place (Levelt 1974: 96).

1.2. The concept of syntactic category is defined by abstraction. The definition follows that of Ajdukiewicz, the only difference being the change of "semantic category," present in the original text, to "syntactic category."

"The word or expression A, taken in sense x, and the word or **e**xpression B, taken in sense y, belong to the same syntactic category if and only if there is a sentence (or sentential function) S_A in which A occurs with meaning x, and which has the property that if S_A is transformed into S_B upon replacing A by B (with meaning y), then S_B is also a sentence (or sentential function" (Ajdukiewicz 1967c: 208).

Categories divide into basic categories, which include sentences and names, and functor categories. The term "functor" (funktor) comes from T. Kotarbiński and is common in Polish logical texts; it covers function symbols, predicates, sentential connectives, but also adverbs, prepositions, etc. Thus, it corresponds to what was called syncategorematic expressions in medieval logic, while basic categories overlap with the class of categorematic expressions. The expressions of basic categories either occur as independent units of text (e.g. constituting the text of a sentence), or are arguments of functors. And the functor, having always one or more arguments, can itself be sometimes an argument of a higher-order functor. For there is a hierarchy among functors, which has some analogies to Russell's hierarchy of types (here it is worth mentioning Leśniewski's contribution to the method of preventing antinomies which was an alternative to the theory of types).

1.3. There are three methods of characterizing relationships between functors and their arguments within the analyzed expression: a) the traditional system of brackets, b) the so-called quasi-arithmetical notation, c) the method of numbering syntactic positions which shares its underlying idea with Łukaszewicz's bracket-free notation. The two latter methods come from Ajdukiewicz, however only the quasi-arithmetical notation was studied further and improved by other authors (Bar-Hillel 1970).[1] It shall be applied here in Ajdukiewicz's original version, because the mentioned improvements are useful only when a positional language is not used, and the present considerations shall use this language significantly.

1.4. Here follow examples of syntactic analyses that employ the three mentioned notational methods.

(S) John eats sour apples.

Distinguishing arguments by means of brackets will result in:

(S_1) (John) eats [sour (apples)],

[1]This notation was used in notational systems in P. Geach (1970) and M. J. Croswell (1973).

or

(S_2) (John) {eats [sour (apples)]}.

The difference between the two depends on what is assumed to be the sentence-forming functor for (S) — the two-argument predicate "eats," or a one-argument complex predicate "eats sour apples." The latter could be further analyzed as the functor-forming functor "eats" and the argument "sour apples," which in turn is a complex name formed by means of the name-forming functor "sour."

1.5. In quasi-arithmetical notation, named so because indicators of functor categories look like common fractions, the category of functor is characterized in such a way that what is put in the numerator is the category of expression formed by a particular functor, while in the denominator — categories of arguments of this functor. Let the letter s stand for the category of sentences, and n the category of names. The components of (S) get the following indicators of categories. According to the analysis in (S_1):

John: n, eats: s/nn, sour apples: n;
apples: n, sour: n/n.

According to the analysis in (S_2):

John: n, eats sour apples: s/n;
sour apples: n, eats: s/n//n;
apples: n, sour: n/n.

Note: the expression s/n//n is a linear equivalent of a complex fraction: $\frac{\frac{s}{n}}{n}$.

For typographical reasons, linear equivalents shall be used here. A bigger number of slashes reflects what is conveyed by a greater length of the line in complex fractions.

1.6. The third method of syntactic description refers to the idea of positional language, that is, such in which the position of an expression indicates its syntactic role, just as it is in Łukaszewicz's bracket-free notation. An expression, let's say W, can be notated by means of positional notation because it has a main functor which, together with its arguments, forms the whole of expression W; these components are called first-order elements (of expression W). If any of the first-order elements are complex expressions, they are further decomposed into second-order functor and arguments, etc.[2]

[2]This method also comes from K. Ajdukiewicz (1967a, 1967b).

(S_1)	John	eats	|sour apples;|	sour	apples.
	1	0	2	2.0	2.1

(S_2)	John	|eats sour apples;|	eats	|sour apples;|	sour;	apples
	1	0	0.0	0.1	0.1.0	0.1.1

Let first-order elements be marked with one-position numbers, for example let's mark the functor with 0, and its consecutive arguments from 1 up. Second-order elements are characterized by numerical expressions composed of two digits. The first digit indicates that the component belongs to a particular element of the former order (e.g. first-order), while the other digit indicates the position o this component in the analyzed element, etc. Application of this method to the considered examples results in:

The result of the analysis of (S_1) can be presented as a linearly ordered set by means of a relation which could be defined as a relation of syntactic subordination (or dependence). Here follows the set:

0, 1, 2.0, 2.1.

By analogy, the expression (S_2) is assigned with the following sequence of indicators of syntactic positions:

0.0, 0.1.0, 0.1.1, 1.

If we ordered individual words in the expressions (S_1) and (S_2) according to the sequence of indicators of syntactic positions that correspond to these words, then we would translate these expressions into a positional language (such a translation could be called a syntactic translation). The language of logic provides simple illustrations of the translation of this type. For example, the expression:

$$CApqCNpq$$

is a positional equivalent of the expression:

$$(p \vee q) \rightarrow (\sim p \rightarrow q).$$

The possibility of employing positional notation to any expression is an important feature of categorial grammar. It shows that there is a linear

order in the set of simple expressions that are components of any complex expression.

1.7. A synthesis of the two approaches discussed above, namely that of assigning expressions with indicators of categories and linear ordering of a sequence of expressions, is the algorithmic method of testing the syntactic cohesion of expression, also coming from Ajdukiewicz (1935). This algorithm consists of the following rules: 1° when testing the cohesion of expression W, begin with ordering the set of indicators assigned to particular elements of W, according to the order of expressions which would occur in a linear version, that is in the positional notation. 2° Going from left to right, observe if there is in the sequence a pair of neighboring indicators in which the first element is a fraction indicator and the following indicator is such as the denominator of this first fraction indicator. 3° Whenever you find such a pair, cross both identical elements out but leave the numerator. 4° Continue the procedure described under 2° and 3° as long as the sequence can no longer undergo reduction. 5° If this final result is a one-element sequence, that is a single indicator, then the analysed expression is syntactically cohesive and has the category of this final indicator; if the final sequence has more elements, then the analyzed expression needs to be regarded as syntactically incohesive, that is grammatically incorrect.

Examples of such a procedure, widely known and much popularized, shall be provided further in the article.

2. The combinatorial method of syntactic description of reported speech

2.1. If one and the same expression (in terms of shape) can have more than one syntactic structure, then a question arises about how many such structures can be assigned to expressions belonging to the structure called reported speech. It needs to be highlighted that such a "pluralistic" approach to the problem is not common among authors dealing with the subject of reported speech. Most of them take the stance that reported speech can have only one structure and only this one, which is believed to be appropriate, is analyzed. There are a few such "monistic" approaches; let's focus on three of them as they are the most widespread.

The view initiated by Frege (1892; see also Church 1956) can be reconstructed in categorial grammar in the following way: the word "that" functions as a name-forming functor of one sentential argument: "that (it) rains" is the name of the proposition which, in direct speech, is expressed by "(it) rains;" "that (it) rains" can be also interpreted as the name of a state of raining. This approach is significantly supported by grammatical properties of the German language, where the structure "dass p" (with p

standing for a sentence) not only can function as grammatical subject (as is also the case in English), but can also naturally occur with a preposition, for example in the structure of the type "ohne dass p...;" thus, it occurs in all functions characteristic of names. Contemporary authors who postulate such an analysis of reported speech are, among others, R. Jeffrey and Y. Bar-Hillel (Bar-Hillel 1965, 1970; Jeffrey 1965).

A different approach, present in certain comments of Bar-Hillel (in writings other than the ones quoted above), is taken by K. Ajdukiewicz who treats "that" as a constituent of the functor "says that," "thinks that," etc., and not as a part of the construction "that p" (Ajdukiewicz 1967a) It is, then, a functor of one name argument and of one sentential argument. Ajdukiewicz does not comment on the internal structure of this functor: namely, if it is one simple word in which "that" functions as a sort of a syllable devoid of its own syntactic function, or if it is a complex expression in which "that" plays the role of a functor-forming functor of one functor argument. It is worth noticing that in both, the former and the latter variant of Ajdukiewicz's approach, there is no sense in putting a comma between "that" and the preceding verb;[3] it would be a syntactic error that requires disapproval, whereas — so far — our linguists and editors-purists disapprove of those who occasionally drop the comma in this position. Let's add here that if we accept that one and the same expression (in its graphic shape) can have many syntactic structures, and that punctuation is, similarly to accent in speech, treated as a means to express the author's intention with regard to the syntactic structure, then this controversial comma should be treated as optional in reported speech: let it be used by those who want to give their utterance the Fregean structure in reported speech, let it be dropped by those who want to use Ajdukiewicz's structure.

There is, finally, a traditional view of grammarians that the word "that" is a conjunction.[4] A conjunction is usually a sentence-forming functor of two sentential arguments. Here, however, it is not clear if we are dealing with such a classical conjunction, or a special variant in which the first argument is a certain functor for which there is an isomorphic expression which is a sentence but which occurs then outside of the context of reported speech. Indeed, "John says" can be a full sentence, for example, as an answer to "What does John do every day?," yet it is not the component of "John says that it rains," in the latter case, the expression "John says"

[3]As it is required in Polish language — translator's note.

[4]See, among others, S. Szober (1953). This view is further discussed on the grounds of modern formal linguistics (Wojtasiewicz 1972).

is not a full sentence, but rather a two-argument functor that occurs with its first argument and without the other. However, the view assumed in the considerations to follow, is that "that" is a functor of two sentential arguments, which is one of the acceptable descriptions of syntactic structure; whoever does not agree with the view can cross out an appropriate structure from the obtained list, by means of adding one more eliminating assumption to the set of accepted assumptions.

An approach similar to the traditional one was introduced by R. Martin who treats the structure "e That p" as a two-argument relation between an event e, which is an intentional act, and an event described by the sentence p. Admittedly, "that" is not a conjunction, but, similarly to a conjunction, constitutes an "axis of symmetry" of the syntactic construction with one argument on each side (Martin 1979).

2.2. The multitude of approaches on reported speech represented in the literature leads naturally to two questions: what is a criterion of choice between those approaches (if one had to choose), and how can one be sure that no structure that could be assigned to reported speech has been omitted? These questions are mutually independent, however, for the sake of clarity, it will be more convenient to start with the second one. In the first stage, then, we will try to obtain a full set of all structures, which can be obtained by applying combinatorial methods to a set of words that contribute to reported speech, no matter whether the obtained structures satisfy well-formedness conditions of the grammar of our language. The second stage will involve discussing and forming the assumptions which concern well-formedness, in other words, what will be the basis for elimination in the set of all (combinatorially) possible structures.

2.3. Let the sentence below be an example of an expression representing reported speech:

Jan	myśli,	że	pada.
John	thinks	that	it rains.

The Polish sentence is composed of four expressions. If the sentence after "że" was not a single word, then there would be more components. However, since it is not necessary to take the internal structure of this sentence into consideration, it will be always represented by one symbol. Let the four component expressions be represented, accordingly, by the following letters:

Jan — x, myśli — e, że — o, pada — p

John — x, thinks — e, that — o, it rains — p

The letters x, e, p can be regarded as variables of the, accordingly, name, functor and sentence category; the letter o is a constant, a symbolic equivalent to the word "że" (that) (chosen for mnemotechnic reasons, as an association to "oratio obliqua"). What we are dealing with, then, is a set of component elements of reported speech, that is: {x, e, o, p}. As indicated by the notation, the set is not ordered, thus the order of elements does not indicate any structure, and can be freely changed.

How many structures can be obtained from these four elements, taking into consideration all possible ways of ordering? In order to answer that question, let's introduce the notions of a linearly ordered structure and a dendritic structure (that is a certain partial ordering); I shall show that, for the four-element set, there are 24 linear structures, and that each of them is assigned 10 dendritic structures different from one another, which gives 240 as a maximal number of structures. It will be possible to reduce this alarmingly large set, by using the eliminating assumptions of the second stage, which will result in only 11 acceptable structures of reported speech left. Among them will be all structures presented in section 2.2.

2.4. The number of linear structures is 24, for they are obtained by permutations of a four-element set; thus it is 4!. Out of these 24 sequences we can immediately eliminate 12 — 6 starting with x, and 6 starting with p. For it is well known that, in positional notation, the word in the initial position should be zero or a sequence of zeros, thus it should belong to the category of functors without being — in a given expression — an argument of any functor. This category is not met by either the name x, nor the sentence p, for they belong to basic categories.

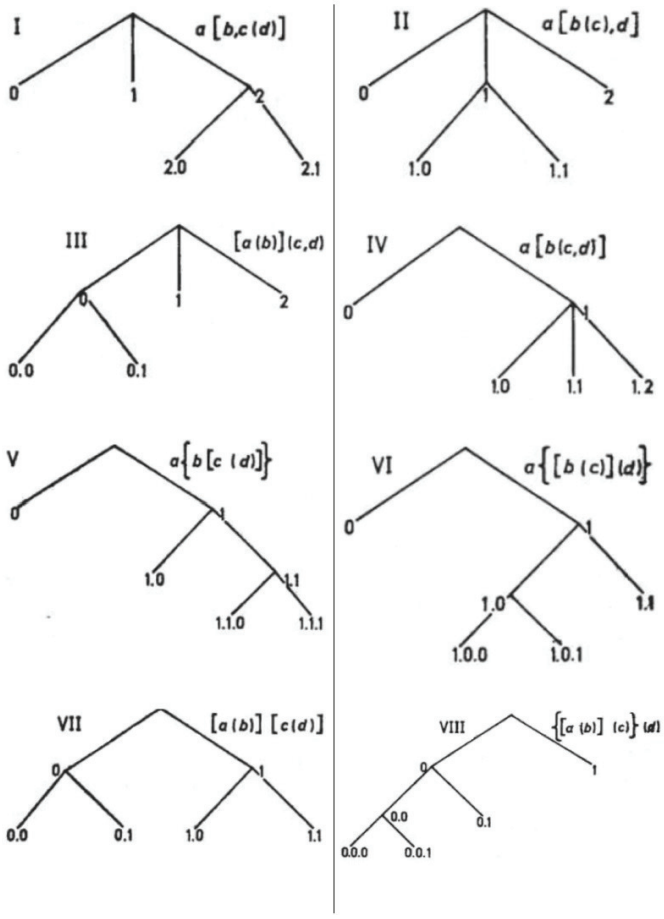
Having eliminated the above mentioned structures, there are still 12 following sequences:

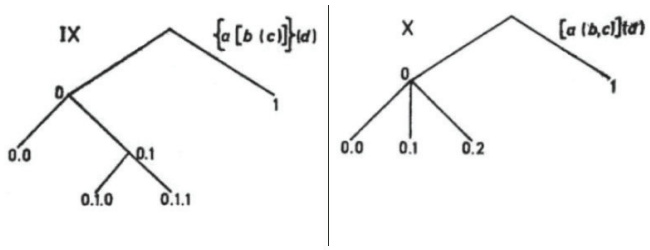
Tab. 1.

(1)	*o e x p*	(7)	*e o x p*
(2)	*o e p x*	(8)	*e o p x*
(3)	*o p e x*	(9)	*e x o p*
(4)	*o p x e*	(10)	*e x p o*
(5)	*o x e p*	(11)	*e p o x*
(6)	*o x p e*	(12)	*e p x o*

2.5. Each of the above linear structures is assigned with ten tree-structures, which I shall call dendritic. This follows on from the considerations below. All four elements of each linear structure can be grouped in various ways that can, in turn, be presented as trees. The number 10 results from the fact that the four-element structures here always split into two or three components, that is a functor with one argument, or a functor with two arguments. Let's consider all such alternatives. Each dendritic structure will be accompanied by a notation with brackets in which the brackets appropriately group the structure's components. The variables a, b, c, d can be substituted with the words from the set {x, e, o, p} in such an order as they appear in the sequence under consideration. If the sequence is e.g. o e x p, then the bracket notation for the first dendritic structure is o[e, x(p)].

Tab. 2.





Thus, each of the 120 structures that is a result of the Cartesian product of the set (1) ... (12) and the set I ... X can be marked with a two-digit symbol indicating a pair of elements in which each comes from a different set: (1)I, (1)II ... (2)I ... (2)X, (3)I ..., etc. By means of such symbols it will become possible to indicate which structures will be rejected as not meeting the syntactic conditions, and which will resist the elimination.

3. Syntactic assumptions adopted as a basis for elimination of wrong structures and the results of elimination

3.1. Categorial grammar of any artificial language, such as the language of logics, is included in its dictionary, in which each word is assigned with one particular syntactic category; a word thus characterized in every context appears in the permanently assigned syntactic category and unequivocally indicates, together with the remaining context elements, its syntactic structure (Levelt 1974: 107). Natural languages are different as along with the phenomenon of polysemy, there is the phenomenon of syntactic ambiguity which consists of one (isomorphically and semantically equivalent) word that can be assigned to more than one category, playing — depending on the context — one or another role. Thus, for example, the word "and" is sometimes a sentence-forming functor of sentence arguments (s/ss), but sometimes a name-forming functor of name arguments (n/nn); we can talk here about semantic equivalence in the sense that both cases concern the same operation of logical multiplication. By analogy, adjectives play the role of names when they function as (subject) complements, or the role of name-forming functors when they function as attributes; semantic equivalence here consists in the connotative equivalence. This syntactic ambiguity has usually been regulated to some extent, that is it has been established that a given word or a class of words has such-and-such syntactic category in one context, and such-and-such syntactic category in other contexts.

There are, however, such words in natural languages that not only are assigned to more than one syntactic category, but also their set of syntactic categories is not sufficiently defined; Thus, what we are dealing with here

is an equivalent of what is called polysemy with reference to meaning, and at the same time an equivalent of vagueness (semantic instability). This is exactly the case of the Polish że "that." It does not mean that interpretation possibilities of such words are unlimited. It is reasonable to assume that the issue of belonging to the basic or the functor category will be resolved (in a particular type of context). However, where there is a multitude of functor categories in a given language, and different functors of functors among them, there is a possibility of linking elements in more than one way, and different ways of linking particular functors with various syntactic categories. In such a situation, in order to define the class of possible syntactic interpretations, it is necessary to clearly state what is assumed about the analyzed complex expression (or a particular class of expressions) and its component elements. These assumptions can even be controversial, but being controversial will not be a drawback if the consequences of each assumption is shown. Then whoever does not accept a questioned assumption does not have to accept its consequences; if one does not accept one of the consequences, however, they have to reject the assumptions which constitute the premises.

3.2. The following assumptions are postulated as a test for all 240 structures obtained here; in practice only 120 will remain since half will be immediately eliminated on the basis of assumptions A.1 and A.2 on the list below:

A.1. An expression represented by variable p is not a functor.
A.2. An expression represented by name variable x is not a functor.
A.3. When p or x is one of the arguments of a two-argument functor, then x is the first and p is the second argument.
A.4. An expression represented by variable e, and also by formula e(x) — if it can be interpreted as a functor, is a predicate, that is a sentence-forming functor of name arguments, and in the case of e there are one or two arguments, and in the case of e(x) there is one argument.
A.5. Expression o (a symbolic equivalent of "that") is never an argument of a functor.
A.6. Expression o never has name arguments.
A.7. If an expression represented by o(p) is a functor, then it has no name arguments.

What can be explained by adopting particular assumptions from the list above? A.1 and A.2. are true ex definitione: variables of basic categories cannot represent functors. The remaining assumptions are based on verdicts

of authority called native speaker; if the author of this article is wrong in assigning this or that judgment to this authority, then of course he can be corrected by anybody who has the right to play the role of a native speaker. The following restrictions seem to be valid for the matters in question.

Ad A.3. The word order in reported speech requires the following order of arguments: 1 — name, 2 — sentence, and a change in this word order, for example "it rains thinks that John," is always felt to be a violation of syntactic rules.

Ad A.4. Expressions e and e(x) are not names, and also e is not a sentence, thus it needs to be a functor. The issue of arguments of this functor is more controversial. On the grounds of the Polish language it could be claimed that the first argument is not necessary as it is not present (at least in the surface representation) in impersonal expressions, e.g. "sądzi się, że" (it is thought that). In languages such as English and German, which always have personal (one, man) or impersonal (it, es) subjects, this claim fails. On the other hand, the second argument in these languages needs not be a name argument; this is at least how it looks like "on the surface." In reported speech, whenever "that" or "dass" is omitted, what immediately follows the functor e is a sentence. The problem whether the finding should be made on the basis of surface structure or rather a default hidden linguistic object that binds e with p will remain open, though such an object could be a zero symbol or a syntactic construction itself. If such solutions are not adapted, then A.4 remains valid only for such languages as Polish, where omitting o before p violates syntactic rules.

Ad A.5 There is no such functor expression with which "że" (that) as an argument would create a coherent whole or at least a coherent piece of a whole. Coherent pieces that include o are always of the form eo or op; in both cases o functions as a functor. By the way, this indicates that there are more than two levels of completeness or incompleteness. It is not only among basic categories that a diversification into the more complete (sentences) and the less complete (names) occurs; the diversification also occurs among functors, though by a different principle: there are such functors that can be arguments of other functors, and such that cannot. The latter have the lowest level of completeness; it seems that except for "że" (that) also the functor "bardzo" (very) is a convincing example.

Ad 6. Again, we refer to the linguistic competence of a native speaker, who would say that no construction of the type "że x" (that John) is a coherent piece of a sentence, thus it cannot make, together with other expressions, a grammatically correct sentence. However, the constructions eo and op are

examples of a coherent piece with ”że” (that).

Ad 7. This assumption concerns a rather special interpretation according to which o(p) is a functor functioning similarly to a post-verb adverb answering the question ”how” (the so called adverbial of manner). For it is possible to consider the content of an intentional act to be one of its property, thus finding this property would answer the question of what this act is, that is (in the adverbial version) how it is represented, what it is about, etc., which is reflected in the following colloquial conversation.

— I think she will not come. And what do you think?
— I think she will, though.

What is claimed in A.7 is not that such an interpretation is (sometimes) accurate, but rather something less: namely, certain restrictions are imposed on this interpretation in case somebody wants to accept it; what is claimed is that even if o(p) is a functor, then it is not a functor of a name argument. Indeed, any native speaker will admit that sentences ”John that (it) rains” or ”(it) rains that John” are nonsensical.

3.3. Let’s apply assumptions A.1 — A.7 to the list of 120 constructions obtained through multiplication of table 1 by table 2. The list of the obtained pairs whose first element is an Arabic number from table 1, and the second — a roman number from table 2, is also most conveniently presented in a table, whose first column contains Arabic numbers, while the first row — roman numbers. The intersections of the two contain the numbers of assumptions on the basis of which a given construction can be rejected if it can be rejected; if a construction does not fail any of the seven tests, then the appropriate intersection is marked with ”+.”

Tab. 3

	I	II	III	IV	V	VI	VII	VIII	IX	X
(1)	2	+	+	4	2	4	2	+	+	6
(2)	1	3	3	6	1	4	1	+	4	+
(3)	3	1	3	1	1	1	+	4	1	3
(4)	2	1	7	1	1	1	2	7	1	6
(5)	6	2	6	2	2	2	4	6	2	6
(6)	1	2	6	2	2	2	1	6	6	6
(7)	2	6	5	6	2	6	2	5	6	5
(8)	1	3	5	6	1	3	1	5	+	5
(9)	+	2	5	2	2	2	+	5	2	5
(10)	1	5	5	5	5	5	5	5	5	5
(11)	6	1	4	1	1	1	6	5	1	5
(12)	5	5	5	5	5	5	5	5	5	5

The reasoning leading to the conclusions of table 3 will not be presented here, since they are so simple that the reader who wants to check the results can easily reconstruct the reasoning, except for the positive results (the variants that survived the elimination tests) which deserve a more detailed treatment. In the list below they are given in three notations: with brackets, with numbers (that indicate the syntactic position and can be easily transformed into a tree), and with indicators ordered according to the principle of positional notation, that is the principle assumed by the bracket-free notation.

(1)II	*o*	*e*	*x*	*p*	*o*[*e*(*x*), *p*]	1
	0	1.0	1.1	2		
	s/*ss*	*s*/*n*	*n*	*s*		
(1)III	*o*	*e*	*x*	*p*	[*o*(*e*)](*x*,*p*)	2
	0.0	0.1	1	2		
	s/*ns*//*s*/*n*	*s*/*n*	*n*	*s*		
(1)VIII	*o*	*e*	*x*	*p*	{[*o*(*e*)](*x*)}(*p*)	3
	0.0.0	0.0.1	0.1	1		
	s/*s*//*n*///*s*/*n*	*s*/*n*	*n*	*s*		
(1)IX	*o*	*e*	*x*	*p*	{*o*[*e*(*x*)]}(*p*)	4
	0.0	0.1.0	0.1.1	1		
	s/*s*//*s*	*s*/*n*	*n*	*s*		
(2)VIII	*o*	*e*	*p*	*x*	{[*o*(*e*)](*p*)}(*x*)	5
	0.0.0	0.0.1	0.1	1		
	s/*n*//*s*///*s*/*n*	*s*/*n*	*s*	*n*		
(2)X	*o*	*e*	*p*	*x*	[*o*(*e*,*p*)](*x*)	6
	0.0	0.1	0.2	1		
	s/*n*//*s*/*n*,*s*	*s*/*n*	*s*	*n*		
(3)VII	*o*	*p*	*e*	*x*	[*o*(*p*)][*e*(*x*)]	7
	0.0	0.1	1.0	1.1		
	s/*s*//*s*	*s*	*s*/*n*	*n*		
(8)IX	*e*	*o*	*p*	*x*	{*e*[*o*(*p*)]}(*x*)	8
	0.0	0.1.0	0.1.1	1		
	s/*n*//*n*	*n*/*s*	*s*	*n*		
(9)I	*e*	*x*	*o*	*p*	*e*[*x*,*o*(*p*)]	9
	0	1	2.0	2.1		
	s/*nn*	*n*	*n*/*s*	*s*		
(9)VII	*e*	*x*	*o*	*p*	[*e*(*x*)][*o*(*p*)]	10
	0.0	0.1	1.0	1.1		
	s/*n*//*n*	*n*	*n*/*s*	*s*		
(1)III*	e_0	*x*	*p*		e_0(*x*,*p*)	11
	0	1	2			
	s/*ns*	*n*	*s*			

The last item on the list is not represented in table 3 because the morpheme "że" (that) does not have the status of an expression but is reduced to the role of a syllable. However, since, in every other respect, it resembles construction (1)III, this variant is marked as (1)III*.

3.4. Among these 11 resultant structures, there are all the familiar structures discussed in written sources, namely:

(1)II — inter alia, traditional grammar
(1)III — K. Ajdukiewicz
(9)I — inter alia, G. Frege.

Most of the remaining structures can be assigned with some intuitive readings. Thus, for example, according to (1)VIII, a sentence of the type "John thinks that it rains" resembles sentence-forming modal functors: this modality of sentence p consists in that it is thought, and it is thought by John. Similarly, (1)IX, though the internal structure (the order of constructing particular elements) of the latter is different in this quasi-modal functor.[5]

(2)VIII and (2)IX predicate about x that x thinks (or says, believes, etc.) this-or-that; the difference between these two variants of structure (2) again consists in the internal structure of this complex predicate. Also, (8)IX represents a predication about x, yet the structure of the predicate is different than in variants (2).

Contrary to this, in construction (9)VII the subject is the expression "że p" (that p) which is a name of a proposition created from the sentence (the proposition is expressed by the sentence), while what is predicated about the proposition is, for example, that John is in the act of thinking about the proposition; instead of a proposition (judgment), it would be also possible to talk here about a proposition (state of affairs).

This review omits (3)VII which seems to be a construction totally devoid of intuitiveness, but has been incorporated into the list according to game rules: it cannot be rejected since it survived the elimination criteria. Whoever wants to cross it out due to the intuitiveness of interpretation, should express their own intuitions in the form of an additional assumption on the list of adopted criteria.

4. Conclusions, hypotheses, problems

[5] An extremely interesting analysis of such constructions of modal expressions, accompanied by a rich bibliography on modal logics and deontic logic, was presented by G. Kalinowski (1973, cf. especially p. 183-190).

4.1. The presented considerations lead to two conclusions: one — more detailed and about the content, concerning reported speech, and the other — more general and about the method, concerning a certain research method of language.

The first conclusion is: there is more than one acceptable structure of reported speech. Whoever does not agree with this carries the burden of proof (onus probandi), that is, needs to supplement the list of assumptions in such a way as to eliminate all syntactic structures except for one. The author of this article does not consider it likely unless some restrictions, which are contrary to linguistic practice and artificial, are imposed on the language. Still the way to such attempts, even while adopting the method presented here, is wide open.

The so-formulated conclusion gives rise to a problem. The structures considered here belong to the sphere of language (la langue) which, in the field of linguistic research, is complemented by the sphere of speech that is utterance (la parole). The question is: do the described syntactic structures and the differences between them correspond to a diversification of phenomena in the sphere of parole, e.g. changes in word-order, pauses, stress. If the answer is affirmative, then another question arises: what is this correspondence? What, for example in the manner of speaking, corresponds to differences between the structures:

(1) II o[e(x), p] ("traditional")
(1) III [o(e)] (x, p) (based on Ajdukiewicz)
(9) I e[x, o(p)] (based on Frege).

The hypothesis posed here, which is an indirect link between the input (in this article) conclusion and the present question, is the affirmative answer to the previously posed question. I claim, namely — hypothetically — that at least in some cases, the differences in speaking or writing can be recognized. In the case of written utterances there are two phenomena that come into play: punctuation and word-order. The punctuation matters include the exciting issue of a comma before "że" (that). Those who advise the use of it in Polish or German sentences, express, at least, that they oppose treating the cluster of words oe (e.g. "mówi, że" (says that)) as one syntactic whole (when the rule that a comma separates is taken seriously). Here I oppose treating this punctuation rule as universal, for it would be contrary to the presented multitude of acceptable structures of reported speech, though the mere fact that some language experts feel that there is a caesura before

”że” (that), seems to be, in relation to a particular language, in favor of ”traditional” structure or the one based on Frege, and those against the structure, among others, based on Ajdukiewicz. What speaks in favor of the view that the approach based on Frege is sometimes executed in speech and writing seems to be the characteristic separation of ”że” (that) in the case when it is followed by a conjunction or a disjunction sentence. It is possible to say ”Jan mówi, że pada i grzmi” (John says that it rains and thunders), but equally it is possible to say ”Jan mówi, że pada i że grzmi” (John says that it rains and that it thunders). The same separation will apply to ”Jan mówi, że pada lub mży” (John says that it rains or drizzles).

In German, combining ”dass” with the sentence that follows results in a more autonomous whole, namely a construction of the nature of a complex name, which is shown in the possibility of combining this construction with prepositions, for example in the sentence: ”Man kann gelernt haben, logisch korrekt zu schliessen, ohne dass man dabei irgendein Gesetz der Logik explizit können müsste.”

Are there also, on the phonetic level, recognizable differences corresponding to e.g. structural differences between (1)I, (1)II and (9)I? As a hypothesis for further study which seems to be a result of common and immediate observations, I would claim that pauses have different distributions in different variants of phonetic representation of reported speech, which results in differences in stress, and can be assigned to the three discussed structures in the following manner:

syntactic structure	phonetic structure
(1)II $o[e(x), p]$	$x\ e \mid o \mid p$
(1)II $[\ o(e)](x, p)$	$x\ e\ o \mid p$
(9)I $e[x, o(p)]$	$x\ e \mid o\ p$

4.2 If the above assumption turned out to be correct, the following question should be raised: What semantic differences are assigned to these syntactic differences that can be represented in different phonetic variants? Solving this issue requires treating the discussed utterances as answers to particular questions; for a question variable determines semantic stress in the answer, and the stress determines pauses that can indicate syntactic split. For example, in the sentence: ”What is John’s (belief) relation to life after death?,” the question variable belongs to the set of possible relations: beliefs, assumptions, doubts, rejections, etc. And the answer could be, for example:

John believes | that life after death exists,

with a strong stress on "believes," which requires a pause before continuing, and in which "that" merges smoothly with the phrase following the pause. If we accept that phonological properties of this type correspond to any of the possible syntactic structures (not necessarily unambiguously), then the most natural candidate in this case seems to be the structure based on Frege.

An informative example of stressed "that" can be found in a particular text by K. Ajdukiewicz, in which stress is marked with italics: "But it does not suffice if he [the teacher] tells the pupil that he [the pupil] committed an error; he also ought to point out where the error is and in what it consists" (Ajdukiewicz 1974: 1). It is visible in the context that when sentences starting with "where" and "in what" are answers to certain open-ended questions, the sentence starting with "that" answers the open-ended question "Did the student make a mistake?" Thus, the context with known or inferred questions determines stress in reported speech, and stress determines (not necessarily unambiguously) the syntactic structure. In the given example, stress indicates that it would be unusual to treat "że" (that) as a dependent syllable in "powiedzieć, że" (say that); because it is unusual for semantic stress to fall on a meaningless morpheme.

4.3. Another conclusion that comes to mind is a certain methodological program for syntactic description; an illustrative (but not necessarily "exemplary") application of this program was presented here with regards to reported speech, and consisted of the following five steps:

1° differentiating the most simple elements in the studied type of utterance; 2° listing all ordered sequences of these elements as well as the variants resulting from different groupings of these elements; 3° formulating assumptions specifying the conditions that a syntactic structure of the given type needs to meet; 4° eliminating sequences that do not meet the criteria; 5° indicating semantic and phonological variants assigned to the remaining syntactic structures.

In the present study on reported speech, the first four steps have been conducted, while the fifth remained in the sphere of hypotheses and questions.

It is beyond doubt that such a program for syntactic description, if considered purposeful, can be applied to all types of complex expressions. A question arises, however, if an analogous method can be applied to units of higher-order than sentences, namely multi-sentence texts. Also to these texts the condition called cohesion is intuitively applied, though it is not the

same cohesion, and it is not as easy to characterize as syntactic cohesion sensu stricto. For convenience, in the first stage of the task, we will restrict the type of text to not too extensive and relatively less complex one, such as a paragraph.

Let's notice that except for the "intra-sentence" conjunctions forming complex sentences, there are "compound-sentence" conjunctions, which can be called text-forming functors. In Polish, these are: "więc" (so), "bowiem" (for), "ponadto" (moreover), etc. They are usually part of the "inside" of one of adjoining sentences, but their function does not hold when the sentence is in isolation; this function is to refer to the preceding sentence. If there are no such linking words between sentences in a text, then the "text-forming" functor is possibly analogous to conjunction; this function is performed as if "in substitution" by the full stop separating sentences.

In more than one case, a multi-sentence text such as a paragraph could be transposed into a single complex sentence, and then the criteria for sentence syntactic cohesion could be undoubtedly applied to it. Let the following text be an illustration of this possibility:

Let's assume that somebody hears or reads the wording: "If Caesar had not fallen victim to Brutus' conspiracy, he would have pronounced himself the absolute emperor of Rome." Even if this somebody does not know who Caesar was or who Brutus was, and even if does not understand the term "absolute emperor," then knowing the rules governing the English language, they can infer that the wording states that this Caesar fell victim to the conspiracy of this Brutus. They may also infer that as a consequence, Caesar did not manage to pronounce himself this absolute emperor of Rome.

Simple stylistic tricks allow us to transform the above paragraph into a single-sentence text, which is as follows:

If somebody hears or reads the wording: "If Caesar had not fallen victim to Brutus' conspiracy, he would have pronounced himself the absolute emperor of Rome," then even if they do not know who Caesar was or who Brutus was, and even if they do not understand the term "absolute emperor" — knowing the rules governing the English language, they can infer that the wording states that this Caesar fell victim to the conspiracy of this Brutus and as a consequence, Caesar did not manage to pronounce himself this absolute emperor of Rome.[6]

[6]The example drawn from O. Wojtasiewicz (1971).

4.4. In such cases as the above, the problem of the difference between sentence cohesion and text cohesion disappears. However, it is worth formulating a further-reaching generalization that would allow us to extend the notion of cohesion to texts that cannot be transposed into one (complex) sentence. The problem can be reduced to whether formal structures, such as trees in table 2, can be interpreted as text structures. These trees have the property that elements of the same-level-splits are in asymmetric, intransitive and irreflexive relation: functor — argument. This additional property is a problem here, since the fact that a text structure can always be represented as a tree is obvious and commonly used in text division, for example, in a table of contents. However, a question arises if such a dendritic structure that characterizes all partial orders, and hence the order of thought-development in the text, will also encompass the equivalent (with the same formal properties) of the specific relation functor — argument.

A hypothesis that the searched equivalent of the syntactic relation functor — argument is the textual relation question — answer is worth consideration. It does not mean that any question that comes into play needs to be explicitly formulated in the text. Questions are most often inferred, but it is not difficult to reconstruct them in such texts that are intuitively recognizable as coherent. Provided that we can present a text structure in the form of a tree in which, except for the characteristic partially ordering relation with an upper bound, there is this "cross" asymmetric, intransitive and irreflexive relation, then the methods of analysing cohesion developed for syntactic description can be transposed into a text structure description.

Since this hypothesis is noted only as a subject for further study, only a certain illustration will be provided together with an assumption that the compared types of structures (syntactic and textual) are interpretations of certain abstract algebra, i.e. the theory of lattice.

Let the first paragraph of Pragmatic Logic by Ajdukiewicz be an example illustrating the above hypothesis. The text division, sentence numbers and questions interpolated in capital letters are mine; the rest is a translation of the text by Ajdukiewicz (1974: 1).

0. WHAT IS THE TASK OF SCHOOL?

1. The task of the school is not only to convey to the pupils information in various fields,

2. but also to develop in them the ability of correctly carrying out cognitive operations. (...)

2.0. IN WHAT WAY CAN ABILITIES TO CORRECTLY PERFORM COGNITIVE OPERATIONS BE TRAINED?

2.1. But to be able to train his pupils in a correct performance of cognitive operations the teacher must himself provide the standard of correct thinking

2.2. This, however, does not suffice: the pupils must carry out the operations themselves, and the teacher must see to it that they do that properly.

2.2.0. IN WHAT WAY SHOULD THE TEACHER CARE ABOUT A CORRECT PERFORMANCE OF THE STUDENTS' COGNITIVE OPERATIONS?

2.2.1. Hence if a pupil, when proving a theorem, or explaining a phenomenon, or defining a concept, commits an error, the teacher must draw his attention to the fact.

2.2.1.0. WHAT DOES CORRECTING ERRORS CONSIST IN?

2.2.1.1. But it does not suffice if he tells the pupil that he committed an error;

2.2.1.2. he also ought to point out where the error is

2.2.1.3. and in what it consists.

2.2.1.1.0. WHAT ABILITY SUFFICES TO NOTICE THE ERROR?

2.2.1.1.1. The practical ability to think correctly, which every teacher should have, will alone suffice for him to notice that the pupil made an error,

2.2.1.2.0. WHAT ABILITY SUFFICES TO INDICATE WHERE THE ERROR OCCURRED?

2.2.1.2.1. it will probably also suffice for him to explain to the pupil where the error is.

2.2.1.3.0. WHAT IS NECESSARY TO BE ABLE TO TELL WHAT THE NATURE OF THE ERROR IS?

2.2.1.3.1. But it need not necessarily suffice for him to be able to tell in what that error consists and what its nature is. To be able to do so he must know those concepts and terms which make it possible to discuss cognitive operations and their types, properties, etc.

2.2.1.3.2. He also must have a theoretical knowledge of the conditions which the various cognitive operations must satisfy in order to be correct.

As can be seen, it turned out to be possible, within the analyzed text, to convey the split analogous to syntactic structure split. Although no split starting from (or ending in) an element with the number 0 (i.e. a "functor" element) occurred, it is not difficult to observe such elements in many other texts whose authors start with a question to raise additional problems which are the equivalents of functor-forming functors. Also, the above text can be

represented by a structure with a split in zero if the questions are formulated "from below," that is starting with the most detailed statements, for example:

0. SHOULD THE TEACHER KNOW THE CONCEPTS AND TERMS TO DISCUSS COGNITIVE OPERATIONS?

0.0. WHAT ARE CONCEPTS AND TERMS ALLOWING TO DISCUSS COGNITIVE OPERATIONS FOR?

0.1. They are for indicating in what the error committed by the pupil consists.

Thus, the analogy to syntactic description encompasses the multitude of acceptable structures, which correspond to different interpretations of a given text. Let's consider one more possible variant. The starting point in the analysed text could be the question:

0. WHAT IS THE TASK OF SCHOOL BESIDES CONVEYING INFORMATION?

1. To develop the ability of correctly carrying out cognitive operations.

In such an approach, question 0 would have one argument, which would express the fact that the person posing the question does not know only one of two things to be known. The choice of one particular structure interpreting the text depends on what knowledge about the described reality and the receiver of the text is attributed to the author (speaker), as well as on the intentions attributed to the author. This interpretation consisting in assigning the text with a particular structure can be expressed to a great extent by means of vocal interpretation, that is on the phonological level.

4.5. The proposed method of text analysis, that is its structure, seems not to be meaningless for such text operations (or text transformations) as a summary or indexing. Having split the text in the form of a tree (with a zero element on each level), it is possible to sum up the text to any degree of detail measured by specifying the last of the considered levels of split. Thus, for example, the most detailed summary of the analyzed paragraph will leave out the information of the last level, marked there by numbers with five positions. Yet, one can approach a summary in a different way, focusing on details rather than general ideas, and arriving at a different type of a summary, namely, by including mostly data of the lowest level (which will be, however, at the top of the structure if just below the tree's upper bound, as a zero point, will be posed a question as detailed as in the example with 0.0 and 0.1). Indeed, either one or the other approach is employed in practice, for example, when the summary gives only the starting assumptions, or only the final conclusions. However, both approaches can be merged, and then the middle levels of the tree will be left out.

Indexing is analogous: the depth of indexing can be characterized by specifying split levels, which determine indexing terms.[7]

These various suggestions resulting from the method of text analysis used here, which could be called a quasi-syntactic method on the basis of formal analogy to semantic description, seem to encourage the use of categorial grammar as a theoretical basis for text analysis. The present article employed categorial grammar to approach one of the most complex syntactic problems — reported speech, which involved designing methods of syntactic description in detail. Formal, that is highly abstract, characteristics of these methods allowed us to notice even further applications for the methods, which were presented here as a material and an inspiration for further research.

Bibliography

1. Ajdukiewicz, Kazimierz (1935) "Die syntaktische Konnexität." Studia Philosophica I: 1—27. In English: Ajdukiewicz 1967c.

2. Ajdukiewicz, Kazimierz (1967a) "Intensional expressions." Studia Logica 20: 63—86.

3. Ajdukiewicz, Kazimierz (1967b) "Proposition as the connotation of sentence." Studia Logica 20: 87-98.

4. Ajdukiewicz, Kazimierz (1967c) "Syntactic Connexion." In Polish logic 1920—1939, Storrs McCall (ed.), 207—231. Oxford: Oxford University Press.

5. Ajdukiewicz, Kazimierz (1974) Pragmatic Logic. Synthese Library 62. Dordrecht — Boston: Reidel.

6. Bar-Hillel, Yehoshua (1965) "A Quasi-arithmetical Notation for Syntactic Description." In Language and Information. Jerusalem: Addison-Wesley.

7. Bar-Hillel, Yehoshua (1970) "Decision procedures for structure in natural language." In Aspects of Language. Jerusalem — Amsterdam: Magnes and Morth-Holland.

[7]The issues of a summary, indexing and semantic organization of the text (also called "cohesion") were among my interests in the articles published in *Studia semiotyczne*: Marciszewski 1970, 1972, 1974, 1975. Certain remarks on semantic aspects of reported speech were included in Marciszewski 1971.

8. Church, Alonzo (1956) Introduction to Mathematical Logic. Vol. 1. Princeton: Princeton University Press.

9. Cresswell, Maxwell John (1973) Logics and Languages. London: Methuen.

10. Frege, Gottlob (1892) "Über Sinn und Bedeutung." Zeitschrift für Philosophie und philosophische Kritik 100: 25—50.

11. Geach, Peter Thomas (1970) "A program for syntax." Synthese 22: 3—17.

12. Jeffrey, Richard C. (1965) The Logic of Decision, New York: McGraw-Hill.

13. Kalinowski, Georges (1973) "Norms and Logic." The American Journal of Jurisprudence 18: 165—197.

14. Levelt, Willem J. M. (1974) Formal Grammars in Linguistics and Psycholinguistics. Vol. II. The Hague: Mouton.

15. Marciszewski, Witold (1970) "Sposoby streszczania i odmiany streszczeń." Studia Semiotyczne 1: 151—167.

16. Marciszewski, Witold (1971) "O funkcjach semantycznych mowy zależnej." In Semiotyka polska 1894-1969, Jerzy Pelc (ed.), Warszawa: PWN. In English: Pelc, Jerzy (ed.) (1979) Semiotics in Poland, 1894-1969. Dordrecht, Holland: Reidel.

17. Marciszewski, Witold (1972) "Formalna charakterystyka dziedziny rozważań jako podstawa indeksu rzeczowego." Studia Semiotyczne 3: 85—101.

18. Marciszewski, Witold (1974) "Analiza semantyczna pytań jako podstawa reguł heurystycznych." Studia Semiotyczne 5: 133—146.

19. Marciszewski, Witold (1975) "Organizacja semantyczna tekstu." Studia Semiotyczne 6: 105—124.

20. Martin, Richard Milton (1979) "On Prepositional Protolinguistics." In Konstruktionen versus Positionen: Beitrage zur Diskussion um die konstruktive Wissenschaftstheorie, Band II. Festschrift for Paul Lorenzen, Lorenz, Kuno (ed.), 184—208. Berlin: de Gruyter.

21. Szober, Stanisław (1953) Gramatyka języka polskiego. Warszawa: Nasza Księgarnia.

22. Wojtasiewicz, Olgierd Adrian (1971) "O pewnej interpretacji pojęcia spójności tekstu." In O spójności tekstu Maria Renata Mayenowa (ed.). Wrocław: Ossolineum.

23. Wojtasiewicz, Olgierd Adrian (1972) "Formalna i semantyczna analiza polskich spójników przyzdaniowych i międzyzdaniowych oraz wyrazów pokrewnych." Studia semiotyczne 3: 109—144.