Philosophica 18, 1976 (2), pp. 129–138.

ON SOME GENERAL PRINCIPLES OF SEMANTICS OF A NATURAL LANGUAGE

H. Hiż

The problem of the relation which language bears to reality is old, venerable, and afflicted with the hereditary disease of obscurity. At least from the time of Frege, it has been a central problem for philosophers, while linguists have tried to keep away from it as a rule. In recent years, however, several influential theories of semantics of a natural language have been proposed, and some linguists as well as philosophers take part in their advancement. One of the theories which is gaining momentum, mainly through the work of Richard Montague, assigns objects of set theory to various phrases of a natural language like English or Italian. If somebody doubts the nature or the existence of these objects, he may choose to follow those who assign logical formulae to sentences of a natural language and claim that a sentence says the same as the corresponding formula. This is practiced by linguists who call themselves "generative semanticists." A somehow similar idea is found in the genotype language of S. K. Shaumian. Furthermore, if one doubts that whatever is expressible in a natural language is expressible in logic as well, he may side with the interpretativists, according to whom there are interpretative rules for a language and these give rise to the semantics of the sentences. If I understand what these interpretative rules are, what I have to say later does not depart very far from the interpretative theory, though here it will be twisted in a new way.

Linguists as a rule practice far more restrained semantics than in the theories just mentioned, and I will side with that tendency. In this respect it is instructive to reflect on the attitude of Harris towards semantics. In his work he tries both to reduce semantic considerations in linguistics to its minimal and fundamental role and to capture in his syntax the main general semantic properties of a language. Contrary to general belief, in his early book *(Methods in Structural Linguistics,* 1950) Harris did not eliminate semantics. The semantic input into his linguistics was reduced to a single, simple, and testable question : are two utterances repetitions of each other, or do they contrast. *Book* and *hook* contrast and from this the linguist reconstructs two different phones /b/ and /h/. To contrast means to not be a repetition, to say or to mean something else. This rudimentary semantic element was never eliminated from Harris' work, and — I may add — is at least tacitly assumed in all linguistic efforts. It is present in phonetics, in syntax, in discourse analysis, in field methods, in comparative studies. On the other hand, the entire effect of Harris' syntax, including the latest work (On a Theory of Language, *Journal of Philosophy,* 73, May 20, 1976), is oriented toward rendering semantic differences by syntactic means. His syntax is always semantically motivated, in spite of the changes in form throughout the years. It is not that the result of the syntax, the derived sentences will later receive semantic interpretation, but that each syntactic step reflects or records a semantic property.

A restrained semantics may use only truth as a link between language and the world. A still more restrained semantics would not talk about truth but about relations of equivalence (simultaneous truth) or of consequence (preservation of truth), and would not say which sentences are true but only that if a sentence is true, then another sentence is true also. It is clear that linguists traditionally do not engage in a strong semantics. But it is less clear whether a very weak semantics suffices for the purposes of a linguist. It may be that to make an insightful syntax and discourse analysis one has to know that some sentences are true or, at least, that some sentences are considered true by the speech community whose language is under study. In what follows, I will use the concept of truth as primitive. I will not define it but I will use it in axioms and those axioms will to some extent tell how I use that concept, and how it relates to other concepts used in the axioms. But this information is only very partial. The axioms are of course far from complete. They are only some starting principles for understanding the semantics of a language. There will be two kinds of axioms for semantics; those which are independent of particular phrases, of a particular vocabulary of the language and those which do relate the general semantic concepts to some particular vocabulary of a particular language. The first kind of axioms are therefore general principles of semantics of any language. I will discuss here six such principles. All of them use the concept of truth, the concept of consequence or both. They of course also use the concept of a sentence. Each of the three concepts are to be taken with proper caution. To avoid misunderstanding, I must say that I take truth to be a property of some sentences (like Tarski) and of some finite strings of sentences.

Not all sentences of a natural language are true or false if taken out of context. Some of them are like sentential functions of a formal language, like '$x=s$' which is neither true nor false; it has a form of a sentence but is not a sentence, since it has a variable, a blank. In a natural language variables do not occur but other devices are used which make a sentence not a suitable argument for truth or falsity. For instance *It was John* or, *They revolted* contain referential phrases (*it* and *they*) which refer the listener to occurrences of other phrases in the same utterance. Only after knowing to what occurrence of a phrase *it* refers, it is reasonable to ask whether *It was John* is true in that context. For instance, if the text is *Jane did not give me the book. It was John,* we conclude that *It was John* contains a cross-referential to *gave me the book* and that what we are asked to judge the truth of is *John gave me the book.* One may still ask what book is spoken about in this sentence. *The book* is a referential to some previous phrase. One must distinguish, therefore, between the truth of a single sentence and the truth of that sentence as a segment in a string of sentences. The former may be improper whereas the latter may apply correctly. We may then easily say that a sentence with a cross-referential is true in a text, if the text which also contains the referend is true.

Similar complications are encountered when one analyzes the concept of consequence. *John gave me the book* does not follow from *It was John,* but it does follow from *Jane did not give me the book. It was John.* One has to distinguish between a sentence α being a consequence of the string of sentences $\beta_1 {}^\frown \beta_2 {}^\frown \beta_3$ (where the arc is concatenation), from α being a consequence of the set of which β_1, β_2, and β_3 are members. Thus, if $\alpha \in \mathrm{Cn}\,(\{\gamma_1\} \cup \{\gamma_2\})$. then it is not necessarily the case that $\alpha \in \mathrm{Cn}\,(\gamma_1 \cup \gamma_2)$. Here, as usual, '$\cup$' is a sign of logical sum, '$\in$ Cn' is for 'is a consequence of'. Some of the axioms are due to Tarski, others are slight extensions and modifications of his ideas.

Axiom of Sentencehood :

(A1) $\alpha \in \mathrm{Cn}\,(X)$, then ($\alpha \in S$ and $X \subset S$)

If α is a consequence of X, then α is a sentence and X is a set of sentences.

As the relation of consequence rarely holds between a sentence and a single sentence, we rather say that $\alpha \in \mathrm{Cn}\,(\{\beta\} \cup X)$ in cases where α is a consequence of β with the help (presuppositions, obvious truth, general sentences known in the community, logical rules) of sentences of X.

The following two axioms are clear and are the expression of the facts that the set T of all true sentences is consistent and complete :

(A2) Not $(T = S)$

(A3) If $(\alpha \in S$ and not $(\alpha \in T))$, then $Cn\ (T \cup \{\alpha\}) = S$

One may be tempted to claim that if $\alpha \in Cn\ (\{\gamma_1\} \cup \{\gamma_2\})$, then $\alpha \in Cn\ (\{\gamma_1\}^\frown \{\gamma_2\})$. But this does not necessarily hold. For, in a natural language, not any two sentences can be conjoined by a pause (represented by a period in the written form). The most we can claim is that any two sentences can be conjoined by one of many available conjunctive phrases and that the resulting text implies each of the conjuncts (but is not necessarily equivalent to the set of them).

Both a sentence, even if it contains an unresolved cross-referential, and a string of sentences are here called sentences. A string of sentences is therefore understood with a pause (period) between consecutive sentences. The term *sentence* is ambiguous. Sometimes it is used for an intonational unit (between pauses), at other times for a syntactic unit (a terminal string of a generative grammar), or for a logical or semantic unit. These (and other) understandings of the term are usually confused with each other. Here, only the third is intended. The first argument of the relation Cn, (i.e. a consequence) is a sentence. Also, every member of the second argument of that relation (i.e. a premiss) is a sentence. Also, a part of a premiss which gives a consequence of the premiss after proper replacement of cross-referentials by their referends is a sentence. In the previous examples *It was John* is a sentence in the premiss *Jane did not give me the book. It was John.* So *John gave me the book* is a consequence. (It takes some mechanism to describe the change from *It was John* to *John gave me the book.* One way may be to consider that in *It was John* there is a zero occurrence of a referential at the end, so that replacement of that zero occurence by the referend gives *It was John who gave me the book.* Here *who* enters as a necessary adjustment of grammatical categories. From *It was John who gave me the book* one obtains *John gave me the book* by a usual consequence transformation.

The following is the Axiom of Conjunction :

(A4) if $\alpha_1, \alpha_2, \beta_1, \beta_2, \alpha_1{}^\frown\alpha_2, \beta_1{}^\frown\beta_2$, are sentences and X is a set of assumptions, then there is a sentence γ such that $\alpha_1{}^\frown\alpha_2 \in Cn\ (\{\alpha_1{}^\frown\gamma\} \cup X)$ and $\beta_1{}^\frown\beta_2 \in Cn\ (\{\beta_1{}^\frown\gamma\} \cup X)$ and $Cn\ (\{\alpha_1{}^\frown\gamma\} \cup X) \neq S$ and $Cn\ (\{\beta_1{}^\frown\gamma\} \cup X) \neq S$.

The axiom of conjunction is a very useful tool for linguistic study. It allows one to consider one sentence γ instead of two sentences α_2 and β_2; then γ says at least as much as α_2 and β_2 together. The other factors α_1, and β_1, are to provide the contexts in which α_2, β_2 and γ appear, so that the principle applies to sentences with anaphoric referentials whose referends occur in the contexts. If α_2 is *John goes to school* and β_2 is *2+2=5* , then γ can be, e.g., *John goes to school in order to learn that 2+2=5*. Both *John goes to school* and *2+2=5* follow from γ. But γ does not follow from these two sentences, and their conjunction is not an acceptable English sentence. Here, *in order to learn that* is a conjunctive functor which forms γ from α_2 and β_2. In other cases it may take some ingenuity to find a proper conjunctive functor. The last two conditions in the axiom of conjunction insure that γ be chosen in a way consistent with what was said and known. Otherwise the axiom will be trivially true. The axiom of conjunction permits one to reduce grammatical problems which deal with a (finite) set of sentences to problems with a single sentence. Thus, when Harris and others derive a sentence by a binary transformation from two kernel sentences, it could also have been derived from a single sentence. It is important in practice to find as weak a γ as possible. The best is *and* (or a pause), for then we may have, barring some complications, in addition $\alpha_1{}^\frown\gamma \in \mathrm{Cn}\ (\{\alpha_1{}^\frown\alpha_2\} \cup X)$ and $\beta_1{}^\frown\gamma \in \mathrm{Cn}\ (\{\beta_1{}^\frown\beta_2\} \cup X)$. A γ which is equivalent to the set composed of α_2 and of β_2 does not always exist. Often, a conjunction says something more than the two conjuncts (like in the example with *in order that*).

For some pairs of sentences α_2, β_2 not only an inferentially equivalent sentence does not exist but neither there exists the weakest sentence that gives both α_2 and β_2 as consequences. There may be sentences δ_1, δ_2... such that each of them is a consequence of the preceeding one, but not vice versa, each of them has α_2 and β_2 as consequences and neither of them is a consequence of the set $\alpha_2 \cup \beta_2$. In the example above, *John goes to school in order to learn that 2+2=5* has as a consequence *John goes to school where he may learn that 2+2=5*. The first sentence is not a consequence of the second and *John goes to school* and *2+2=5* are again consequences of the second sentence. In this case it is hopeless to find the closest conjunction, the best fit. When for given α_2, β_2 and X there is a sentence γ such that $\alpha_2 \in \mathrm{Cn}\ (\{\gamma\} \cup X)$, $\beta_2 \in \mathrm{Cn}\ (\{\gamma\} \cup X)$ and for every δ if $\alpha_2 \in \mathrm{Cn}\ (\{\delta\} \cup X)$ and $\beta_2 \in \mathrm{Cn}\ (\{\delta\} \cup X$, then $\gamma \in \mathrm{Cn}\ (\{\delta\} \cup X)$. we may say that γ is the least upper bound for α_2 and β_2. The existence of the least upper bound for any two sentences would lead to a possibility of treating the consequence relation algebraically, forming a semi lattice. But the least upper bound hypothesis does not seem plausible. Therefore for linguistics purely algebraic methods will not suffice. Rather some topological methods may be needed.

For logic, and for other mathematical systems, one often asserts the principle of compactness which says that if a sentence is a consequence of a set of sentences, then it is a consequence of a finite part of the set :

(A5) if $\alpha \in$ Cn (X), then there is a Y such that Y $\subset$ X, Y is a finite and $\alpha \in$ Cn (Y).

For some systems of logic this principle fails. Whether the principle of compactness holds for a natural language or not has not been sufficiently examined. The situation in this respect is not exactly the same as for logic. One reason for this is that a natural language contains its own metalanguage and allows reasonings going from premisses to metalinguistic conclusions which in turn may give consequences in the object language. In practice the compactness axiom is often used. However, any time it is used, it should be noticed that a certain risk is taken, or a restriction is imposed.

The next general principle connects the concept of consequence to that of interpretation. Interpretation is a three-place relation between a phrase, a sentence, and a phrase in that sentence. There is a similarity, and a fundamental difference, between interpretation and satisfaction. A sequence of objects satisfies a sentential function (a matrix), if the replacement of the variables by names of the successive objects of the sequence (the conforming free variables being replaced by names of the same object) results in a sentence which is true. We can generally say that the sequence $\langle a_1, a_2, \ldots \rangle$ satisfies '$f(x_1, x_2, \ldots, x_n)$' if and only if $f(a_1, a_2, \ldots, a_n)$, provided that a_1 is the kind of thing which is named by phrases which the variable 'x_1' represents, a_2 the kind of thing which is named by phrases which are represented by 'x_2' etc.. In such Tarskian constructions we assume that we know which kind of things are named by which kind of phrases represented by this or that variable. It is an assumption which for a natural language is hard to make. Take one example. *Blind* is usually labeled as an adjective and in set-theoretical view as a name of a property, a mapping from individuals to individuals, from a man to a blind man (or from a class of men to a class of blind men). But in the sentence *This is a home for the blind, blind* is used as a substantive and therefore as a name of a class rather than a mapping. French requires a plural substantive : *C'est la maison pour les aveugles.* In Polish both the singular and the plural form can be used : *To dom niewidomego* and *To dom niewidomych* (both in genitive, but the plural can also be with *dla* ('for') : *To dom dla niewidomych.*). More importantly, the set theoretic model for a language requires that there be a set of individuals; all relations are either between the individuals or between relations between individuals, and so forth. But in natural language often there is no obvious way to base it on a single set of individuals. For one thing, mass nouns (*water, wine, flesh*) constitute another basic set, and not a subset of count nouns (*a man, a book, a house*). (About the problems presented by the axiom of foundation in set theory see my forthcoming paper "The Pretense of Existence.") Instead of satisfaction one may take a different relation, namely interpretation which holds between some linguistic phrases and does not involve any other objects. It relates to the world only by truth. *A phrase α interprets the sentence β at γ* means, roughly, that if α replaces γ in β, the result is true. The result of the replacement of γ by α in β will be represented by *repl* $(\beta\ \gamma/\alpha)$, and the result of the replacement of $\gamma_1, \gamma_2, \ldots$ by $\alpha_1, \alpha_2, \ldots$ in β will be represented by *repl* $(\beta, \gamma_i/\alpha_i)$. Thus, the word *sonnet* interprets the sentence *Petrarch wrote a novel* at the occurrence of *novel* because it is true that Petrarch wrote a sonnet. The formulation is only a rough one, for the phrase which interprets a sentence at a place must play the same grammatical role in the result of the replacement as the replaced phrase did. Otherwise, the matter will go out of hand and the theory be useless. For instance, if we replace *novel* in the sentence *Petrarch wrote a novel* by *sonnet and was considered a great poet,* we obtain a true sentence as well but *sonnet and was considered a great poet* is not of the same (if any) grammatical category as the substantive *novel* was. We must make our theory of interpretations relative to a theory of grammatical categories. Also, of course, one has to generalize the concept of intepreting to cover the case of a sequence of phrases to interpret a sentence or a set of sentences at proper places. Generally, f is an interpretation of the set of sentences X if f is a sequence of pairs $\langle \alpha_i, \gamma_i \rangle$ and every sentence in X results in a true sentence after replacing each γ_i by α_i.

$f \in \text{Int}(X)$ if and only if $X \subset S$ and there is a sequence $\langle \alpha_i, \gamma_i \rangle$ such that $f = \langle \alpha_i, \gamma_i \rangle$ and for all β if $\beta \in X$ then repl $(\beta, \gamma_i/\alpha_i) \in T$. (For more about interpretation see my paper "Alethetic Semantic Theory", *Philosophical Forum,* I, 1969).

With the help of the concept of interpretation, one can formulate the Axiom of Consequence :

(A6) $\alpha \in \text{Cn}(X)$ if and only if for all f, if $f \in \text{Int}(X)$, then $f \in \text{Int}(\{\alpha\})$

A sentence α is a consequence of the set X of sentences exactly when every interpretation of X is also an interpretation of α. This axiom looks like a definition of consequence in terms of interpretation. But this is an illusion. For the concept of interpretation involves the concept of grammatical category, and grammatical category, I would claim, requires, in turn, the concept of consequence. A grammatical category is just a role that a fragment of a sentence plays in consequences. The structure of a sentence is imposed by the fact that consequences are drawn from the sentence. An adjective, approximately, is a phrase α in the following rule of inference, where 'Det' stands for 'determiner' : ψ (Det$\alpha\beta$) $\rightarrow$ Det β *is* α. (E.g., *I wear a blue sweater* $\rightarrow$ *A sweater is blue.)*

From the Axiom of Consequence, taking the identity as an interpretation. it follows that if $X \subset T$ and $\alpha \in Cn(X)$, then $\alpha \in T$. Thus, consequence is truth preserving. Of course one can build a formal system in which the formula used here as the axiom of consequence will be a definition and interpretation will be accepted as a primitive concept and grammatical categories will be given in a lexicon. Such a system may be elegant and instructive. But in empirical work consequences are given and categories (even the category of a sentence) are constructs on the basis of a mass of material on consequences.

Besides the general axioms (A1 — A6), which are independent of any particular language, there should be axioms which connect the concept of consequence, and therefore other concepts appearing in the general axioms, with the specific phraseology of a language under study. For that purpose one may choose some phrases of the language as (grammatical) constants. The choice of constants is, in principle, arbitrary but some choices are better than others. For the language of logic, ordinarily, *if, not,* and *for every* are chosen. Sometimes *is a* is added; or, alternatively, *set,* or *class* and *set.* One may add *may be* or *but* or *although,* each time obtaining a variety of modal or relevance logic. The chosen constants are not subject to the reinterpretation referred to in the axiom of consequence. If α is a (grammatical) constant, then the only interpretation allowed for it in (A6) is identity; if α is a constant, then for the purpose of (A6) $Int(\{\alpha\}) = \alpha$. If a language contains a negation, N, then it would be reasonable to add

(A7) If $X \subset S$ and $\alpha \in X$ and $N(\alpha) \in X$, then $Cn(X) = S$.

For languages with the conditional (*if, then,* symbolically : $\supset$) one adds

(A8) If $X \subset S$ and $\alpha \in X$ and $(\alpha \supset \beta) \in X$, then $\beta \in X$.

or

(A8') $\beta \in Cn(\{\alpha\} \cup \{\alpha \supset \beta\})$

Also, the rule of deduction is added (or derived) for languages with the conditional :

(A9) If $X \subset S$ and $\alpha \in Cn(X \cup \{\beta\})$, then $(\beta \supset \alpha) \in Cn(X)$.

For English one may choose as constants those words which play specific roles in those transformations that have the import of consequence rules. These are words which appear on one side of a transformation only, i.e., phrases added or deleted by a transformation. Typically, there should be a number of negations (*no, not, few, neither, hardly, by no means, un-*, etc.), determiners, propositions, affixes (*-s, -ed, -en, re-*), auxiliary and modal verbs. But the list of grammatical constants may vary greatly depending on which rules are to be described and with what degree of generality.

Some writers (the present writer included, in 1964 and 1965 papers) consider paraphrase to be the basis for grammatical analysis. But consequence is more handy. Paraphrase is a rare phenomenon, if taken very strictly, and it admits degrees of accuracy which are hard to notice. It is similar to measurement in the physical sciences. We say that two things are of the same length, if measurements fell within the allowed accuracy. But for paraphrase the accuracy is difficult to grasp. We know that there are closer and looser paraphrases but we do not have a measure of the closeness. In principle, paraphrase is consequence both ways. But in practice we are often much more certain of the consequence one way than the consequence the other way around. From *It was at 5 o'clock that Jane arrived from Boston* follows the sentence *Jane arrived from Boston at 5 o'clock.* But the first sentence gives some other message. We expect that there is a problem as to when she arrived, or that she also arrived from a different direction at a different time.

One has to remember that consequences, and paraphrases, are rarely drawn from a single sentence. More often they are obtained from a sentence together with a set of assumed sentences (e.g. those which in a discourse appeared just before the sentence). Logicians are aware that two theorems may be equivalent, if there are some assumptions present, but not otherwise. For instance, Peirce's law (CCCpqpp) and (CCCpqrCqr) are inferentially equivalent if the two usual axioms of C-calculus are used : the syllogism (CCpqCCqrCpr) and the simplification (CpCqp). If one requires for a paraphrase that two different sentences are so related only if they are inferentially equivalent in all contexts, with any assumptions, then presumably there are no paraphrases. Two sentences nearly never satisfy so strong a requirement. In logic only sentences which lead to contradiction are so related. In a natural language even this is not certain. The conclusion that paraphrase is an unusual phenomenon which we reach in transformational grammar is quite parallel to a similar conclusion of distributionalist linguists (e.g., H. Hoeningswald, see his review of John Lyons' *Structural Semantics,* *Journal of Linguistics*, I, 1965), that synonymy is very rare. For, if you ask in what environment a given sentence can be used, then it will be clear that usually its paraphrases, even close ones, cannot occur in all the same environments. There are, therefore, three different criteria for paraphrase. According to a distributionalist, two sentences are paraphrases of each other if they can occur in exactly the same contexts. A logician may say that two sentences are related paraphrasically if they are inferentially equivalent no matter what other sentences are used for deduction. Finally, one may say — and this is the most useful concept linguistically — that two sentences are paraphrases of each other if, for whatever assumptions, they have the same consequences. It is probable that, in either understanding, paraphrase is an exceptional case, if it exists at all. Even so, people say "in their own words" what other people have said, and a linguist seems obligated to describe this phenomenon.

If one takes as the meaning of a sentence, the set of its consequences, then the meaning varies with the assumptions. Therefore, the meaning of a sentence α with respect to a set X of sentences is the set of consequences of X augmented by α, minus all the consequences of X alone. In other words the meaning of α is what α says that is new.

$$M(\{\alpha\}, X) = Cn(\{\alpha\} \cup X) - Cn(X)$$

When we repeat correctly, we approximate the meaning of what was said by a set of sentences whose set of consequences, with the common assumptions, is similar enough to the consequences of the original utterance.