Marek Tokarz

SEMANTICS WITHOUT THE CONCEPT OF DENOTATION

Originally published as "Semantyka bez pojęcia denotacji," Studia Semiotyczne 14–15 (1986), 133–146. Translated by Maja Wolsan.

This article is composed of two distinct parts: the syntactical part (I and II) and the semantic part (III). Their topics are quite far from each other, but they are combined into a single article because there exists a formal construction of language **L** (II) common for both these parts. The syntactic part is an attempt at taking a new approach to the syntactic function of common nouns and indicative pronouns. The analysis will lead to the said formal construction in section II, broadly referring to one of the works by A. Nowaczyk (1971). In the final part of the text (III), the semantic system for the formal language has been presented. It differs from other semantics in that here the problem of interpretation of this language comes down to indicating true statements among particularly simple expressions, referred to as atomic sentences, while not requiring the specification of object references for names of any type — neither individual nor general.

I. THE SYNTACTIC ROLE OF GENERAL NAMES

In this article, under 'general names' we understand all names that are not individual, while under 'individual names' we understand those which correspond to proper nouns, such as: *Aristotle*, *Vistula*, *Moon* and those that are created by indicating one object as their reference, e.g. *this man here*, *that running dog*, etc. (these expressions are an attempt to direct the reader's thoughts in a more or less relevant direction; they are not exact definitions and should not be treated as such). The logical division along the line 'general names' — 'individual names' is somewhat similar to the division into 'common nouns' and 'proper nouns' existing in grammar, although it is not totally equivalent. The set of general names will be marked as **G**, and the set of individual (singular) names as **J**.

Usually we identify **G** and **J** as both semantically different (**J** refers to objects, **G** — to a class of objects) and syntactically different (**G** can be used as both the subject of a sentence and the predicative nominal of sentences like *x is y*, **J** — only as the subject). We can easily explain the semantic basis of the syntactic difference. The following sentences, which are isomorphic in terms of syntax:

(a) *Aristotle is a philosopher*,
(b) *The dog is a mammal*,
(c) *The Morning Star is the Evening Star*

are interpreted as describing different kinds of relations. Sentence (a) corresponds to x $\in$ y, sentence (b) to $x \subseteq y$, while (c) to $x = y$.

We have already mentioned the sentence pattern *x is y*. Further on, we will need a more general concept of pattern, more specifically: SENTENCE PATTERN. By sentence pattern, we shall understand an expression containing a VARIABLE (or variables) that represents an entire class of sentences of a similar structure. These sentences are formed by replacing the variables in the pattern with expressions belonging to a specified set, called the SCOPE OF THE VARIABLE. Let us stress that the scope of a variable shall be understood herein as a set of expressions, i.e. as something of language origin; thus, it is a syntactic, not semantic term, although in practice the METHOD OF DEFINING this set may be semantic. For example, in the pattern $x + 5 = y$ the scope of both variables is NOT a set of natural numbers but rather a set of DIGITS, i.e. signs of numbers. In the pattern *x believes that Earth is a sphere* the scope of *x* is a set of personal proper nouns, and in the pattern *John believes that x*, the scope of *x* is a set of declarative sentences. In the nominal sentence pattern (in the sense adopted in logic) *x is y* the scope of *x* is $\mathbf{J} \cup \mathbf{G}$, and the scope of *y* is **G**.

The fact that a pattern (an expression containing variables which may be replaced, the so called FREE VARIABLES) must implicitly provide an instruction on the scope of the variables that it contains should not be questionable, if we analyse the example (+) *x believes that y*. If no scope was established, (+) would produce fully grammatical sentences, such as *John believes that Earth is a sphere*, but it would produce ungrammatical sentences as well, for example *Earth is a sphere believes that John*.

A sentence created according to pattern **A** by replacing the variables contained in **A** with expressions from the scope(s) of these variables will be called the REALISATION of A. Another type of sentences based on patterns are some particular sentences created by QUANTIFICATION. They establish (or define) the frequency of true realisations in the set of all realisations of a given pattern.

Let us sum up what has been said so far. The characteristic features of sentence patterns are that: (a) they represent a certain class of sentences; (b) the class is created by replacing the variable contained in the pattern with expressions from the scope of the variable; (c) the scope is defined in advance for every pattern; and (d) replacing a variable by an expression from beyond its scope leads to ungrammatical expressions, i.e. such that it is pointless to ask about their truth value; (e) sentences based on patterns can be also formed by quantification.

Let us consider what plays the role of variables in expressions (sentence forms) of natural language (naturally, they are not letter variables, such as x, y, z, ...; p_1, p_2, ..., etc.). It is usually believed that some types of pronouns function as these variables, e.g. *someone*, *something*, *this*, *that*, etc. Let us take a closer look at this statement. We might suppose that the 'variability' of the indicative pronoun *this*, for example, results from the actual variability of its meaning. If someone is turning around with his arm stretched out and the index finger pointing straight ahead, while saying *this!*, then the denotation of *this* is constantly changing, depending on the object which is currently being pointed at.

This solution is therefore related to a special, colloquial definition of the term 'variable', originating in physics, or rather from the stage in the development of mathematics in which theories were still indistinguishable from their practical applications. In the colloquial use, the word 'variable' is, namely, understood as 'something that can change in terms of quantity' or 'something that we can change in terms of quantity and possibly observe the quantitative effects of this change'. A variable is thus something of an independent variable. This point of view is not accepted in present-day logic. As a matter of fact, it gives rise to paradoxes pointed out already by Frege. Its persisting existence in the logical theory of syntax seems to be a copy of the solutions of traditional grammar. To back up this thesis, we can refer to a passage from the book *Gramatyka języka polskiego* [*Grammar of the Polish language*] by Stanisław Szober (1953):

The individual content contained in proper nouns makes them similar in terms of semantic value to pronouns, as these also always have individual content. The difference is that the content related to a noun is constant, while the individual content of a pronoun changes, as is commonly known, depending on the circumstances in which we use it. Consequently, pronouns have an unlimited scope of usage, while in proper nouns the scope is strictly related to the set content.

It seems unquestionable that in some contexts, some relative and indefinite pronouns (e.g. *any*, *someone*, *something*, *every*) play the role of variables and at the same time of operators binding them (quantifiers), as it is in the statement *Everybody loves somebody.* A very interesting and subtle analysis of such contexts was presented by A. Nowaczyk in his work *Zaimki zamiast zmiennych i operatorów* ([*Pronouns in the place of variables and operators*], Nowaczyk 1971). Many ideas for the present article were taken from this work, as e.g. introducing internal negation, which cannot be easily eliminated from some expressions. In our terminology, the symbol corresponding to the word *is* will be E, therefore 'internal negation' will be marked as $\bar{E}$, read as 'is not'; $\bar{E}$ corresponds to the term *e$\bar{s}$t* used by Nowaczyk. However, our task in the present part of this article is completely different from the above. We are namely trying to convince the reader of two theses: (I) that the INDICATIVE pronouns are constant-like rather than variable-like, and (II) that the grammatical function of variables in many statements of natural language, and most probably in a vast majority of sentences, is fulfilled by general names.

Let us consider the statement *człowiek jest ssakiem* [*man is a mammal*]. It does not say anything specific about anything, as it is not indicated whether it is about all human beings or only some of them, or maybe a specific human being. This example has been selected on purpose, to confuse two meanings of the noun *man* — the one referring to class and the individual one. It might give the impression that the above statement is a sentence, and one accurately describing the reality at that. If someone claims that the statement *man is a mammal* is simply a true sentence, he says that because he unconsciously identifies this statement with the sentence *every man is a mammal.* We can prove that this identification is unjustified by quoting an isomorphic (structurally identical) expression *człowiek jest blondynem* [*man is a blond*] which clearly requires an 'interpretation': either 'every man is a blond' or 'a man is a blond', or 'this man here is a blond' (when 'the man here' means e.g. John Smith, just indicated by the speaker), or 'every second man is a blond', or yet something else. The role of the word *man* in this example [*man is a blond*] is obvious. It marks the place that can be filled with various individual names. These names, however, only include names of people (the expression *sprawiedliwość jest blondynem* [*justice is a blond*] would be ungrammatical), therefore we have to assume that the scope of the variable is specified. The term *człowiek* is also subject to quantification — we can say *każdy człowiek* [*every man*], *pewien człowiek* [*a certain man*]. . . , etc. It is clear, then, that the term plays the role of a variable in the sense described above: it has a scope, it can be replaced by individual names from this scope and lends itself to quantification. These characteristics distinguish general names from individual names in a more fundamental way than others, usually incidental characteristics quoted as distinctive features.

What was said above refers only to those names which are used in the position of a subject (here: *man*), and not general names used as predicatives, i.e. not names such as *mammal* or *blond*. Although the phrases *is a mammal*, *is a blond* can be broken down using grammatical methods, they cannot be broken down in logical terms. From the logical perspective, both these phrases could look like this: *ssakuje* like in *człowiek ssakuje* [*man is mammaling*] and *blondynuje* like in *człowiek blondynuje* [*man is blonding*]. We have the right to adopt this arbitrary solution, as the aims of logical analysis are different than the aims of grammar, and while the aim does not justify the means, it certainly defines them.

On the other hand, in the sentence *ten człowiek jest znanym chirurgiem* [*this man is a famous surgeon*], probably neither the pronoun *this*, nor the phrase *this man* can be treated as marking the grammatical position of the whole class of acceptable replacements. We perceive the above statement as SENTENTIAL, not as a PATTERN of possible sentences of a certain shape. It shows much more resemblance to sentences such as *Ryszard Wójcicki jest znanym metodologiem* [*Ryszard Wójcicki is a famous methodologist*], *Wrocław jest dużym miastem* [*Wrocław is a large city*] than e.g. to the statement *człowiek jest znanym chirurgiem* [*man is a famous surgeon*]. The structure of the latter is not that of a sentence, but of a sentence form. In a half-formal language, it would correspond to the pattern *x is a famous surgeon*, in which the scope of *x* would be defined as a set of personal proper nouns.

It is also clear that expressions such as *ten człowiek* [*this man*], as opposed to the phrase *łysy człowiek* [*bald man*] do not lend themselves to quantification. We can say *każdy łysy człowiek* [*every bald man*], but the strings of words *każdy ten człowiek* [*every this man*] and *ten każdy człowiek* [*this every man*] would be ungrammatical in all possible contexts.

The above analysis has revealed the role of the indicative pronoun, which, when placed before a general name (a common noun) creates a phrase which syntactically corresponds to individual names (proper nouns). The seeming variability stems from pragmatic aspects: in real life, the actual content of the word *this* depends on what is currently being pointed at, i.e. on circumstances. The said pronoun is thus a so called indexical expression. However, it shares its incidentality with a vast majority of natural language expressions. Therefore, there is no reason to attach any special significance to it, at least no greater significance than to other typical incidental statements (*it is raining*, *I'll be back in five minutes*, *today is Friday*, *tomorrow is my birthday*, etc.).

It is clear that, for instance, in the sentence *Ten człowiek jest sumienny, a ten jest nieodpowiedzialny* [This man is diligent, and this one is irresponsible], the first pronoun *ten* [*this*] corresponds to a different person than the one indicated by the second pronoun. In formal language models, seeming variability is avoided by, for instance, attaching indexes to repeating indicative pronouns. In its initial formalised form, the above example would have the following structure: *Ten*$_1$ *człowiek jest sumienny, a ten*$_2$ *jest nieodpowiedzialny.* In the written version of Polish, the actual realisation of the latter structure could be as follows: *Ten pierwszy człowiek jest sumienny, a ten drugi jest nieodpowiedzialny* [The first man is diligent, and the other one is irresponsible]. Therefore, we can sometimes say, without any contradiction: *Tamten człowiek jest szatynem i równocześnie tamten człowiek jest łysy* [That man has dark hair and, at the same time, that man is bald], namely when this statement is one of the possible realisation of the deeper structure *Tamten*$_1$ *człowiek jest szatynem i równocześnie tamten*$_2$ *człowiek jest łysy.*

Eventually, we must assume that the role of an indicative pronoun, making a general name an individual one, is to transform sentence forms (patterns) with a common noun as a free variable by replacing the variable by a certain individual name (in the form: an indicative pronoun + a general name).

The aim of the first two sections of this article is, as has been said, to make a general analysis of purely syntactic problems related to general names. Now, we intend to build simple formal language employing this type of names in — as it seems — a way syntactically typical of them. The simplest languages of this type, $\mathbf{L}_1$ and $\mathbf{L}_2$, do not contain proper nouns; $\mathbf{L}_1$ has no individual names at all. $\mathbf{L}_1$corresponds to the language of syllogism, based on general sentences as primary sentences. The role of patterns (formulae which are not sentences) is played there by the expressions *S is P* and *S is not P*, which in our construction take a slightly different form.

Our further discussion will be relatively highly formalised. It will include the standard logic and set-theoretic notation. Readers who are unsure whether they understand a certain symbol correctly should refer to a handbook of formal logic (e.g. Bańczerowski, Pogonowski, Zgółka 1982: 98—100).

II. FORMAL LANGUAGE CONSTRUCTION

The vocabulary of $\mathbf{L}_1$contains $n + 4$ symbols:

$$g_1,\ g_2,\ \ldots\ g_n,\ E,\ \bar{E},\ K,\ \sim.$$

We shall be using the following abbreviation:

$$\mathbf{G} = \{g_1, \ldots,\ g_n\}.$$

G shall be called a set of GENERAL NAMES. If $x \in$ **G,** then Kx, read as 'every x' is called a QUANTIFIER PHRASE (of language $\mathbf{L}_1$). The set of all quantifier phrases of language $\mathbf{L}_1$shall be marked as $\mathbf{QP}_1$. General names and quantifier phrases are jointly called noun phrases (of language $\mathbf{L}_1$); the set of noun phrases of language $\mathbf{L}_1$shall be marked as $\mathbf{NP}_1$, i.e. $\mathbf{NP}_1 = \mathbf{G} \cup \mathbf{QP}_1$. The set of expressions taking the form Ex and $\bar{E}x$, where $x \in$ **G**, is called the set of VERB PHRASES of language $\mathbf{L}_1$ and shall be marked as $\mathbf{VP}_1$. Ex is read as 'is an x'; $\bar{E}x$ is read as 'is not an x'. Now we shall define (by induction) the terms PATTERN, SENTENCE and FORMULA (of language $\mathbf{L}_1$):

If $x \in$ **G** and $y \in \mathbf{VP}_1$, then xy is a pattern. If $x \in \mathbf{QP}_1$ and $y \in \mathbf{VP}_1$, then xy is a sentence. If **A** is a sentence, then $\sim A$ (read as 'it is not true that A') is also a sentence. Sentences and patterns are jointly called FORMULAE; a set of formulae of language $\mathbf{L}_1$shall be marked as $\mathbf{For}_1$. Elements of the set $\mathbf{For}_1 \cup \mathbf{NP}_1 \cup \mathbf{VP}_1 \cup \{E,\ \bar{E},\ K,\ \sim\}$ are called CORRECTLY BUILT EXPRESSIONS.

EXAMPLES. (a) PATTERNS: g_iEg_j, g_iEg_i, $g_i\bar{E}g_j$, $g_i\bar{E}g_i$, where $i, j \leqslant n$. (b) SENTENCES: Kg_iEg_j, $Kg_i\bar{E}g_j$, $Kg_i\bar{E}g_i$, $\sim\sim Kg_iEg_j$, etc. (c) INCORRECTLY BUILT EXPRESSIONS: g_iE, KEg_i, $Kg_ig_jEg_i$, $\sim K\bar{E}$, $\sim g_i\bar{E}g_j$, $K{\sim}g_iEg_j$, etc.

Categorical sentences in the standard syllogistics correspond to the following expressions of language $\mathbf{L}_1$:

sentence xay ('all x are y'/ 'every x is y') corresponds to sentence $KxEy$
sentence xey ('no x is y') corresponds to sentence $Kx\bar{E}y$

sentence xiy ('some x are y') corresponds to sentence $\sim Kx\bar{E}y$
sentence xoy ('some x are not y') corresponds to sentence $\sim KxEy$.

EXAMPLE OF A REALISATION OF $\mathbf{L}_1$

The vocabulary in our example shall be composed of the following expressions:

G =	{*filozof, dramaturg*}	
	[*philosopher, playwright*]	(general names)
	jest, nie jest [*is, is not*]	(copulas)
	każdy [*all/every*]	(quantifier)
	nieprawda, że [*it is not true that*]	(sentential operator, so called negation)

(In the examples quoted below, words are inflected according to the rules of Polish grammar.)[1]

EXAMPLES. (a) PATTERNS: *filozof jest dramaturgiem* [a philosopher is a playwright]; *filozof jest filozofem* [a philosopher is a philosopher]. (b) SENTENCES: *każdy dramaturg jest filozofem* [every playwright is a philosopher]; *nieprawda, że każdy dramaturg nie jest dramaturgiem* [it is not true that every playwright is not a playwright]. (c) INCORRECTLY BUILT EXPRESSIONS: *nieprawda, że dramaturg nie jest filozofem* [it is not true that a playwright is not a philosopher]; *każdy jest filozofem* [every is a philosopher], *nieprawda, że filozof jest* [it is not true that a philosopher is].

INFLECTIONAL FORM OF $\mathbf{L}_1$

In order to keep the expressions of language $\mathbf{L}_1$ in agreement with the rules of Polish grammar, we had to change the literal form of some formulae. Now we will show that it is possible to build a formal language in which no such changes are necessary. It will be called the INFLECTIONAL FORM OF $\mathbf{L}_1$. Further in this text, however, we shall not build any inflectional forms of analysed languages, as they would have to be immensely complex, and

[1] Translator's note: please note that Polish is an inflective language. To make the examples more clear for non-Polish speakers, corresponding phrases in English have been provided in square brackets, however, the English phrases should not be viewed as examples for the purpose of this article.

this complexity would not be justified by its value for scientific purposes. Let us just state the fact that a construction of this kind is indeed possible.

The vocabulary of the language in question has $2n+5$ symbols:

$$g_1^1,\ g_1^2,\ g_2^1,\ \ g_2^2,\ \ldots,\ g_n^1,\ g_n^2,$$
$$E,\ \bar{E},\ K^1, K^2,\ \sim$$

Expressions g_i^1, $g_i^2$are called INFLECTED VARIANTS of the general name g_i; K^1 and $K^2$are variants of the quantifier. We introduce the following abbreviations:

$$\mathbf{G}_1 = \{g_i^1\colon 1 \leqslant\ i \leqslant\ n\};\ \mathbf{G}_2 = \{g_i^2\colon 1 \leqslant\ i \leqslant\ n\}.$$

In this language, the inductive definition of a set of all formulae (including sentences) is as follows:

1. If $x \in\ \mathbf{G}_1$ and $y \in\ \mathbf{G}_2$**,** then the expressions xEy, yEx, $x\bar{E}$y, $y\bar{E}x$ are formulae.

2. If $x \in\ \mathbf{G}_1$ and $y \in\ \mathbf{G}_2$**,** then the expressions K^1xEy, yEK^1x, $K^2x\bar{E}$y, $y\bar{E}K^2x$ are sentences.

3. If A is a sentence, then $\sim A$ is also a sentence.

EXAMPLE OF A REALISATION OF AN INFLECTIONAL FORM OF $\mathbf{L}_1$

The vocabulary in our example of realisation of an inflectional form of $\mathbf{L}_1$ shall be composed of the following expressions:

filozof^1, *filozof*^{2}, *dramaturg*^{1}, *dramaturg*^{2}
jest, *nie jest*
każdy^1, *każdy*^{2}
nieprawda, że

The indexes mark various versions of the above words, as required by Polish grammar:

filozof^1 $=$ ⌜filozof⌝

folozof^2 = ⌜filozofem⌝
dramaturg^1 = ⌜dramaturg⌝
dramaturg^2 = ⌜dramaturgiem⌝
każdy^1 = ⌜każdy⌝ [all/every]
każdy^2 = ⌜żaden⌝ [no].

EXAMPLES. (a) FORMULAE: *filozof jest dramaturgiem*; *dramaturgiem jest filozof* [a philosopher is a playwright[2]]; *każdy dramaturg jest dramaturgiem* [every playwright is a playwright]; *nieprawda, że nieprawda, że żaden filozof nie jest dramaturgiem*; *nieprawda, że nieprawda, że dramaturgiem nie jest żaden filozof* [it is not true that it is not true that no philosopher is a playwright]. (b) INCORRECT EXPRESSIONS: *filozof jest filozof* [a philosopher$_{NOM}$ is a philosopher$_{NOM}$]; *każdy filozof nie jest dramaturgiem* [every philosopher is not a playwright]; *nieprawda, że dramaturgiem nie jest filozof* [it is not true that a philosopher$_{NOM}$ is not a playwright$_{INSTR}$]; *nieprawda, że filozofem jest każdy* [it is not true that every is a philosopher].

ADDING INDICATIVE PRONOUNS

The vocabulary of $\mathbf{L}_2$is an extended version of the vocabulary of $\mathbf{L}_1$, created by adding the following symbols:

$t_1, t_2, t_3, \ldots$ (indicative pronouns in unlimited quantity)
$\wedge, \vee, \rightarrow$ (sentential connectives)

The set of all indicative pronouns shall be marked as $\mathbf{T}$, i.e. $\mathbf{T} = \{t_1, t_2, \ldots\}$. The expression xy, where $x \in \mathbf{T}$ and $y \in \mathbf{G}$ shall be called an INDIVIDUAL NAME; a set of all individual names shall be marked as $\mathbf{J}$. Just as for $\mathbf{L}_1$, we will provide the definitions of a QUANTIFIER PHRASE, NOUN PHRASE and VERB PHRASE of language $\mathbf{L}_2$. Let as assume that:

$$\mathbf{QP}_2 = \mathbf{QP}_{1,} \mathbf{VP}_2 = \mathbf{VP}_1,\ \mathbf{NP}_2 = \mathbf{NP}_1 \cup \mathbf{J}.$$

[2] Translator's note: in Polish, both these phrases have the same meaning. The subject and the predicative can be placed on either side of the copula, yet the meaning remains the same because the inflection indicates which one is the subject and which one is the predicative (the predicative is in the instrumental case). Compare: *dramaturg jest filozofem* [a playwright is a philosopher].

The term pattern shall remain in its former meaning, but the definition of sentence will naturally change (as will the term formula). Just as earlier, the definitions of these terms shall be given in an inductive form:

If $x \in \mathbf{J}$ and $y \in \mathbf{VP}_2$, then xy is a sentence of $\mathbf{L}_2$.

If A is a sentence in $\mathbf{L}_1$, then A is also a sentence in $\mathbf{L}_2$. If A and B are sentences of $\mathbf{L}_1$, then the expressions $\sim A$, $(A \vee B)$, $(A \wedge B)$, $(A \rightarrow B)$ are also sentences of $\mathbf{L}_2$. The last three expressions should be read, respectively: 'A or B', 'A and B', 'if A then B'.

An example of the realisation of $\mathbf{L}_2$ can be a concrete language with the following vocabulary:

G = {*filozof*, *dramaturg*}
T = {*ten pierwszy*, *ten drugi*, *ten trzeci*, ...} [this first one, this second one, the third one, ...]
jest, *nie jest*
każdy
i, *lub*, *jeśli, ... to nieprawda, że* [*and*, *or*, *if ... then*, *not true that*]

If there is only one indicative pronoun (e.g. only *ten ósmy* [this eight one]) and, in addition, it occurs only in one place, then the index added to this pronoun (in this case the numeral *ósmy*) can be omitted, i.e. we can say *ten filozof jest dramaturgiem* [this philosopher is a playwright] instead of *ten ósmy filozof jest dramaturgiem.* As it was the case with the inflective version of $\mathbf{L}_1$, the quantifier in contexts with the word *nie jest* is rather read as *żaden* than *każdy*, i.e. e.g. the sequence of words *każdy filozof nie jest dramaturgiem* should rather be read as *żaden filozof nie jest dramaturgiem.*[3]

EXAMPLES. (a) SENTENCES: *ten pierwszy filozof jest dramaturgiem*; (*jeśli ten dramaturg nie jest filozofem, to żaden dramaturg nie jest filozofem*); (*ten pierwszy dramaturg nie jest folozofem i ten drugi dramaturg nie jest filozofem*); (*jeśli* (*ten pierwszy filozof jest filozofem i nieprawda, że każdy filozof jest filozofem*), *to ten pierwszy filozof jest dramaturgiem*). (b) INCORRECT EXPRESSIONS: *ten każdy filozof jest dramaturgiem*; *każdy ten drugi filozof jest filozofem*; *filozof jest tym dramaturgiem*; *filozof jest dramaturgiem lub filozof nie jest dramaturgiem*; etc.

[3]Translator's note: in cases such as this, double negation is required in Polish.

The very simple languages $\mathbf{L}_1$and $\mathbf{L}_2$ are initial stages of the construction of language **L**, which is our target. Apart from groups of terms, this language also includes proper nouns, i.e. individual names which are not formed by using indicative pronouns, e.g. *Sokrates*, *Eurypides*, and an internal general quantifier, expressed by words such as *jakikolwiek* [any] or *dany* [given]. Proper nouns make it possible to form sentences such as *Sokrates jest filozofem* [Socrates is a philosopher]; *Jeśli Eurypides jest filozofem, to nieprawda, że żaden dramaturg nie jest filozofem* [If Eurypides is a philospher, then it is not true that no playwright is a philosopher]. The internal quantifier allows for building sentences such as *Jeśli jakikolwiek filozof jest dramaturgiem, to ten filozof nie jest filozofem* [If any philosopher is a playwright, then this philosopher is not a philosopher].

The above example shows that introducing the quantifier *jakikolwiek* creates an additional problem — the possibility of an indicative pronoun appearing as an anaphora. This leads to the need to introduce a new group of terms to the language, as the possible occurrence of the symbols t_1, t_2, ... in a double role would create great technical difficulties in semantic analysis.

LANGUAGE **L**

The vocabulary of language **L** is composed of the following symbols:

$g_1, g_2, \ldots, g_n$ — general names
$a_1, a_2, \ldots, a_m$ — proper nouns
$t_1, t_2, t_3, \ldots$ — indicative pronouns
$o_1, o_2, o_3, \ldots$ — anaphoric pronouns
K; $J_1, J_2, J_3, \ldots$ — quantifiers
E, $\bar{E}$ — copulas
$\wedge, \vee, \rightarrow, \sim$ — connectives
(,) — brackets

We introduce the following abbreviations: $\mathbf{G} = \{g_1, g_2, \ldots, g_n\}$, $\mathbf{NJ} = \{a_1, a_2, \ldots, a_m\}$, $\mathbf{T} = \{t_1, t_2, t_3, \ldots\}$. Any finite string of symbols from the vocabulary, including the empty string $\varnothing$, will be called an expression (of language **L**). A concatenation (combination) of expressions x and y will be marked as xy. We say that expression x OCCURS in expression y or that it IS A PART OF y, marking it as $x \varepsilon y$, when there exist expressions v and w, of which at least one is not empty, such that $y = vxw$. If A, x, y are expressions, then $A[x\| y]$ means an expression created from A by replacing each instance of x with y; if x is not part of A, then $A[x\| y] = A$. Letters i, j,

k are variables across digits 1, 2, ..., n, letter l is a variable across digits 1, 2, ..., m, letters p, r, s are variables across the signs of all natural numbers digits 1, 2, 3, ... By **J** we mean a set of INDIVIDUAL NAMES, defined as **NJ** $\cup \{xy: x \in$ **T** and $y \in$ **G**$\}$; therefore J is a set of proper nouns and names taking the form $t_r g_i$.

The set of SENTENCES of language **L** is marked as **ZD** and defined as the smallest set of expressions fulfilling the following criteria:

1. For all l and all j, $a_l E\mathrm{g}_j \in$ **ZD** and $a_l \bar{E}\mathrm{g}_j \in$ **ZD**;

2. For all i, j, r, $t_r g_i E\mathrm{g}_j \in$ **ZD** and $t_r g_i \bar{E}\mathrm{g}_j \in$ **ZD**;

3. For all i, j, $Kg_i E\mathrm{g}_j \in$ **ZD** and $Kg_i \bar{E}\mathrm{g}_j \in$ **ZD**;

4. Let us assume that A, $B \in$ **ZD** and that no expression $J_s g_i$ is contained both in A and B. Then $(A \wedge B) \in$ **ZD,** $(A \vee B) \in$ **ZD,** $(A \rightarrow B)$ $\in$ **ZD,** $\sim A \in$ **ZD**;

5. Let us assume that $t_r g_i \ \varepsilon\ A$, $A \in$ **ZD**, and at the same time $J_s g_j \notin A$. Then $(J_s g_j E g_k \rightarrow A[t_r g_i \| o_s g_j]) \in$ **ZD** and $(J_s g_j \bar{E} g_k \rightarrow A[t_r g_i \| o_s g_j])$ $\in$ **ZD**;

6. Let us assume that $(B_1 \rightarrow B_2) \in$ **ZD**, $t_r g_i \ \varepsilon\ B_2$, $J_s g_j \notin (B_1 \rightarrow B_2)$. Further, let $A = J_s g_j E g_k$or $A = J_s g_j \bar{E} g_k$. Then $((A \wedge B_1) \rightarrow B_2[t_r g_i \| o_s g_j]) \in$ **ZD**, $((A \wedge B_1[t_r g_i \| o_s g_j]) \rightarrow B_2[t_r g_i \| o_s g_j]) \in$ **ZD**, $((B_1 \wedge A) \rightarrow B_2[t_r g_i \| o_s g_j]) \in$ **ZD**

In order to make the content of this definition more clear, let us quote four examples of expressions which are not sentences:

(a) $(J_1 g_1 E g_2 \rightarrow o_1 g_1 E g_3) \wedge (J_1 g_1 E g_2 \rightarrow o_1 g_1 E g_4)$ — in this expression, $J_1 g_1$ is repeated twice, and we can infer from the above definition that no $J_r g_i$ phrase can be repeated in a sentence;

(b) $(J_1 g_1 E g_2 \rightarrow (J_2 g_1 E g_3 \rightarrow o_2 g_1 E g_4))$ — no $o_1 g_1$ phrase in the consequent;

(c) $(J_1 g_1 E g_2 \rightarrow (J_2 g_1 E g_3 \rightarrow o_1 g_1 E g_4))$ — the consequent is not a sentence, which excludes the application of points 4 and 5 of the definition;

(d) $((o_1 g_1 E g_2 \wedge J_1 g_1 E g_3) \rightarrow o_1 g_1 E g_4)$ — let us assume that this expression is a sentence; it was not formed under point 4, as its antecedent is not a sentence, thus it would have to be based on rule 6°, with $A = J_1 g_1 E g_3$; but then $B_1 = o_1 g_1 E g_2$and for certain r, j, $B_2 = t_r g_j E g_4$; then, however $(B_1 \rightarrow B_2)$ is not a sentence, thus rule 6 also does not apply.

If a sentence does not contain the symbols J_2, J_3, …, o_2, o_3, …, t_2, t_3, …, then if it contains J_1 and o_1, we read them as *jakikolwiek* [any] and *ten* [this] respectively. The symbol J_i plays the role of a general quantifier where the symbol K cannot be used. The role of the 'internal' quantifier J_i is best explained with an example. The expression corresponding to sentences such as *Każdy palący mężczyzna jest zagrożony rakiem* [Every smoking man is threatened by cancer] is obviously not $K(xEy \rightarrow\ xEz)$, as this is not even a sentence, but $(J_1xEy \rightarrow\ o_1xEz)$. If in the latter expression we read x, y and z as *mężczyzna*, *palący* and *zagrożony rakiem* respectively, we will form the following sentence: *Jeśli jakikolwiek mężczyzna jest palący, to ten mężczyzna jest zagrożony rakiem* [If any man is a smoking man, then this man is threatened by cancer]. In the sentence $(J_1g_iEg_j \rightarrow\ o_rg_iEg_k)$ the individual content of the phrase $o_rg_i$is defined by the occurrence of the internal quantifier J_rg_i before it. The phrase o_rg_i occurring in this kind of context is sometimes called an ANAPHORA.

An example of a sentence can be the following expression $((J_1g_1Eg_2\wedge J_1g_2Eg_3) \rightarrow (o_1g_1Eg_3 \vee\ o_1g_2\bar{E}g_4))$, which we read according to the rules: 'if any g_1 is g_2 and any g_2 is g_3, then this g_1 is g_3 or this g_2 is not g_4.' This sentence is formed in the following way: according to rule 2°, $t_1g_1Eg_3$ and$t_1g_2\bar{E}g_4$ are sentences, but do not have a common element of the J_ix type, therefore, the alternative $(t_1g_1Eg_3 \vee\ t_1g_2\bar{E}g_4)$is a sentence under rule 4°; this alternative does not contain J_1g_2, thus under rule 5° we can replace $t_1g_2$with $o_1g_2$and place the phrase $J_1g_2Eg_3$ before this whole expression, thus forming the following sentence: $(J_1g_2Eg_3\rightarrow (t_1g_1Eg_3 \vee o_1g_2\bar{E}g_4))$; as this sentence does not contain the phrase J_1g_1, under rule 6° the following expression is also a sentence: $(J_1g_1Eg_2\wedge\ J_1g_2Eg_3) \rightarrow\ (o_1g_1Eg_3 \vee o_1g_2\bar{E}g_4))$; in point 6° of the definition we place $B_1= J_1g_2Eg_3$, $B_2= (t_1g_1Eg_3 \vee o_1g_2\bar{E}g_4)$, $t_rg_i\ = t_1g_1, J_sg_j\ = J_1g_1,\ A = J_1g_1Eg_2$.

If $x \in\ \mathbf{J}$, $y \in\ \mathbf{G}$, then the expression xEy is called an ATOMIC SENTENCE, i.e. atomic sentences have the form $a_iEg_i$or $t_rg_iEg_j$. The set of all atomic sentences of language L will bemarked as **AT**.

III. SEMANTICS

In the conception presented herein, atomic sentences are the only expressions directly 'connected' with the reality. They are the only ones that have a purely empirical content and logic cannot discuss their truth value. Some of them express truth about the world that they were created to describe, others do not tell the truth about this world, and thus are false. We start with the assumption that the interpretation of a language is given when some true sentences have been selected among its atomic sentences. Thus, SEMANTICS is a certain subset **S** of set **AT** of atomic sentences. The set **AT-S** is marked as **F**; elements of **F** are false atomic sentences. A given semantics **S** defines the set of all true sentences of language **L**, marked as **S***, in the following way (in the formulae below, 'iff' is used as an abbreviation of 'if and only if'):

A If $A \in AT$, then $A \in$ **S*** iff $A \in$ **S**;

B If $x \in$ **J**, then x$\bar{E}g_i \in$ **S*** iff $xEg_i \notin$ **S;**

C $Kg_iEg_j \in$ **S*** iff for any $x \in$ **J**, if $xEg_i \in$ **S**, then $xEg_j \in$ **S**;

$Kg_i\bar{E}g_j \in$ **S*** iff for any $x \in$ **J**, if $xEg_i \in$ **S**, then $xEg_j \notin$ **S**;

D Let us assume that A and B are sentences that do not contain a common occurrence of the phrase J_sg_i. Then $(A \wedge B) \in$ **S*** iff $A \in$ **S*** and $B \in$ **S***; $(A \vee B) \in$ **S*** iff $A \in$ **S*** or $B \in$ **S***; $(A \rightarrow B) \in$ **S*** iff $A \notin$ **S*** or $B \in$ **S***; $\sim A \in$ **S*** iff $A \notin$ **S***;

E Let us assume that $t_rg_i \; \varepsilon \; A \in$ **ZD** and $J_sg_j \notin A$. Then

$(J_sg_jEg_k \rightarrow A[t_rg_i \| \; o_sg_j]) \in$ **S*** iff for any $x \in$ **J**, if xEg_j **S** and $xEg_k \in$ **S**, then $A[t_rg_i \| \; x] \in$ **S***;

$(J_sg_j\bar{E}g_k \rightarrow A[t_rg_i \| \; o_sg_j]) \in$ **S*** iff for any $x \in$ **J**, if xEg_j **S** and $xEg_k \in$ **S**, then $A[t_rg_i \| \; x] \notin$ **S***;

F Let us assume that $(B_1 \rightarrow B_2) \in$ **ZD**, $t_rg_i \; \varepsilon \; B_2$, $J_sg_j \notin (B_1 \rightarrow B_2)$. Then

$((J_sg_jEg_k \wedge B_1) \rightarrow B_2[t_rg_i \| \; o_sg_j]) \in$ **S*** iff $((B_1 \wedge J_sg_jEg_k) \rightarrow B_2[t_rg_i \| \; o_sg_j]) \in$ **S*** iff for any $x \in$ **J**, if $xEg_j \in$ **S** and $xEg_k \in$ **S** and $B_1 \in$ **S***, then $B_2[t_rg_i \| \; x] \in$ **S***;

$((J_sg_j\bar{E}g_k \wedge B_1) \rightarrow B_2[t_rg_i \; \| \; o_sg_j]) \in$ **S*** iff $((B_1 \wedge J_sg_j\bar{E}g_k) \rightarrow B_2[t_rg_i \| \; o_sg_j]) \in$ **S*** iff for any $x \in$ **J**, if $xEg_j \in$ **S** and $xEg_k \notin$ **S** and $B_1 \in$ **S***, then $B_2[t_rg_i \| \; x] \in$ **S***;

$((J_sg_jEg_k \wedge B_1 \; [t_rg_i \| \; o_sg_j]) \rightarrow B_2(t_rg_i \| \; o_sg_j]) \in$ **S*** iff for any $x \in$ **J**, if $xEg_j \in$ **S** and $xEg_k \in$ **S**, then $(B_1 \rightarrow B_2) \; [t_rg_i \| \; x] \in$ **S***;

$((J_sg_j\bar{E}g_k \wedge B_1 \; [t_rg_i \| o_sg_j]) \rightarrow B_2(t_rg_i \| o_sg_j]) \in$ **S*** iff for any $x \in$ **J**, if $xEg_j \in$ **S** and $xEg_k \notin$ **S**, then $(B_1 \rightarrow B_2) \; [t_rg_i \| \; x] \in$ **S***.

The set **ZD** — **S*** is marked as **F***; **F*** is a set of false sentences of language **L** with interpretation (semantics) **S**. As a consequence of the adopted definitions ZD ≡ S* ∪ F*, according to which each sentence is either true of false. We could call this equivalence an ASSUMPTION OF BIVALENCE. There is, however, no reason not to consider also those semantic theories which do not include this assumption. One of the possible unorthodox ways would be as follows: SEMANTICS is any pair (**S**, **F**) where $\mathbf{S} \subseteq \mathbf{AT}$, $\mathbf{F} \supseteq \mathbf{AT}$ and $\mathbf{S} \cap \mathbf{F} = \varnothing$, without requiring that $\mathbf{S} \cup \mathbf{F} = \mathbf{AT}$. We would then have to modify the definition of a set of true sentences, and the new definition would have to provide for the construction of two disjunctive sets, **S*** and **F***. An intuitive basis for this assumption could be the observation that in real languages we can empirically establish the truth value of only some atomic sentences, while it would be impossible to do it empirically with some other sentences.

In concrete realisations of language **L** the possible semantics are not all equally applicable. When we use a language, it is important to us which atomic sentences are considered true; we distinguish one of the possible semantics as the 'right' one. For example, in a language containing the proper nouns *Sokrates* and *Eurypides* and the general names *filozof* and *dramaturg*, the right semantics would have to cover the sentences *Sokrates jest filozofem* [Socrates is a philosopher] and *Eurypides jest dramatugiem* [Eurypides is a playwright], as in our concrete world they are simply true, and it could not include sentences such as *Eurypides jest filozofem* [Eurypides is a philosopher]. If one of the possible semantics of language **L** is distinguished as the right one, the language is considered interpreted; formally: an INTERPRETED LANGUAGE is a pair (**L**, **S**), where **S** is a semantics for **L**.

Not all expressions of the set **S***, i.e. not all true sentences, are of equal interest to a logician. As opposed to a sociologist, physicist, historian, etc., a logician is interested mainly in those sentences of which the truth value is a non-variable of interpretation that is those which are true in any semantics. The existence of such sentences is the most fundamental characteristic distinguishing human languages from other communication systems. No reasoning would be possible without them.

The sentences which remain true regardless of the selected semantics are called TAUTOLOGIES; a set of all tautologies will be marked as **TAUT**. If the expressions A_1, A_2, $(A_1 \rightarrow A_2)$ are sentences, then we say that sentence A_2IS A CONSEQUENCE of sentence A_1, if $(A_1 \rightarrow A_2) \in \mathbf{TAUT}$. An example of a tautology of a given realisation of language **L** can be the following sentence: (*Jeśli każdy dramaturg jest filozofem, to* (*jeśli Eurypides jest dramaturgiem,* *to Euripides jest filozofem*)) [(If every playwright is a philosopher, then (if Eurypides is a playwright, then Euripides is a philosopher))]. Thus, the sentence *Każdy dramaturg jest filozofem* [Every playwright is a philosopher.] implies *Jeśli Eurypides jest dramaturgiem, to Euripides jest filozofem* [If Eurypides is a playwright, then Euripides is a philosopher].

The set **TAUT** is decidable, i.e. there exists an effective procedure (algorithm) which allows us to decide, in a finite number of steps, whether any selected sentence of language **L** is a tautology. It can be proved by reconstructing the set **ZD** in a set of formulae of language **M** of monadic predicate calculus and using the theorem that the set of laws of this calculus is decidable. The proof of decidability of **TAUT** will be only shortly outlined below, without uninteresting technicalities.

LOGIC OF MONADIC PREDICATES

Language **M** of this logic has a vocabulary composed of the following symbols:

$x_1^1, x_1^2, \ldots\ x_2^1, x_2^2, \ldots, x_n^1, x_n^1, \ldots$ — individual variables
$a_1, a_{2,\ldots,} a_m, b_1^1, b_1^2, \ldots, b_2^1, b_2^2, \ldots, b_n^1, b_n^2, \ldots$ — individual constants
$Q_1, Q_2, \ldots, Q_n$ — one-argument predicates
$\forall$ — quantifier
$\wedge, \vee, \rightarrow, \sim$ — connectives
(,) — brackets

The set of all variables shall be marked as **ZM**, the set of all constants as **ST**. The set of formulae of language **M**, marked as $\mathbf{FOR}_M$, is defined inductively: (i) if $x \in$ **ZM** $\cup$ **ST**, then $Q_i(x) \in \mathbf{FOR}_M$,(ii) if $A, B \in \mathbf{FOR}_M$, then $(A \wedge B)$, $(A \vee B)$, $(A \rightarrow B)$, $\sim A \in \mathbf{FOR}_M$, (iii) if $A \in \mathbf{FOR}_M$, then $\forall x_i^r A \in \mathbf{FOR}_M$. If A is a formula of language **M** and $x, y \in$ **ZM** $\cup$ **ST**, then $A[x\| y]$ means the result of replacement of every occurrence of x in formula A by y.

A MODEL for language **M** is any family $\mathbf{R} = \{U_1, U_2, \ldots, U_n\}$ of subsets of **ST**, such that $\cup\mathbf{R} = \mathbf{ST}$. An INTERPRETATION of language **M** in model **R** is any function I from the set **ZM** OVER **ST**, i.e. I is an interpretation when I: **ZM** $\rightarrow$ **ST** and every constant is an image of a certain variable. For some interpretations I and some formulae A, we will say that I FULFILS A, which we note as $I \models A$. The definition of the fulfilment is an inductive one:

I° $I \models Q_i(x_s^j)$ iff $I\ (x_j^s) \in U_i$; if $x \in$ **ST**, then $I \models Q_i(x)$ iff $x \in U_i$;

II° $I \models (A \wedge B)$ iff $I \models A$ and $I \models B$;

$I \models (A \vee B)$ iff $I \models A$ or $I \models B$;

$I \models (A \rightarrow B)$ iff, if $I \models A$, then $I \models B$;

$I \models \sim A$ iff it is not true that $I \models A$;

III° $I \models \forall x_i^s A$ iff for any $x \in \mathbf{ST}$, $I \models A[x_i^s \| x]$.

Formula A is a TAUTOLOGY OF MONADIC PREDICATE CALCULUS, if for any model $\mathbf{R}$ and for any interpretation I of language $\mathbf{M}$ in this model formula A is fulfilled by I. A set of monadic tautologies is marked as $\mathbf{TAUT}_M$. (The definitions of model and fulfilling, and consequently also of a tautology of language $\mathbf{M}$ are provided here in an untypical form, to make it easier to prove the theorems quoted below).

DECIDABILITY THEOREM

We shall now define a certain function f, which assigns a special formula of language $\mathbf{M}$ to every sentence of language $\mathbf{L}$:

a° $f(a_l E g_i) = Q_i(a_l)$; $f(a_l \bar{E} g_i) = \sim Q_i(a_l)$; $f(t_r g_i E g_j) = Q_j(b_i^r)$; $f(t_r g_i \bar{E} g_j) = \sim Q_j(b_i^r)$;

b° $f(K g_i E g_j) = \forall x_1^1 (Q_i(x_1^1) \rightarrow Q_j(x_1^1))$; $f(K g_i \bar{E} g_j) = \forall x_1^1 (Q_i(x_1^1) \rightarrow \sim Q_j(x_1^1))$;

c° If A, B are sentences that do not contain a common occurrence of the phrase $J_s g_i$, then $f(A \wedge B) = f(A) \wedge f(B)$; $f(A \vee B) = f(A) \vee f(B)$; $f(A \rightarrow B) = f(A) \rightarrow f(B)$; $f(\sim A) = \sim f(A)$;

d° If $t_r g_i \,\varepsilon\, A \in \mathbf{ZD}$ and $J_s g_j \not\varepsilon A$, then $f(J_s g_j E g_k \rightarrow A[t_r g_i \| o_s g_j]) = \forall x_j^s ((Q_j(x_j^s) \wedge Q_k(x_j^s)) \rightarrow f(A)[b_i^r \| x_j^s])$;

$f(J_s g_j \bar{E} g_k \rightarrow A[t_r g_i \| o_s g_j]) = \forall x_j^s ((Q_j(x_j^s) \wedge \sim Q_k(x_j^s)) \rightarrow f(A)[b_i^r \| x_j^s])$;

e° If $(B_1 \rightarrow B_2) \in \mathbf{ZD}$, $t_r g_i \,\varepsilon\, B_2$, $J_s g_j \not\varepsilon (B_1 \rightarrow B_2)$, then $f((J_s g_j E g_k \wedge B_1) \rightarrow B_2[t_r g_i \| o_s g_j]) = f((B_1 \wedge J_s g_j E g_k) \rightarrow B_2[t_r g_i \| o_s g_j]) = \forall x_j^s (((Q_j(x_j^s) \wedge Q_k(x_j^s)) \wedge f(B_1)) \rightarrow f(B_2)[b_i^r \| x_j^s])$;

$f((J_s g_j E g_k \wedge B_1) \rightarrow B_2[t_r g_i \| o_s g_j]) = f((B_1 \wedge J_s g_j \bar{E} g_k) \rightarrow B_2[t_r g_i \| o_s g_j]) = \forall x_j^s (((Q_j(x_j^s) \wedge \sim Q_k(x_j^s)) \wedge f(B_1)) \rightarrow f(B_2)[b_i^r \| x_j^s])$;

$f((J_s g_j E g_k \wedge B_1[t_r g_i \| o_s g_j]) \rightarrow B_2[t_r g_i \| o_s g_j]) = \forall x_j^s ((Q_j(x_j^s) \wedge Q_k(x_j^s)) \rightarrow f(B_1 \rightarrow B_2)[b_i^r \| x_j^s])$;

$f((J_s g_j \bar{E} g_k \wedge B_1[t_r g_i \parallel o_s g_j]) \rightarrow B_2[t_r g_i \parallel o_s g_j]) = \forall x_j^s\ ((Q_j(x_j^s) \wedge \sim Q_k(x_j^s)) \rightarrow f(B_1 \rightarrow B_2)\ [b_i^r \parallel x_j^s]);$

The consequence of these definitions is the following

LEMMA: (a) For every interpretation I there exists a semantics S_I, such that for any sentence A of language **L**

$$A \in S_I^* \text{ iff } I \models f(A).$$

(b) For every semantics **S** there exists an interpretation I_S, such that for any sentence A of language **L**

$$I_S \models f(A) \text{ iff } A \in \mathbf{S}^*.$$

From the lemma we conclude that

THEOREM: For any sentence A of language **L**,

$$A \in \mathbf{TAUT} \text{ iff } f(A) \in \mathbf{TAUT}_M.$$

The decidability of the set of tautologies of language L can now be easily inferred from the above theorem and from the known fact that the set $\mathbf{TAUT}_M$ is decidable.

Bibliography

1. Bańczerowski, Jerzy, Pogonowski, Jerzy and Tadeusz Zgółka (1982) *Wstęp do językoznawstwa*. Poznań: Wydawnictwo Naukowe UAM.

2. Nowaczyk, Adam (1971) ”Zaimki zamiast zmiennych i operatorów.” *Studia Semiotyczne* 2: 163-194.

3. Szober, Stanisław (1953) *Gramatyka języka polskiego*. Warszawa: PWN.