Witold Marciszewski

SEMANTIC ORGANISATION OF A TEXT

Originally published as "Organizacja semantyczna tekstu," Studia Semiotyczne 6 (1975), 105–124. Translated by Maja Wolsan.

Good order is the foundation of all good things.
Edmund Burke

1. The purpose of the concept of semantic organisation

We sometimes refer to some utterances as chaotic or disorganised. These two synonyms imply a negative approach. There is no term in the terminology of logic to negate chaos, i.e. to denote a certain order of thoughts which should be appreciated. In this paper, I will refer to this idea of an order of thoughts as 'semantic organisation of a text' or just as 'semantic organisation'.[1]

[1]Adopting the term "organisation" is explained by the fact that it is convenient to use a word which has two more grammatical versions: the adjective "organised" and the verb "to organise". The words "order" and "ordering", although also having the same advantage, are reserved to be used a bit further, namely to describe one of the elements or aspects of semantic organisation: they will be used precisely in the meaning ascribed to them in logic, in the theory of relations, when we refer to ordering relations. The adjective "semantic" is applied in the narrower meaning, which is becoming increasingly popular after Morris. It concerns the aspect of language which analyses the relations between the language and the reality described by it. Semantics implies syntactics, but it can leave aside pragmatics. In this paper, we will also leave aside the problems which could be called pragmatic organisation of a text, i.e. organisation taking into account the reactions of the recipient (for instance, many digressions which are not justified from the semantic point of view can be justified in the context of the pragmatic organisation, as they evoke a desired reaction of the recipient). An illustration of the semantic nature of these deliberations is the way the terms "concept" and "proposition" are applied here. Although these terms can be interpreted pragmatically, in this paper they are used in their semantically defined meaning — the basis for the definition of these terms is the concept of logical equivalence; they refer to the concept of the semantic model. Cf. Carnap 1956.

Edmund Burke (1729—1797), the British philosopher and politician who praised order as the basis of all good, would certainly agree that order in utterances brings about particularly valuable things: the knowledge that only ordered speech can convey; communication between the author and the reader, between the speaker and the audience; mutual respect between people, stemming from the fact that the concern for the order in utterances shows the author's respect for the recipient and gives rise to the recipient's respect for the author; and finally, aesthetical contentment provided by each lucid work.

However, it must be said that the idea of such an order of thoughts in speech is quite far from a clear definition which would make this concept lucid and operational. In the article, we will make an attempt at explication, i.e. we will put forward a sort of regulating definition, taking into account at least two applications: one related to the understanding of texts, and the other to evaluation by reviewers, editors, polemicists — in other words, all critical recipients of texts.

Any understanding, whether of a text, of human behaviour, or of the construction of a machine, consists in isolating elements of the whole and identifying the relations between these elements. In other words, it is about capturing the structure, the internal organisation of the analysed system. With regard to texts, there are at least two types of elements which can be distinguished in order to analyse the various relations between them: propositions and individual concepts. However, not all propositions expressed in a text and not all concepts deserve the same attention. Among propositions we distinguish the main theses — those which should be mentioned in, for instance, an abstract of the given text. Among concepts, we distinguish key concepts — those, which deserve to be mentioned in the subject index to the text.

The above distinction between an abstract or summary on the one hand and a subject index on the other hand is only the first approximation. If a subject index is expanded by adding to every term its relations with other terms (I propose to name such an index a relational index), then for every index there will exist an abstract of the given text which will be equivalent in content and different only in terms of graphical layout. This, of course, does not eliminate the point of drawing up both an index and an abstract to the same text, as each of them serves a different practical purpose.[2]

[2]I will not go deeper into the problems of indexing and drawing up abstracts, as I elaborated on these subjects separately in my two other works (Marciszewski 1970, 1972). In the latter, I introduced the notion of a relational index as information not

The coincidence that some indexes are equivalent to some abstracts does not change the fact that there are two ways of achieving these equivalent results: by choosing entire sentences and examining the relations between them, and by choosing individual concepts and examining the relations between them. The purpose of both these methods is to capture the structure of the text, i.e. to help understand it. However, none of them is universal, and thus both are worth applying if we aim at the fullest possible understanding of a text. As a matter of fact, observing a relation between some concepts is nothing more than making a proposition, and as in our analysis we only take into account the key concepts, what we get as a result is a set of the most elementary propositions, which in turn constitutes a kind of abstract, a list of key statements. Now, if we go deeper into examining the relations between these statements, we will get a more refined abstract, which is not replaceable by a relational index. However, if we start from this type of abstract, omitting the indexation stage, we might fail to notice some noteworthy relations between concepts. Only an index — a list of key concepts — makes it possible to juxtapose every concept with every other concept and then to attempt to find all essential relations using a combinatorial method. And only this can be called a full understanding, i.e. an exhaustive presentation of the semantic organisation of a text.

The same methods serve two more goals: the construction and the critical analysis of texts. The former is necessary for authors, who can use it to obtain a sort of algorithm for constructing texts, based on the concept of logical organisation. The latter is a tool for reviewers, who have to evaluate texts in terms of logical organisation, as well as for editors, who decide whether to publish a text, and for all other people whose task it is to evaluate written works in terms of their formal characteristics.

Explication, which will be employed here, consists in the following procedure: an unclear or vague term, used intuitively, i.e. without clearly formulated criteria of application, is replaced with a new term, of a technical nature, which is given explicit conditions of application, as precise as possible. The terms expressing intuitions which we are referring to are expressions such as ”order of thoughts”, ”transparency”, etc. The technical term which is supposed to replace them is the ”semantic organisation of a text”. In search of a precise definition of this concept, we obviously have to try to emphasize and respect the intuitions underlying the terms replaced by our new technical term.

only on concepts, but also on the theses of the analysed text.

It seems that the idea of the order of thoughts in a text is related to at least four intuitions, with relations between them yet to be examined. First of all, in order to write and speak in an orderly manner, one has to write or speak on a given subject, sticking to this subject from the beginning to the end. Secondly, a well constructed text is always divided into parts, such as chapters, sections, paragraphs, etc. Of course, the division cannot be mechanical or arbitrary — in given units of text it should group sentences concerning the same subject; thus, another condition for the organisation of a text is the logical division of the set of sentences of which the text is composed. Thirdly, there should exist (or at least it is desirable for it to exist) an order in the narrower sense, as it is understood in the theory of relations: in our set of sentences, or at least in some of its subsets, there should be an ordering relation between the elements. An example of such a relation is the time sequence in historical narration or an inferential relation in a deductive system. Fourthly, there should be something which linguists refer to as coherence of a text, i.e. the connections (not necessarily of an ordering relation type) between the elements of a text. For example, in historical narration, the connecting element is the name of the main character of the events in question, in syllogism it is the middle term, etc. I will call this feature coherence.

In the next sections, we will focus on each of these four aspects of organisation of a text, starting from coherence, which is, in a way, a fundamental characteristic of a text. In order to describe it, we will use the tool of graphical representation of a text in a coordinate system, in which relations between concepts are transformed into spatial relations. I will call the result of such a transformation the table of relations (for the given text). Every sentence containing key concepts is marked as a binary relation — on one axis we mark the elements of the domain of the relation, and on the other the elements of codomain, and on each intersection we mark the relevant relational predicate. The numerical indicator of coherence of the text will depend on the number of intersections, as its function. Other configurations presented in the table will provide indicators for some other characteristics of the text.

2. Text coherence

Let us use a very simple example of a coherent text — a conclusion method traditionally known as a polysyllogism. Let us assume, in order to adapt this example to our problem, that all terms with which the name variables would be replaced in this formula express key concepts. Below a

simple (only 'three-tiered') type of polysyllogism:

Each A is B
EACH B IS C
Each A is C
EACH C IS D
Each A is D
EACH D IS E
Each A is E

Each sentence of this text describes a relation of inclusion between the scopes of the terms. In order to note these relations in the table of relations, we shall introduce the following symbols: i for the relation of inclusion, $\bar{i}$ with a horizontal line above it for a converse relation, i in parentheses for the relations which are not mentioned in our polysyllogism but which may be inferred from it (this includes the relations of inclusion between several terms and all their converses). The vertical column shows the first argument, and the horizontal row the second argument of the inclusion.

Table 1.

	A	B	C	D	E
A	$=$	i	i	i	i
B	$(\bar{i})$	$=$	i	(i)	(i)
C	$(\bar{i})$	$(\bar{i})$	$=$	i	(i)
D	$(\bar{i})$	$(\bar{i})$	$(\bar{i})$	$=$	i
E	$(\bar{i})$	$(\bar{i})$	$(\bar{i})$	$(\bar{i})$	$=$

The number of intersections visible in the table, compared with the number of all possible intersections (n^2, where n is the number of terms), would be a natural indicator of the coherence of a text, intuitively understood as the 'density' of the conceptual network. It will be more convenient, however, to take into account not all possible intersections but only the ones grouped on one side of the diagonal, i.e. excluding everything that is on the other side of the diagonal and the diagonal itself, as the diagonal consists only of simple equal signs, and one side shows solely the converses of the relations which are on the other side. Both these types of data are trivial. After this modification, the maximum number of intersections p for n terms is:

(1) $max(p) = \frac{n^2---n}{2}$

As can be seen in Table 1, the text of the polysyllogism has maximum coherence if we take into account the relations of inclusion which can be inferred from the relations described explicitly in the text.

Minimum coherence would mean that each term is connected with only one other term, e.g. *A* only with *B*, but not with *C*, not with *D*, etc. For *n* terms the number of intersections will then be *n* minus one, as illustrated by the table below:

Table 2.

	A	*B*	*C*	*D*	*E*
A	=	+			
B		=	+		
C			=	+	
D				=	+
E					=

We subtract one from *n* because the first intersection will always be a trivial equality, and only the next ones will express some non-trivial relations. Thus, the minimum number of intersections, i.e. minimum coherence, is described by the formula:

(2) $min(p) = n - 1$

The coherence of the texts that we usually deal with will be somewhere in between the minimum and the maximum. For instance, a narration which speaks about the characteristics of one person or about what happened to that person, can be presented as the following table of relations, in which *N* is the name of the character, while *A*, *B*, etc. refer to the respective characteristics or events. Our relational predicate for statements about the character will be the symbol of belonging to a certain class, in this case *e* (from *est*), and the symbol for the direct preceding of events will be *a* (from *ante*).

Table 3.

	N	*A*	*B*	*C*	*D*
N	=	*e*	*e*	*e*	*e*
A		=	*a*		
B			=	*a*	
C				=	*a*
D					=

The coherence of this text is characterised by the number of intersections (7), which is between the maximum (in this case 10) and the minimum (4). For a comparative characterisation of coherence, it would suffice to use the indicators provided above, which make it possible to put every text on a scale between the two extremes. However, it would also be advisable to have at our disposal a classifying concept of coherence, i.e. one that would divide the set of texts into coherent and non-coherent. The natural demarcation line seems to be the medium value between the two extremes:

(3) $med(p) = \frac{n^2+n-2}{4}$

Now it is for us to decide (a bit arbitrarily), whether to assume as the condition of coherence that the number of intersections is to be greater than $med(p)$, or greater than or equal to it. It seems that we can adopt the second, more liberal approach. The fragment of narration analysed above seems coherent (it tells the subsequent events which happened to the same character), and the number of intersections is 7, so it is equal to $med(p)$.

We have already defined the notions of minimum, maximum, and medium coherence, and the classifying concept of coherence — by dividing a set of texts into coherent and non-coherent. For texts containing the same number of key concepts, here symbolised by n, there also exists a way of comparing their coherence if we take it as a gradable feature: out of two texts having the same number of key concepts, the one for which the table of relations shows more intersections is more coherent. We only have to define the method of comparing the coherence of texts with unequal numbers of key concepts. The natural measure of this will be the relation between the number of intersections in a given table and the maximum number of intersections for this table, here symbolised by $max(p)$. Moreover, it would be desirable to express the maximum and the minimum always with the same number, regardless of the value of p and n — for instance 1 and 0, respectively.

These conditions are fulfilled by the following formula for z (the grade of coherence of a text):

(4) $z = \frac{p - min(p)}{max(p) - min(p)}$

This function ascribes value 0 to $min(p)$ and value 1 to $max(p)$, and values from the range $0 \leqslant z \leqslant 1$ to the numbers between these two extremes; $med(p)$ receives the value 0.5. For example, if $n = 6$, the values are as follows:

Table 4.

value of argument p	5	6	7	8	9	10	11	12	13	14	15
value of function z	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1

The essence of the operation of moving from p to z consists in choosing as many numbers in the range [0,1] as there are in the range $[min(p), max(p)]$, where the subsequent numbers from the latter are attributed to subsequent numbers selected from [0,1], maintaining equal spaces between them (i.e. equal spaces between the numbers of one of these sequences are reflected by equal spaces between the relevant numbers of the other).

Considering that for the same text we will get different indicators of coherence, depending on whether we take into account solely the relations explicitly mentioned in the text or also those which can be inferred from it, actual coherence should be distinguished from potential coherence. The latter also covers the relations which can be expressed in the language of the text by using sentences inferred from the sentences explicitly belonging in the text. For example, in the polysyllogism examined earlier in this paper, potential coherence reached its maximum, while actual coherence had a value less than one, namely 0.5.

The relation between the actual and potential coherence of a text can be considered a measure of the feature for which there is actually no term either in every-day or academic language, but its opposite being colloquially referred to as pathology. We could call it coherence, but this term is ambiguous and only one of its meanings are close to the concept which we are dealing with here. The function attributing this feature, or a certain degree of it, to a text could be the difference between the potential coherence indicator and the actual coherence indicator. It equals zero when the text does not leave anything to the speculations of the recipient.

3. Monotopicality

A well organised text keeps to the topic. A more refined composition may contain several intertwining topics but we will focus on the simplest case, when there is one topic elaborated on throughout the whole text. We shall call this feature of a text 'monotopicality'. The table of relations for any topic instantly shows whether a text is monotopical or not. Monotopicality exists when there is a sequence of predicates in one of the columns or one of the rows which is stronger than all other sequences and in extreme cases even spans the whole row or column.

This is illustrated by the following example, in which symbols from A to F in the column always represent the first argument of the relation, the symbols in the upper row represent the second argument, and the "+" sign at the intersection shows that there exists a certain relation between the arguments (it may repeat or may be different each time).

Table 5.

	A	*B*	*C*	*D*	*E*	*F*
A	=	+				
B	+	=	+			
C	+		=	+		
D	+			=	+	
E	+				=	+
F	+					=

The topic of the text in the above example is A, as this concept enters in relations with all the other key concepts of the text in question. For instance, the topic of the present article, i.e. logical organisation, is in various relations to all the other key concepts discussed here, such as coherence, monotopicality, ordering, etc.

What is the relation between complete monotopicality, which consists in filling the whole row or column with predicates, and the feature of a text discussed in the previous section, called coherence? It can be easily observed that when the number of concepts $n \geqslant 6$, monotopicality is the sufficient condition of coherence. For Table 5, in which n equals 6, the coherence indicator is 0.5, and thus the text presented in this table should be classified as coherent (in the classifying sense of the term). When n is less than six, monotopicality contributes even more to coherence. For instance, if $n = 5$ and the text is monotopical, the coherence indicator will be 2/3. As shown in Table 5, after cutting out the sixth symbol (F) in both dimensions, the number of intersections is 8, which is a higher share of all possible

intersections than when $n > 5$, as the maximum in this case is 10 and the minimum is 4, which leads, according to the formula (4), to $z = 2/3$.

The condition of monotopicality is fulfilled by every text in which the same things are said repeatedly about the same object (which would be a pathological case) or different things are said about the same topic (which is a correct situation). The things 'said about something', in turn, can be grouped in classes. For instance, talking about the subsequent events in someone's biography, we group them into chapters according to certain periods of time (childhood, youth, 'the age of defeat', etc.) or according to certain topics. Such a division of a text can be multi-tiered if we divide larger parts of the text into smaller units, e.g. chapters into sections and then into paragraphs. Each of these units should be monotopical, and the topics of a fragment being a part of a larger unit should be related in a certain way to the topic of the higher unit. We will deal with these problems in the next section.

4. Structural adequacy

It is natural to look at a text as a set of sentences. In logic, an operation of dividing a set into subsets or the result of this operation is called partition.[3] Although in this case it is hard to define disjointness and completeness, which are to characterize the logical partition of a set, both are in a sense guaranteed in advance even in the most chaotic of texts, if we consider subsets of sentences grouped next to each other and separated from the adjacent groups (for example by a paragraph indent) as the result of partition. This way, each sentence belongs to a subset, i.e. to a fragment of text, and thus the partition is complete; furthermore, it is physically impossible for one and the same sentence, understood as an inscription (hence something material) to be in two places at the same time, which would violate the condition of disjointness. However, when talking about a correct division of a text, we do not have in mind any such trivial correctness. The intention behind such statements could be described as follows (which is probably not the only possible way).

In order to be able to examine the structure of a text, expressed by what is commonly called a division of this text, let us distinguish the physical and the thematic structure. Both structures can be presented as a tree. For

[3]Translator's note: the author uses one Polish word "podział", which corresponds to both "partition" and "division". I will use the word "partition" when talking about a division of a text wherever it is treated as a set of sentences, and "division" when it is used in the common meaning.

the physical structure, the top of the tree, marked by any number, e.g. 1, is the whole text, and the largest units (e.g. chapters) are the tree's upper branches; we will assign to them numbers consisting of two digits: 1.1, 1.2, 1.3, etc. Lower branches will be fragments of chapters, thus the third level from the top will feature numbers consisting of three digits: 1.1.1, 1.1.2, etc., and further 1.2.1, 1.2.2, etc.

In the thematic structure, the top of the tree is the general topic of the whole text, that is what the title of the text should describe. The subsequent levels of branches represent the topics which, if the correct line of thought is followed, should be the topics of chapters, and then parts of chapters, etc. We assume that for each main topic there is at least one correct way to divide it into branches and that not every branching is correct. What is the factor determining whether they are correct or not?

In order to find the answer, we have to notice that every topic can be assigned a question or a set of questions. For instance, the topic: "Semantic organization of a text" is equivalent to the following question: "What is semantic organization of a text?"; while the topic: "Development of speech in children who are deaf" can be read as the following question: "How to develop speech in children who are deaf?".

The transformation of a topic into a question is the first step towards creating a thematic tree. The question, in turn, is assigned a disjunctive answer, which is a part of the supposition of the question. The supposition itself is also a disjunction, but in the case of the *wh*-questions the disjunction usually has so many elements and is so hard to formulate by just enumerating all these elements, that we usually formulate it as an existential sentence.[4] The supposition of the question: "What events (in the life of a student) cause failures at school?" is the following sentence: "(Ex) (x causes failures at school)", in which the variable x belongs to the set of events.

However, what we start with as a given when solving a problem is a significantly limited disjunction, i.e. such a disjunction that entails only elements which can potentially — as far as we know — become answers to the question. For instance, nobody will put events such as the coronation of Charlemagne or the eruption of Vesuvius among the possible reasons for school failures. With such a far-reaching limitation of disjunctions of all answers, reducing them to the disjunctions of only these hypotheses which our current knowledge allows us to make, the disjunction defines the plan of the analysis (i.e. the plan of the process of problem-solving) and

[4]In Marciszewski 1974, I further developed the idea of a supposition of a question as a disjunction of answers which is the starting point for research,.

indirectly also determines the thematic structure of the text reporting on the results and possibly also the course of such research. Namely, the research process consists in eliminating the elements making up the disjunction of possible (as far as we know) answers, until only one answer is left. Thematic branches are formed by new questions appearing over the course of examining the individual elements of the disjunction. These questions concern data necessary for accepting or rejecting the examined element of the disjunction.

Naturally, not all steps of the analysis must be reflected in the text reporting on the research. Some elements of the disjunction which were taken into consideration and then eliminated by the researcher may be so uninteresting for the recipient that it is better to just pass them over in silence. The analysis includes some operations on a trial and error basis, while the report usually refers to only those steps which are necessary to understand the final result. For example, in the course of the reflections underlying this section of the article, there has appeared an idea to present the discussed feature of a text as the result of a logical partition of the set of sentences forming a text. This idea was therefore one of the elements of a disjunction of answers. It was rejected, which was reflected in some brief critical comments on this proposal at the beginning of this paragraph. These comments present only a certain final state; they do not reflect the stage of the analysis in which the said disjunction of answers was temporarily treated as the right answer. To sum up, the thematic structure of a text is a transformation of the structure of the analysis, formed by omitting some elements which are irrelevant for communication with the recipient, the structure of the analysis (in other words — the plan of the analysis) being determined by the initial problem leading to a disjunction of possible answers (formed by confronting the supposition of a question with our knowledge on the subject). This disjunction is branching into new questions generated by its elements, which must be answered in order to accept or reject the given element of the alternative.

After roughly outlining the method of obtaining/forming the thematic structure of a text, we can already evaluate the structural adequacy of a text in the following way: a text is structurally adequate if and only if there exists a thematic structure which represents the branching of the main topic of the text, illustrated by a tree which is isomorphic to the tree representing the physical structure of the text. This isomorphism consists in the fact that the images of both trees are identical, and where on the tops (points of intersection) of the tree of physical structure there are the numbers of text parts, in the tree of thematic structure there are the formulated topics of

the corresponding parts of the text.

One of the transgressions against structural adequacy is the existence of redundant elements in the physical structure of a text, i.e. elements with a topic not present in the thematic structure, for instance digressions which are not justified by anything other than the one incidental fact that the author had this particular association (however, if they are justified solely by, for example, didactic reasons, they may be justified in the pragmatic organisation of a text but will not be determined by its semantic organisation). Another transgression against structural adequacy — opposite, in a way, to the previous one — is the physical structure missing a fragment which should be there, according to the thematic structure of the text. This would be the case when a given thematic structure contains a topic which has no counterpart in the physical structure (a corresponding fragment of the text). We call it a gap in the flow of thoughts or skipping a thought. Yet another error occurs when there exist corresponding elements of both structures — i.e. each topic formed by branching the main problem has a corresponding fragment of text devoted to it and, at the same time, each fragment is devoted to one of the topics forming the thematic structure — but the corresponding elements are in different places in each of the structures, which means that they have different numbers in the numbering system describing their positions in the trees.

Naturally, in order to assess how structurally adequate a given text is, the two structures must be analysed separately. Thus, the thematic structure cannot be constructed on the basis of the physical structure. A person attempting to conduct a critical analysis of a text must construct, independently from the author, different possible thematic trees formed by branching the main problem (there may be more than one tree), in order to check whether one of them is isomorphic to the tree of physical structure. Undeniably, this would require a huge amount of work and skill on the part of the critic, but it seems that there is no other way to evaluate the structural adequacy of a text. It is worth pointing out (and may be of some solace) that the more a text fulfils the conditions of structural adequacy, the easier the work of a critic will be. Furthermore, the critic's ambitions might not be as high as to require her to propose alternative thematic structures in the event of detecting a defect in the structure of the analysed text: she may just be content with identifying the ambiguities or difficulties, which is much easier.

A numerical measure of structural adequacy can be formed in the following way. If the physical and thematic structures are exactly isomorphic

to each other, structural adequacy is 1. If there is no similarity between the two structures, as for instance in an unlikely case when a text entitled "The Tides of the Sea" would be about squaring the circle or breeding pigeons, structural adequacy is minimal and equals 0. The number of possible grades between 0 and 1 depends on the complexity of the thematic structure concerned. After calculating by a combinatorial method the number of possible differences between the thematic and physical structure, we could estimate how far the text is from maximum structural adequacy, measured by the number of differences actually occurring between the structures in relation to all possible differences.

5. Ordering of a text

Looking at a text as a set of sentences, we would say that it is linearly ordered — in accordance with logical terminology — if there is a linear ordering relation between the elements of the set, i.e. a relation which is transitive, asymmetrical, and total in the given set. We will say that it is partially ordered if there are no expectations as to the condition of totality, while the two other conditions are met. An example of a linearly ordered text (very hard to find) is the beginning of the Gospel of St. Matthew, which is a kind of lineage of the Messiah: "Abraham was the father of Isaac, Isaac the father of Jacob, Jacob the father of Judah" (etc.). The ordering relation for the set of ancestors is the relation of descent; the text describes this relation, therefore it comprises a relation ordering the set of names. Although the relation directly concerns names, it indirectly defines a certain order in the set of sentences as well. Texts are usually only partially ordered, without fulfilling the condition of totality. As mentioned further in this article, the existence of at least one partial ordering in a set of sentences should be considered a necessary condition to call the set a text. It is worth pointing out that the proposed conception of ordering refers to sentences, and not to a set of terms or concepts as in the previously analysed features of a text, like coherence. The reason for this is that one of the important relations ordering a text is the relation of deducibility, i.e. inferential consequence, which exists between sentences. When an ordering relation directly concerns names or concepts, like in the genealogy quoted above (another example would be historical narration observing a proper chronological order of events), it is usually possible to determine the order between the sentences defined by the order between the terms.

In the case of longer texts, such as an article, essay, or a short story, which are sets of sentences divided (also as regards their physical structure)

into subsets, the notion of linear order must be modified. Usually, linear order exists only within the set (family) of these subsets, while it does not exist in the set consisting of all sentences of the text. Moreover, if within each of these subsets — chapters, sections, paragraphs — the sentences making up these subsets are linearly ordered, the result is the same as with one linearly ordered set of sentences divided into these subsets. Namely, each sentence will have its clearly defined position in a given text — let us call it T — which means that any change of this position would create a text not equivalent to T (the equivalence in question would have to be defined separately, but we can assume, putting it simply, that it is an inferential or logical equivalence, as long as a text is treated as a conjunction of the sentences of which it is composed).

Such a combination of ordering and division is formally similar to e.g. alphabetical ordering of a set of surnames, in which every surname has a clearly defined position due to the linear ordering of sets, which are: the set of surnames starting with A, the set of surnames starting with B, etc. The subsets within each set are ordered according to a different relation, determined by second letters of the surnames, which gives us another sequence of subsets, each of them being further divided according to third letters, etc. Hence, it is not the case that there is one linear ordering relation which orders a set composed of surnames. There are various ordering relations, the difference between them being which letter of the surnames they concern, and each of them ordering not a set of surnames but a set composed of subsets of the set of surnames. However, the result is the same as in the case of one and the same ordering relation in a single and only set composed of certain surnames. It would be advisable to introduce a new term for this type of ordering, but to avoid using too many terms we shall extend the meaning of ordering of a set to include the above case as well.

In order to define a certain numerical indicator of ordering of a text, marked as $U(T)$, it is sufficient to take into account two values, symbolized by the letters z and i. The former, namely z, is the number of sentences making up the text T, while i is the number of texts which can be obtained from T by permutations of elements, without changing the meaning of T. The underlying idea of this condition is that these permutations are semantically irrelevant (hence the symbol i). As regards the criterion of identical meaning, it is rather inadvisable for it to be the same for all types of texts. The conditions for maintaining identical meaning are different in an academic text than in a feature article or in a poem. In academic texts, identical meaning could come down to inferential equivalence (cf. Ajdukiewicz 1960).

If T is linearly ordered, this means that there are no permutations which would lead to the creation of a text with identical meaning as T. In other words, each rearrangement of sentences introduces a shift in meaning, and thus $i = 0$. At the opposite extreme, there is the case when all rearrangements are irrelevant; for example, a sequence of expressions which is a record of a schizophrenic knight's move thinking , where $i = z!$. It seems natural to assume that this type of record does not deserve to be called a text, which is equivalent to assuming that the necessary condition for a text is that i is less than $z!$.

These considerations lead to the following function as an indicator for the ordering of a text:

(5) $U(T) = 1 - \frac{i}{z!}$

The function fulfils the above intuition when minimum order (i.e. no order) is ascribed the value 0, and maximum order is ascribed the value 1. In consequence, for a linearly ordered text, i.e. when $i = 0$, the indicator has the value 1, while for a 'knight's move thinking' type of text, when $i = z!$, the indicator has the value 0. Formula (5) will also provide an answer to the question whether a single sentence can be treated as a text, as long as we know what the value of i is in that case. There are two possibilities: $i = 1$ or $i = 0$, leading to contradictory solutions. Assuming that $i = 1$, a single sentence is not a text, while if $i = 0$, it is a maximally ordered text. The fact that formula (5) reflects this dilemma seems to speak in its favour, as indeed our intuitions in this matter are hesitant — as is often the way with borderline cases, to which our language is not adapted. On the one hand, a single sentence cannot be considered as unordered, just as a person who had no opportunity to sin cannot be accused of sinning. On the other hand, she can barely be called virtuous just because she had no opportunity to act against virtue. The final decision therefore remains a terminological issue, depending on certain practical aspects. For example, if we want to use expressions such as "the text of a signboard" or "the text of a title", it would be convenient to assume that a single sentence is also a text and ascribe value 0 to the i factor. Intuitively, this could be justified by arguing that a single isolated event cannot change its position in the sequence, as there is nothing in relation to which its position could be changed, and thus the only possible change would be to remove the sentence (delete it or cross it out), which would definitely be a relevant change. Therefore, the number of irrelevant changes, symbolized by i, would be 0, and the function $U(T)$

would have the value 1.

The metric notion of order should not be used as an evaluative notion in the sense that the closer the indicator of ordering is to 1, the higher the semantic organization of a text in the given aspect. For it is often the case that there are no factual reasons for a certain order of elements in a given line of thought. For example, when talking about two parallel events, it may be irrelevant in which order I choose to describe them, as by nature the relation in this case is symmetrical and thus is not an ordering relation, so it is not possible to introduce any ordering to the text. On the other hand, some level of ordering is necessary for a positive evaluation of the semantic organization of the text. Since the level of the necessary ordering depends both on the nature of the described object and the conventions governing the given genre, defining the minimum level of necessary ordering (other than zero, based on (5)) must be based on a detailed analysis of various types of texts. Thus a maximum ordering indicator is a sufficient condition to consider a text semantically well organized (in terms of ordering), but it is not a necessary condition. However, the indicator must be always higher than zero.

6. On methodological aspects

This essay, as well as the earlier papers to which it refers, is inspired not only by practical needs, but also by certain philosophical aspects. Let us now direct some attention to this philosophical basis and its methodological implications, comparing them to other current trends in methodology of human science.

I propose (maybe with some exaggeration but with a benefit to clarity) to call the approach specific to this essay 'neophysicalism'. It would be different from physicalism advocated by early logical empirism, in particular by Carnap, in two aspects: in the object and the postulated range of reduction.

The object of reduction in classical physicalism are psychological theses as statements about non-observable (from the outside) states of mind, which it proposes to translate into statements about bodies subject to external observation. Classical behaviourism is an attempt to realise this programme. The essence of what I suggest to call neophysicalism, is the attempt to relate statements about some intentional objects (as defined by Ingarden),[5] such

[5]The numerous texts by Ingarden on the problem of intentional objects include Ingarden 1960a: 141ff.; 1960b: 180ff. I provide more detailed bibliographic information in Marciszewski 1973.

as meanings, thought structures, conceptual apparatus, etc., to observations concerning physical objects or spatial arrangements, such as inscriptions and their configurations on a plane. Using the terminology of Ryle (1957) and Popper (1968), we could say that both physicalisms postulate a reduction to the 'first world', i.e. the material or physical world, but the older type of physicalism attempts to do it only with the 'second world' — the domain of psychical phenomena, while the new one, proposed here, covers the 'third world' — the domain of intentional objects — as well. It should be stressed here that this neophysicalism can, but not necessarily has to be understood as an ontological position. We can restrict ourselves to treating it as a method of analysing cultural phenomena; even if this method has some underlying philosophical theses, they are weaker, less decisive than the thesis of physicalism or ontological somatism. The intention of the above deliberations is to support methodological neophysicalism.

As regards the scope or level of radicalism of the reduction programme, the difference between the two physicalisms is analogous to the one between (classical) behaviourism and neobehaviourism. The radical and unsupportable programme of behaviourism, identical to the programme of physicalism, postulated absolute translation of all psychological concepts into physical ones, thus totally eliminating psychological concepts. Neobehaviourism, in turn, presents a much more moderate postulate: to define the ways of identifying internal states of a human or animal by external states, observable by senses and describable in physical terms; e.g. the force with which a rat struggles to reach its food, measured by the stretching of a spring, may be treated as an indicator of hunger. Therefore, the programme of total reduction, i.e. translatability of some terms into others, was replaced by a programme of partial reduction, which in terms of logic is expressed in the fact that the sentences linking the two systems of concepts, called reductionist definitions, are not equivalences but mere conditionals (cf. Kotarbińska 1966, Przełęcki 1966a, 1966b).

The postulate of total reduction of intentional objects to physical ones has never been advocated under the name of physicalism, but some similar proposals with respect to items related to language, such as propositions or concepts, can be found in the traditional nominalism. Thus, replacing this radical postulate with the programme of partial reduction, we could coin the term "neonominalism" for the latter. Such a term would, however, be less clear, as there are already many historical layers of nominalism. Another possible term would be "formalism", also with the prefix "neo-", but this term has already become too ambiguous.

Neophysical reduction does not imply a negation of the existence of intentional objects. This ontological question may remain open, just as in neobehaviourism it is possible to pursue the methodological programme without getting ontologically involved into the question of the existence of mental phenomena. However, in the name of intersubjectivity and practical effectiveness of some actions performed on texts, it is postulated that the characteristics of intentional objects, such as e.g. the coherence of a thought construction or the significance of an idea in a given moment, should be described and identified by using physical characteristics, such as a spatial arrangement in the table of relations. This way, the operations involved in understanding or creating a text can be presented in a semi-algorithmic way, by using directives of indexing, abstracting, branching a problem, etc. The operations involved in the evaluation of a text will be supported by the methods of calculating the indicators of coherence, ordering, etc.

This approach to texts and to other creations of culture involves a certain conscious deformation of an object. Indeed, in reality, the semantic organization of a text, even a 'dry' academic text, is independent from its pragmatic organisation. It is also not the case that keeping to the same topic in a text will always be reflected in a specific arrangement of entries in the table of relations. Certain side factors may prevent it, like rich vocabulary, using many various synonyms for the same concepts, or digressions motivated by pragmatic reasons. Such observations can be formulated at every step and could even be treated as reservations against the method of text analysis described in this essay.

Any possible objections should generally be answered by the following paradox: that these kinds of deformation of reality are a necessary condition of transformation. The intuition of an object, in the entire richness of this object and in its essence, is something very valuable, which serves the development of our intellect and our sensitivity, but we should not expect it to provide us with any effective directives of conduct. Bergson understood this very well when he attributed to intuition a contemplative value — inspiring to further learning and acting, but far from technical applications. It is the reality-deforming sciences (and only them), with their abstract and fixed conceptual frameworks, that — according to Bergson — can provide data for practical use.

An excellent illustration which seems to support the above thesis are the achievements of phenomenologists, in particular of Ingarden, in the field of reflections on the language and on various types of texts (cf. footnote 6 above). If the addressees of these reflections struggle through the maze

of conceptual distinctions and — trustingly (such an emotional approach seems to be necessary — let the intuitions of the author lead them, what they will get in return is richness of understandings. Their spiritual sight (if we choose to call it in this exalted manner) abounds with various curiosities of this world of intentional beings, which then becomes as tangible as the reality of our body or mind. Enriched by these experiences, we are likely to become more capable, or at least we have the opportunity to increase our skills, to conduct reflections or research on this world, including research of the kind presented in this essay. This impact, however, is indirect — by stimulating our ability to understand, and not by providing direct premises for practical directives. If we are looking only for direct premises, we can ignore the problem of the existence of intentional objects, the problem of their nature, etc.

It might seem paradoxical that the advantages of practical cognition are growing with every simplification or deformation made by us, but it dissolves after analysing any example. When a wild animal attacks me in a forest, the effectiveness of my defence depends on how much I am able to disregard some characteristics of the situation (e.g. the beautiful silhouette of the animal), in order to concentrate on the characteristics which are relevant in the given situation, such as the strength and agility of the attacking animal and its tactics. Consequently, my perception of the animal is incomplete, and thus deformed, like a caricature emphasizing only some features; but it is this deformation that gives it practical productivity. It would also be unadvisable to apply in situations like this the phenomenological postulate of suspending our own approaches and exposing ourselves totally to the impact of the contemplated object. For it is the approach of the subject that defines the point of view on the object.

Let us say it even more clearly — the passiveness of the cognitive subject is beneficial for contemplative cognition, but not beneficial to practically oriented cognition, which performs a vivisection of the object in order to 'cut out' an aspect of answer for specific practical needs. This is also probably one of the demarcation lines between philosophy and science: a philosopher aims at a contemplative description of the entire world, with respect to its complexity, while a scientist acts more unceremoniously, cutting the reality into various aspects, simplifying them when necessary, applying approximations required by his objectives. In this sense, the phenomenological concept of text is a philosophical concept, while the concept developed here is — at least in its intentions — a draft of a scientific theory.

When drawing such comparisons, we have to mention one more approach

to the problem of texts, called the theory of humanist interpretation, and applied for instance by methodologists from the Poznań philosophical circles (cf. e.g. Kmita and Nowak 1968, Zamiara 1974). It is close to the present deliberations as regards the approval of intentional simplifications, in the theory called idealizations. The fundamental idealization thesis of this theory is the assumption that the subject, in this case the author of a text, is rational. We must admit that this assumption plays an effective role in many analyses or interpretations of texts. If in the postulates characterising rationality we included the principle of non-contradiction, as it is usually done, we would get a directive for different types of interpretation, including abstracts. The directive would be as follows: if in the analyzed text there are both sentences S and non-S, we should not assume that the author accepts the contradiction but that in each case the sentence S expresses a different proposition.

However, it would be a misunderstanding to limit the idealizations in the interpretation or analysis of a text only to the assumption of rationality in this version or another. Another idealization underlying these deliberations as their philosophical basis is that some non-physical features of a text are attributed to certain physical features. The table of relations shows the presence of some of them. For instance, sticking to one subject (a non-physical feature) has a corresponding physical feature which is the sequence of entries in the table of relations filling one whole row or column of the table. In this case, idealization consists of approximation, which in some cases is realized in full, and in others creates a supporting question for the interpretation: why is there no full indicator of monotopicality in this case? The reason might be the multitude of parallel topics, as well as ambiguity of some terms, ineptness of expression or of composition, the existence of digressions. Taking into account these kinds of factors, we achieve a crystallization of our idealizing starting point.

In this essay, using ~~a~~ slightly polemical stylistics, which takes into account the point of view of a possible opponent, the above-mentioned idealizations have been called deformations. Indeed, they lead to a different image of a text than that presented by various concrete texts; let us remember, however, that it is not the concrete thinking but the abstract thinking that contributes to cognition which serves practical purposes, among others.

Such an idealization or deformation is related to all characteristics of a text described above, which are, apart from monotopicality: coherence of a text as the frequency of interrelations between the key concepts, structural adequacy understoood as physical structure consistent with the thematic structure, and finally order, being the result of an ordering relation in a text.

The above list is not necessarily complete and is accompanied by awareness that these are all surface features, not reaching into the depths of thought opening in many texts. But if all authors took the effort to ensure that their texts have these surface features and if all critics demanded it from authors, there would be much less texts in the world which owe their ostensible depth to a clouded surface.

Bibliography

1. Ajdukiewicz, Kazimierz (1960) "Język i znaczenie" ["Language and Meaning"]. In *Język i poznanie*, vol. 1, 145—174. Warszawa: Państwowe Wydawnictwo Naukowe

2. Carnap, Rudolf (1956) *Meaning and Necessity*. Chicago (IL): University of Chicago Press.

3. Ingarden, Roman (1960a) *Spór o istnienie świata* [*The Controversy over the Existence of the World*], vol. 1. Warszawa: Państwowe Wydawnictwo Naukowe.

4. Ingarden, Roman (1960b) *O dziele literackim* [*The Literary Work of Art*], Warszawa: Państwowe Wydawnictwo Naukowe.

5. Kmita, Jerzy and Leszek Nowak (1968) *Studia nad teoretycznymi podstawami humanistyki*. [*Studies on the Theoretical Foundations of Human Science*]. Poznań: Uniwersytet im. Adama Mickiewicza.

6. Kotarbińska, Janina (1966) "Ewolucja Koła Wiedeńskiego" ["The Evolution of the Vienna Circle"]. In *Logiczna teoria nauki*, Tadeusz Pawłowski (ed.). Warszawa: Państwowe Wydawnictwo Naukowe.

7. Marciszewski, Witold (1970) "Sposoby streszczania i odmiany streszczeń" ["How to Prepare Abstracts and Types of Abstracts"]. *Studia Semiotyczne* 1: 151—167.

8. Marciszewski, Witold (1972) "Formalna charakterystyka dziedziny rozważań jako podstawa indeksu rzeczowego" ["Formal Description of the Domain of Discourse as the Basis for a Subject Index"]. *Studia Semiotyczne* 3: 85—101.

9. Marciszewski, Witold (1973) "Problem istnienia przedmiotów intencjonalnych" ["The Problem of Existence of Intentional Objects"]. *Studia Semiotyczne* 4: 189—206.

10. Marciszewski, Witold (1974) "Analiza semantyczna pytań jako podstawa reguł heurystycznych" ["Semantic Analysis of Interrogtatives as a Basis for Heuristic Principles"]. *Studia Semiotyczne* 5: 133—146.

11. Ryle, Gilbert (1957) "The Theory of Meaning". In *British Philosophy in the Mid-Century*, C.A. Mace (ed.). London: Allen and Unwin.

12. Popper, Karl R. (1968) "Epistemology without a Knowing Subject". *Logic, Methodology and Philosophy of Science* 3: 333—373.

13. Przełęcki, Marian (1966a) "Pojęcia teoretyczne a doświadczenie" ["Theoretical Concepts and Experience"]. In *Logiczna teoria nauki*, Tadeusz Pawłowski (ed.). Warszawa: Państwowe Wydawnictwo Naukowe.

14. Przełęcki, Marian (1966b) "Operacjonizm" ["Operationalism"]. In *Logiczna teoria nauki*, Tadeusz Pawłowski (ed.). Warszawa: Państwowe Wydawnictwo Naukowe.

15. Zamiara, Krystyna (1974) *Metodologiczne znaczenie sporu o status poznawczy teorii* [*The Methodological Significance of the Controversy over the Epistemic Status of Theories*]. Warszawa: Państwowe Wydawnictwo Naukowe.