The Speaker Simulation Device and the Foundations of Semantics and Pragmatics

Roland R. Hausser

(Universität München)

0. Introductory Remarks

This paper deals with the question: in what sence can logical formulas be regarded as an adequate analysis of meaning in natural language? This question is particularly relevant in the context of Montague grammar, where the meaning of natural language expressions is analysed in terms of translation into intensional logic. A sentence like *John reads a book.*, for example, is systematically translated into the logical formula

$$\cdot\ {}^{\circ}((\text{reads}^{\circ}(\text{a}^{\circ}(\text{book}^{\circ})))(\text{John}^{\circ}))$$

which is called the unanalyzed translation in Hausser (1980c, 1982). Unanalyzed word translations like $John^{\circ}$ are then replaced by analyzed translations of the same logical type, e.g. $\hat{P}P(j')$, resulting in a complex formula (called the analyzed translation) which is lambda-reducible into the so-called meaning^1-formula. In case of the above example, the meaning^1-formula systematiclly derived in the indicated manner is

$$\text{V}x[\text{book}'(x) \wedge \text{read}'(j',x)].$$

This formula consists of three kinds of logical symbols: operators like Vx or $\wedge$, variables like x, and logical constants like **book′**, **read′**, or **j′**. Thereby, the meaning of the logical operators and variables may be regarded as unproblematic, since it is clearly defined in terms of the usual meta-language definitions. These definitions are of an elementary nature, as attested by the fact that they may be easily operationalized (i.e. formulated as computer operations). The more problematic type of logical expressions are rather the so-called unanalyzed logical constants like **read′**, **book′**, or **j′**. In model-theoretic semantics in general, and Montague grammar in particular, the meaning of such unanalyzed constants is specified in terms of a *model-structure*, which defines the extension of a constant at each index or point of reference. This method is sufficient for maintaining a *well-defined* logical analysis, but there remains the question of what the model-theoretic definition of constants is supposed to mean intuitively.

Standard model-theory, as represented for example by Montague, proceeds on the assumption that the formal model-structure is a representation of actual as well as potential reality. This intuitive interpretation of the formal model-structure, however, leads to a number of basic problems, involving the treatment of non-literal uses (e.g. metaphoric, ironic, hyperbolic, etc. uses) of natural language, the nature of the lexicon and the treatment of word meanings, the interpretation of indexicals, etc. We will therefore propose alternatively to treat the model-structure as part of the speaker/hearer in contrast to the standard approach, which treats the speaker/hearer as part of the model-structure. The proposed departure from standard model-theory may seem unacceptable to those who see logic mainly as a tool for stating scientific truth and for this reason tend towards some kind of logical realism or nominalism. To those, on the other hand, who are interested in the formal analysis of meaning as part of an overall theory of communication, the alternative ontology of model theory advocated in this paper (and also in Hausser (1980a, 1979b, 1979c)) will seem like an adjustment to reality that has long been overdue.[1]

This paper consists of 8 sections. Section 1 adresses the question : What is a meaning? By comparing the speech-act approach and the model-theoretic approach, certain complementary limitations of these two theories of meaning are brought to light. This line of inquiry is continued in section 2, which describes specific problems of the standard model theoretic approach, concerning the treatment of lexical meaning and of context-dependency. Furthermore, the sentence semantic version of standard model-theory is compared with the so-called discourse semantic variant.

The presentation of our alternative approach begins in section 3, where it is proposed to treat the literal meaning of expressions in terms of model-theoretic synthesis in a partial model-structure. Section 4 presents the concept of a speaker simulation device or SID, of which model-theoretic synthesis is assumed to be a part in the sense that literal meanings of expressions are formally characterized by the model theoretic structures synthesized inside the SID. The synthesized model satisfying the formal translation of an expression is taken to be the *icon* of its literal meaning. The use of this icon relative to a context is operationalized in terms of matching the token-model with a second model, the so-called context-model, which is defined as a model-theoretic synthesis of what the speaker/hearer perceives and remembers.

The basic problem for model-theoretic semantics in general and model-theoretic synthesis in particular is the definition of the so-called elementary constants. In section 5, it is proposed to define the meaning of the most elementary constants in terms of matching certain sense data. Thus we arrive at a psychologically or neurologically or electronically based interpretation of model-theoretic semantics. Our proposal affects the ontological basis of model theory, but leaves most of the formal achievements of this tradition intact.

One notable feature of our approach is that we use two notions of truth, (i) the semantic truth-values, defined as ϕ and $\{\phi\}$, used in the logical definitions and (ii) pragmatic truth, defined in terms of the matching of the (logically constructed) token-model and context-model. In order to clarify this departure from traditional model-theory, we turn in section 6 and 7 to a discussion of Tarski semantics. Section 6 is of a preparatory nature, and discusses the notions 'T-condition', 'meta-language', 'object-language', 'sematically closed language', and 'essential richness of a language'. In the course of analysing these notions, a mistake in Tarski's construction of the liar paradox is brought to light, which leads to the conclusion that the essential property of a proper meta-language resides in certain well-defined categorial features of the language, and not in the notion of 'essential richness'. After a brief discussiou of the notion of an 'operational definition', we proceed in section 7 to the construction of a *formal* meta-language for the object-language of categorially analyzed English. It is shown that the infinite recursion of meta-languages inherent in Tarski semantics does not affect our formal analysis of meaning if the formal meta-language used to define meanings is a completely operationalized language.

In section 8, finally, we turn to the notion of truth defined in terms of *correspondence*. It is explained why this notion of truth appears in our theory as the *pragmatic* notion of truth. The paper concludes with a discussion of the phenomenon of vaguenss, which we treat on the level of pragmatics together with metaphoric, hyperbolic, etc. uses. The viability of this suggestion is demontrated by a resolution of the so-called Sorites paradox.

1. Meaning and Use

What is a meaning? Of the many answers that have been given to this question two general aroaches stand out, namely

(i) speech-act theory, as presented in various forms by Austin (1962), Grice (1957), Searle (1969), Wunderlich (1976), and others,

(ii) model-theoretic semantics, as developed by Tarski (1936), Carnap (1947), Kripke (1963), Montague (1974), and others.

Speech-act theory defines meaning as what the speaker intends, as what the speaker really meant when (s)he said something. This intentional approach to meaning is closely related to aspects of language use. In the following let us refer to meaning defined in terms of speech acts, rules of conversation, felicity conditions, or use conditions as the ***speaker meaning*** or ***meaning***^{2}.

Model-theoretic semantics, on the other hand, defines meaning as the objects, or sets of objects, to which the expressions refer, or which the expressions are said to denote. The paradigmatic case of this approach to meaning is the logical concept of a ***proper name***. For example, the meaning, or denotation, or referent of the proper name **John** is the actual person so named. A ***predicate*** like **walk**, furthermore, is said to denote a set of individuals, containing elements which have the property of walking. Frege (1892) completed the assignment of kinds of objects to the major parts of speech by proposing that *declarative* sentences should be defined to denote truth-values. This proposal developed into the view (Davidson 1967) that the meaning of a sentence may be equated with its truth-conditions. Recent developments in model-theoretic semantics, finally have led to quite detailed analyses of meaning in natural language by formally specifying the model-theoretic objects which serve as referents, either in terms of complex translations or in terms of meaning postulates (Montague 1974). In the following let us refer to meaning defined as a relation between surface expressions and model-theoretic objects as the ***literal meaning*** or ***meaning***^{1}.

Each of the two approaches to meaning mentioned in (i) and (ii) above captures a legitimate and important aspect of meaning in natural language. But unfortunately, in their present form the two approaches are persued in ways that render them incompatible. Speech-act theory has no account of how the literal meaning of an expression depends on its surface structure. The speaker meaning is furthermore claimed to represent the primary notion of meaning, so that all other accounts are derivative (Grice 1957). Model-theoretic semantics, on the other hand, while providing a highly developed technique to analyze the literal meaning of expressions, is in its present form unable to provide natural accommodations for the ***use-aspect*** of natural language.

Before we turn to the question of how to reinterpret the speech-act approach (meaning^2)

and the model-theoretic approach (meaning^1) in such a way as to make them compatible, let us consider how meaning^1 and meaning^2 should in general be related. Since meaning^1 is defined as the literal meang of expressions and meaning^2 is defined as what the speaker/hearer has in mind in a certain utterance situation it is reasonable to relate them in the following way:

(1) use of meaning^1 = meaning^2

In other words, by using a certain expression with a certain literal meaning (meaning^1) relative to a context we may achieve a communicative effect (meaning^2) which goes far beyond the literal meaning encoded in the token surface. In ironic use, for example, meaning^2 is usually directly contrary to the literal use of the meang^1.

The necessity to distinguish between meanng^1 and meaning^2 may also be illustrated in connection with the somewhat hackneyed example (2):

(2) Can you pass the salt?

Uttered at the dinner table, (2) is used as a request (normally) and the intended response is passing the salt. Uttered to someone disabled by disease or accident, on the other hand, (2) may be used as a *bona fide* question, and the intended response would be 'yes' or 'no'.

So does (2) have two meanings depending on the context? The answer is yes if 'meanings' in the preceding question is read as meaning^2. The answer is no, however, if 'meanings' is read as meaning^1. (2) has only one literal meaning, but this meaning may be *used* in many different ways in many different contexts, creating a whole spectrum of meanings^2.

The use-potential (Hausser 1978a) of an expression depends not only on its literal meaning but also on its form. For example, (3a) and (3b) are semantically equivalent (i.e. have the same literal meaning), but only (3a) can be used as an answer to (3c):

(3a) John read a book.

(3b) It was John who read a book.

(3c) What did John do?

Let us therefore replace formula (1) by the more refined version (4).

(4)

$$\begin{array}{ccccc} & \textit{syntax} & & & \\ & \vdots & & & \\ \textit{use} \text{ of} & \left\{\begin{array}{c} \textit{form} \\ | \\ \textit{meaning}^1 \end{array}\right\} & \text{relative to a } \textit{context} & = & \textit{meaning}^2 \\ \vdots & \vdots & & & \vdots \\ \textit{pragmatics} & \textit{semantics} & & & \textit{speech-act theory} \end{array}$$

The relation between form and meaning^1 indicated in (4) is the so-called surface-meaning^1 mapping (a notion discussed and analyzed in surface compositional syntax in

Hausser 1978b, 1980c, and 1982).

Our distinction between literal meaning (meaning^1) and the speaker meaning (meaning^2) is reminiscent of Austin's (1962) distinction between locutionary acts (meaning^1) and illocutionary acts (meaning^2). But there is an important difference : *while we define meaning^1 as a structural property of expressions, Austin defines it as an act by the speaker.* Thus the dichotomy between locutionary and illocutionary acts is for Austin a contradistinction on the same level, the level of acts. This has the consequence that Austin and his successors treat illocutionary for indicators as something that is either explicit or implicit in the *language.* Searle (1969, 1971), for example, treats a speech act as something consisting of a proposition, expressing the literal meaning, and an illocutionary force indicator, expressing the use aspect.

The question facing this approach, however, is: how is the illocutionary force indicator to be determined in any systematic fashion? The most widely accepted view is to identify the illocutionary force indicator with the performative clause, if present, and otherwise to derive the illocutionary force indicator from the syntactic mood of the sentence. Both methods are clearly insufficient and subject to innumerable counterexamples. For example, 'I promise to give you an F if you don't do your homework.' contains an explicit performative clause, but it is usually interpreted as a threat rather than a promise. And 'Can you pass the salt?' is normally interpreted as a request and not as a question.

The speech act approach fails to provide distinct analyses of (i) the structure of the expression with the structurally encoded literal meaning and (ii) the possible uses of this expression. Rather, it tries to account for the speaker meaning by treating the whole utterance as the basic unit of analysis. However, without a context-independent, structural analysis of the form and meaning of surface expressions we cannot study the use of such expressions relative to a context, much as we cannot study the functioning of a tool relative to a certaion object before we know the tool's exact shape, size, and material.

The approach of referential semantics, on the other hand, does provide an analysis of the form and the literal meaning of surface expressions. As shown by Montague, we may formally describe the literal meaning of expressions in a fragment of English in terms of translation into the model-theoretically interpretable language of intensional logic. Thus, given any linguistic expression in the surface fragment, we may characterize its literal meaning (meaning^1) in terms of the denotation conditions associated with its formal translation. But how can we get in this framework from a formal characterization of meaning[1] to a formal characterization of meaning[2]?

It is curious that the standard model-theoretic approach, as represented by Carnap, Kripke, and Montague (a) completely abstracts from the speaker/hearer and (b) provides no analysis of lexical meaning. Rather, the formal model is seen as a *representation of reality* (including potential reality), and the denotation conditions (truth conditions) are read as if it were the purpose of a formal interpretation to find out whether a formula is 1 (true) or 0 (false) relative to a model structure at an index. *In praxi*, however, the model structure is not independently given, but must be specified by the logician before (s)he can start with a formal interpretation of a formula. In as much as we may imagine different states of affairs, we may define the formal model structure as we see fit. Thus, the explicit specifications of a formal model is logically and empirically unrewarding on the standard approach. The sole purpose for actually defining a formal model structure would be to *illustrate* how a model-theoretic interpretation works (on the compositional or non-lexical level).

In summary, the speech act approach takes the *utterance* as the basic unit of linguistic description, while referential semantics usually takes the *expression* as the object of analysis. Our approach, as formulated in (1) and (4) may be seen as a synthesis of these two traditions in that we recognize both the expression with its form and literal meaning and the utterance of the expressions by the speaker. We differ from the speech act account in that we define literal meaning not as an utterance act (such as Searle's (1969) so-called propositional act), but as a property of expressions. And we differ from referential semantics in that we define denotations as set-theoretic concepts that may be used, and not as objects of real or possible worlds. This latter point will be explained in the following section.

2. Problems of the Standard Model Theoretic Approach

While we may define the formal model to represent any state of affairs we like, there are systematic restrictions on the definition of the model structure imposed by the meaning of the words of the language under interpretation (assuming the model structure is used to interpret a natural language). Compare for example (5) and (6):

(5) The red circle rises.

(6) The square circle rises.

Whereas one may define a model structure such that (5) is 1 (true) relative to one index and 0 (false) relative to another, intuition requires that there should be no index in the model structure relative to which (6) would be 1.[2)]

One way to treat the restrictions induced by the intuitive word meanings on the definition of the model structure is to exclude certain model structures from consideration. This is the *meaning postulate approach*, as used by Montague in PTQ (Montague 1974, chapter 8). Meaning postulates are external restrictions on model structures which delimit the class of what Montague calls 'logically possible' models. This terminology is somewhat misleading, however. What is at issue is not logical possibility but rather the speaker's intuitions regarding the semantic interdependence in the denotation of different words. For example, a model where the denotation of **man** does not overlap with the denotation of **human** would be no more linguistically reasonable than a model where the denotations of **square** and **round** are not disjoint.

While the method of meaning postulates permits to maintain that assumption of the standard approach according to which the model structure is viewed as a representation of reality and the denotation conditions are viewed as instructions to find out whether a sentence is 1 or 0 relative to an index, meaning postulates are an extremely cumbersome method for formally implementing lexical interdependencies (cf. Hausser 1979c for further discussion). This leads to the question: how could we separate the *lexical aspect* of word meaning from what we might call the referential aspect? This question is quite parallel to our earlier question of how to separate the description of literal meaning from the speaker meaning in model theoretic semantics.

The traditional model-theoretic approach, according to which meaning (denotation, reference) is stipulated to be a *direct relation* between expressions and model-theoretic objects not only eliminates the possibility for a well-defined lexical semantics in model-theoretic terms, but also raises serious *ontological problems*. If meaning is the relation between an expression and the object it refers to, must the object be real? If yes (and philosophers in the traditions of nominalism and realism decidedly think so), we are faced with the question of what to do with language expressions for which there simply are no real objects as possible referents. Take for example **the smallest prime number greater than 11**, **John's last hope**, but also expressions other than noun phrases such as **in**, **and**, **to**, etc. There are no real objects to which these expressions may be said to refer. Thus one either has to expand one's notion of what is real in order to give these expressions meanings (Meinong 1904), or one has to deny meanings to incomplete expressions, postulating that only complete sentences have a meaning by themselves (Russell 1905 ; cf. Hausser 1979c for a more extensive discussion of this issue).

Another problem with the traditional model-theoretic approach concerns the treatment of *context-dependent expressions* or indexicals. Compare for example (7) and (8):

(7) Bill saw Mary at the station.

(8) I saw you here.

In (7) the truth-value depends on the denotation of the constants **Bill**, **see**, **Mary**, and **at the station**, as specified by the model. In particular, **Bill**, **Mary**, and **at the station** are to be defined as denoting particular individuals and a particular place, respectively. In (8), however, the situation is quite different insofar as it would be intuitively wrong to assign fixed denotations to the indexicals **I**, **you**, and **here**.

One way to treat indexicals within the standard model-theoretic approach is the so-called *coordinates approach* (Montague 1974 chapter 3, Lewis 1972): in addition to the coordinates specifying a possible world and a moment of time, additional coordinates are defined for each context-dependency aspect to be treated. Lewis (1972), for example, uses a different coordinate for possible speakers (pronoun **I**), possible hearers (pronoun **you**), possible places (pronoun **here**), possible indicated objects (pronoun **this**), and even for possible previous discourse, respectively. In short, the coordinates approach permits to retain the assumption according to which meaning is a direct relation between expressios and referents by defining a context of use as extended point of reference.

The intuitive interpretation of the model structure as a representation of reality, however, suffers under the coordinate approach. Since the model structure is assumed to specify a state of affairs at an index, one would expect that this state of affairs *is* the context. Instead, the coordinates approach introduces a second kind of reference mechanism: while the denotation of regular constants is specified over the denotation function, the denotation of indexicals is specified over numerous additional context-coordinates. Furthermore, to define the context as an arbitray n-tuple of external coordinate fails to capture the highly specific interaction between context-dependent expressions and a coherent context (i.e. situation).[3]

Additional problems raised by the standard approach concern non-literal reference such as vague reference and metaphoric reference. Since the standard approach characterizes the meaning relation as a direct mapping between the expression and the state of affairs provided by the model (denotation), the only way to handle non-literal meaning assignments is to postulate ambiguities. In light of the fact that non-literal uses (i.e. metaphoric, ironic, etc. uses) of a word like **wood**, **table**, **fox**, etc. are as *unlimited* in their variety as the various use-situations in which they may be created by the speaker, the ambiguity treatment imposed by the standard model-theoretic approach is methodologically untenable. Our alternative proposal is to treat these variations as variations of meaning[2] (speaker meaning) rather than meaning[1] (literal meaning). Meaning[2] is described in terms of general principles of use (pragmatic principles). Thus no assumption of *semantic* ambiguity for handling metaphoric, ironic, etc. 'meanings' is to be assumed in our system.

So far, we represented referential semantics in the sense of Montague, that is, oriented towards the analysis of the literal meaning of surface expressions and essentially limited to sentence semantics. But there is another version of referential semantics, represented in various forms by Hintikka (1976), Stalnaker (1970, 1978), Groenendijk & Stokhof (1975), Karttunen & Peters (1978), and others, which we might call *discourse semantics.*

Discourse semantics is in many respects an attempt to treat the domain of speech act theory with model theoretic means. It is oriented toward the *utterance* and models in which an utterance may be understood in different ways. Hintikka (1976), for example, treats a question like 'Who came?' by assigning it a number of interpretations or "readings". One of these is 'Bring it about that I know that Vx [I know that x came]', another is 'Tell me whether Vx [you know that x came]', whereby the latter analysis models the situation of an *examination question,* These paraphrases, which appear as the deep structure of the utterance, are of a formal semantic nature and are regarded as readings of the utterance.[4)]

Another interest of discourse semantics is to model the context as a set of propositions and to study how inferences of expressions may vary in conjunction with different contexts (Stalnaker, Karttunen & Peters, Groenendijk & Stokhof). This approach is also called 'formal pragmatics'.

The question is now: can the discourse semantic or formal pragmatic version of model-theory avoid the difficulties of the sentence semantic approach? The problem of the sentence semantic approach is that meaning is defined as a direct relation between expressions and the model-theoretic 'reality', so that there is no room for the use-aspect of language. Discourse semantics, on the other hand, describes the whole utterance situation-or at least what are deemed the relevant aspects of it. Thus discourse semantics shares the problem of the speech-act approach: there is no separate structural description of the expressions used in the utterance. Consequently, there is no real analysis of the principles of use either. Modeling the utterance situation and the literal meaning of the expression used in terms of the same model, with a direct relation between elementary expressions and their model-theoretic referents, inevitably leads either to extremely standard utterance situations or contexts (thus preserving a resemblance of literal meaning), or to non-standard meaning assignments as soon as non-standard contexts are taken into account.[5)]

In summary, both the sentence semantic and the discourse semantic version of model theory suffers from the same old problem of the standard model theoretic approach, though in different form. This problem is what we have described as the fusion of the lexical and the referential aspect. It is caused by the assumption that 'meaning' should be defined as a direct relation between expressions and the model-theoretic 'reality'. After all, which 'meaning' is in the standard model-theoretic approach supposed to be constituted as a direct relation between the expression and the model-theoretic object, meaning[1] (as a structural property of expressions) or meaning[2] (which in cludes the use-aspect relative to the utterance situation)? It is the weakness of the standard model-theoretic approach that it cannot provide a clear answer to this question. Since sentence semantics has no room for pragmatics, defined as a coherent theory of use, and discourse semantics has no convincing account of the literal meaning of the expressions used in the utterances (which are regarded as the basic units of linguistic description according to the speech-act theory and its formal semantic applications), neither version of the standard model-theoretic approach can provide for a clear distinction between semantics and pragmatics.

Such a distinctions is indispensable, however. Every time we study the meaning of a word or sentence we must decide what to treat as part of the literal meaning and what in terms of use. If we rob the field of pragmatics of its legitimate regularities, we obstruct our ability to develop a viable theory of pragmatics. At the same time we obstruct our ability to arrive at a viable theory of semantics (overloading).

3. Definition of a Semantic Space as a Partial Model Structure

The difficulties of the standard model theoretic approach all stem from problems arising with the semantic treatment of *natural* language. For example, the need for providing interpretations to context-dependent expressions (indexicals) and the problems constituted by vague and metaphoric reference come from nattural language (but are easily avoided in artificial languages). Similarly, the need for a model-theoretic account of the lexical intuitions of the speaker/hearer come from natural language. This has led the representatives of the standard model theoretic approach (or formal semantics in general, e.g. Russell and Quine) to occasionally scoff at natural language as illogical or even beyond any hope of consistent logical analysis. The source of the problem, however, must be sought in the use-aspect of the speaker meaning in natural language.

Let us turn now to an alternative approach to model theory which preserves the formal and descriptive merits of formal semantics (as demonstrated in Montague grammar) while accommodating formula (4):

$$\text{(4) use of} \left\{ \begin{matrix} \textit{form} \\ | \\ \text{meaning}^1 \end{matrix} \right\} \text{relative to a context} = \text{meaning}^2$$

This new framework, first presented in Hausser (1979b), separates the lexical aspect of meaning from the referential aspect (cf. 2 above) by treating

(9i) *literal meaning* (meaning^1) in terms of model-theoretic *synthesis* in a semantic space representing the speaker/hearer's lexical intuitions;

(9ii) *context* in terms of a model-theoretic representation of what the speaker/hearer perceives and remembers in a given utterance situation;

(9iii) *reference* in terms of matching the synthesized literal meaning with the context. Reference is thereby treated as part of pragmatics.

Thus our alternative approach is based on the construction of two models, one representing the literal meaning of the token, the other representing the context. The former model is called the ***token-model***, the latter is called the ***context-model***. The speaker's *use* of a literal meaning (meaning^1) relative to a context is treated in our system as the *matching* of the two model theoretic structures. Thus pragmatics is sandwiched between the token model and the context-mode, inside the head of the speaker/hearer. The process of *reference* is regarded as part of pragmatics, while the construction of the token-and the context-model shares to a degree the goals of sentence- and discourse-semantics, respectively.

We arrive at the token-model by reinterpreting the intuitive role of the formal model-structure. Rather than treating the model-structure as a representation of reality and the denotation conditions as instructions to determine the truth value of formulas relative to an index, let us view the model structure as a representation of the semantic intuition of the speaker/hearer and the denotation conditions of a sentence token as instructions to synthesize or construct a token-model (or set of models) relative to which the sentence would be true. Thus the purpose of semantically interpreting an expression is not to determine its denotation relative to a model (in a model structure at an index) given in advance and regarded as a representation of reality (at that index), but rather to construct a denotation (or model) that would satisfy the expression and that is regarded as a *formal representation of its literal meaning* (meaning^1).

We assume that the synthesis of a token meaning is executed in a partially defined model structure, which is assumed to be part of a speaker simulation device (SID). What is required for the synthesis of a token meaning? While the logical operators like $\sim, \wedge, \lambda$, etc. in the translation of a token receive their meaning in terms of the denotation conditions associated with these operators (where the denotation conditions are specified in a metalanguage or in terms of certain operations), elementary logical constants are assigned their denotations by the partial model structure.

Our model structure is partial in the sense that it specifies the meaning of elementary constants in terms of their denotational *correlations* rather than in terms of explicit denotations. For example, the denotations of RED and BLUE are not characterized by stating explicitly which model-theoretic individuals are in the respective sets at each index. Instead, the set- or function-theoretic correlations of denotation sets is defined in terms of general conditions. As a rough illustration of the definition of the meaning-correlations in a partial model-structure consider the following clause:

$$\text{If } a_0 \in F\ (\text{RED})\ (i,j), \text{ then } a_0 \notin F(\text{BLUE})\ (i,j).$$

For the sake of simplicity, we assume here the general definition of the PTQ model structure @ (Montague 1974, chapter 8), where @ is defined as a quintuple $(A,I,J,\leq,F)$. Thus $a_0 \in A$, $i \in I$, $j \in J$, and F is the usual denotation function.

However, the set-theoretic correlations between constants of equal type like RED, BLUE, and COLOR as defined by the partial model structure constitutes only half the meaning analysis of elementary constants. The remaining task is to characterize precisely in what way, e.g., RED and BLUE differ in meaning. This will be done by defining the meaning of elementary constants in terms of matching certain kinds of sense data of the SID. More specifically, we assume that the SID contains a *semantic space*, consisting of elementary concepts. These elementary concepts are defined in terms of matching certain elementary sense data of the SID. Some of these elementary concepts are denoted by elementary constants like RED or BLUE. The set-theoretic interrelations among these elementary constants in turn is captured by the partial model-structure, which is regarded as a part of the semantic space.

The crucial difference of our approach and the PTQ model structure is that (i) the semantic space is used for the *synthesis* of models which are regarded as set- or function-theoretic representations of meaning[1], and (ii) the semantic space is not regarded as a representation of reality, but functions solely to encode the meaning relations among elementary constants. In other words, the partial model structure representing our semantic space has a purely *lexical* function. The referential aspect of meaning, on the other hand, will be handled as something quite different, as we will see in the following section.

Our notion of a partial model structure must be clearly distinguished from other uses of this term in the recent literature. In Friedman *et al* (1978) the model structure is conceived as a partially defined representation of reality, which means that as new expressions come up in the text, new denotations are defined in the model structure (assuming a computer simulation of a text interpretation). Thus, in order to interpret **John walks** at an index, a denotation is assigned to, e.g. **walk'**, if it has not been specified already. The evaluation of expressions relative to indices in the Friedman model structure is partial because of the limitations of the computer memory. Our model, on the other hand, is partial not because certain aspects of reality have not been filled in yet, but because the model structure specifies only the semantic *interrelations* of elementary constants. A completely specified model (or denotation) comes about in our framework only once the synthesis instructions associated with the logical operators present in the UL-formula of a token-translation have been executed.

Note that the notion of a *model* (and consequently the notion of a 'partial model') may be interpreted in two different ways. One is 'model of a world', the other is 'model of a sentence'. The first interpretation is connected with the concept of a model structure @. @ is defined by Montague in PTQ as (A, I, J, $\geqslant$, F). Such a model-structure specifies the denotation of each unanalyzed logical constant via F relative to each point of reference (i, j), where $i \in I$ and $j \in J$. These denotations are constituted by elements of the power set over A, where A is regarded as the set of entities of the model structure @. The denotations of all constants of the language at a given index are called a *model* (of the world at that index). Thus a model-structure consists of a set of models, the elements of which are associated with the index-structure of the model-structure.

The second interpretation of the term *model* is associated not with a model-theoretic index, but with a formula. Given a sentence p, we may talk about some or all models which satisfy p. Our approach of model-theoretic synthesis of the literal meaning of expressions clearly represents this second notion of a model.

However, models defined as satisfying sentences may be treated not only as synthesized structures, but also as part of a traditional model-structure. This latter possibility underlies the recent work on so-called partial models by Barwise (1981) and Kamp (1981). Barwise and Kamp assume complete model-structures in the traditional sense, intended to represent actual and potential reality. What they call partial models are model-theoretic structuses which satisfy sentences or discourses. These models are called partial because they disregard part of the assumed model-theoretic reality represented by the model-structure.

For Barwise, partial models are a device to accommodate the fact that we don't know what the external reality is really like. The goal of Barwise's approach is to test inferences of sentences relative to a model-theoretic reality which is partial in the sense that it satisfies only a limited set of discourse propositions, and not the whole of reality. Kamp's paper, on the other hand, despite some sweeping claims in the beginning, does not go beyond a rather limited treatment of *anaphoric*[6] pronouns. A sentence containing pronouns is satisfied by a model which is partial in the sense that the interpretation of the pronouns is not yet determined.

The point here is to observe the distinction between *models* (partial models) and *model-structures* (partial model structures). Barwise and Kamp deal with partial models, assuming traditional model-structures. Our synthesized models, on the other hand, are not considered partial. Rather, they are synthesized on the basis of a partial model-structure. The model-structure is partial in the sense that it models only the intuitive semantic interrelations among the basic constants of the logical translation language. As will be explained in the following sections, the partial model-structure is part of the semantic space of the SID, whereby the elementary *constants* of the model-structure are defined in terms of *concepts* of the semantic space. In other words, the meaning of an elementary constant is characterized (i) in terms of the set- or function-theoretic interrelations of different constants, as specified by the partial model structure, and (ii) in terms of the concepts denoted by the constants, whereby the concepts are defined in terms of matching certain kinds of sense data of the SID. For example. the denotation of **red** and **blue** is defined as disjunct sets, which are sub-sets of the denotation of **color** (partial model structure). The specific meaning of **red** as opposed to **blue**, however, is defined in terms of matching a particular kind of sense data as perceived by the SID (concepts of the semantic space).

4. The Speaker Simulation Device (SID)

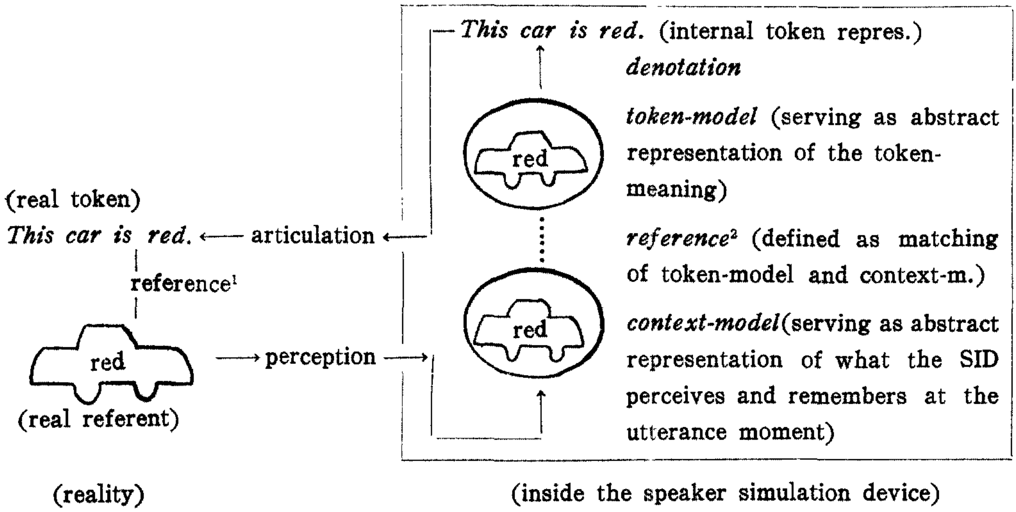
The switch from the "verifying mode" to the "synthesizing mode" in the interpretation of model theory removes the ontological problems of the standard approach outlined in section 2, while still providing for a semantic characterization of the elementary constants (which constitute the smallest units of lexical analysis in our system). Our new form of *lexical* model-theoretic semantics (the grammatical aspect of which is described in Hausser 1981, 1982) is clearly compatible with Montague's sentence semantics (e.g. PTQ, EFL, or UG). All that is changed by our switch from verifying to synthesizing is the interpretation of the *denotation assignment to the elementary constants* in the elementary translation formulas.

However, the characterization of literal meaning in terms of model-theoretic synthesis is only part of a complete meaning^2-analysis in natural language. What we have to attend to now is that aspect of meaning^2 which is constituted by the *use* of an expression by the speaker relative to an utterance context. Furthermore, in order to satisfy the needs and purpose of traditional language philosophy, we must somehow reestablish the connection between expressions and reality, which was severed when we reinterpreted the formal model-structure as a semantic space. The question then is: how do synthesized models relate to reality?

As already indicated in (9), we complement the synthesized token meaning in our sytem with a synthesized *context*. This formal context is regarded as a model-theoretic representation of what the speaker/hearer perceives and remembers at the moment of a token interpretation. The context is the speaker's subjective representation of his momentary *reality*. And the use of an expression relative to the context is formally treated in terms of matching the token-model (which is a set-theoretic representations of the literal meaning of the token) with the context-model (which is a set-theoretic representation of the context). That we use set-theoretically based model theory for these representations, rather than net-work semantic or procedural semantic systems, is at the present point mainly a matter of convenience. One may argue, however, that set theory is the most elementary and most general form of semantics and that net-work and procedural semantic analyses may be translated into set-theoretic representations. Unfortunately, very little is known so far as to how the three types of systems compare (cf. Anderson (1976), p. 231 ff., where the net-work grammar ACT is translated into first order predicate calculus).

The interaction of the token-model and the context-model inside the SID and the reality external to the SID may be schematically indicated as in(10), on the next page. (10) pictures the SID in that kind of speech-act situation which is taken as the paradigmatic case by the standard approach. That is a situation with a declarative expression (i.e. "This car is red.") and a state of affairs containing a 'real' referent (i.e. a car) and a 'real' property (i.e. red) such that there is a *correspondence* between the expression (the real token in (10)) and the model (which is identified with a real situation). The basic goal of the standard approach is to capture the Aristotelian notion

(10)



of truth, which is defined as a corrspondence between what is said and what is (Tarski (1944)).

What is described by the standard approach is indicated in (10) as the relation between the 'real token' and the 'real referent'. Since the standard approach defines meaning as a direct relation between expressions and model structures, it operates with a single, uniform notion of truth. Our alternative approach, on the other hand, takes the relation between the external expression and reality apart into several sub-mappings by routing it through the speaker (SID). This results in two notions of truth: *semantic truth*, which is a purely abstract notion and characterizes the relation between surface expressions and the synthesized models regarded as their formal meaning[1], and *pragmatic truth*, which characterizes the relation (correspondence) between the token-model and the context-model.

We need these two notions of truth in order to explain non-literal use in natural language. For example, if the weather is terrible and someone says: "That's real nice weather today." the speaker uses a certain literal meaning (formally expressed by a synthesized model that makes the translation of the sentence true, in the semantic sense of truth) to say something pragmatically true - if properly understood as ironic use. While our alternative approach may be called more complicated than the standard approach in that we postulate two notions of truth and distinguish between the literal meaning of expressions and their use by a speaker relative to an utterance context, our approach to model-theoretic semantics can treat a number of phenomena the standard approach is not naturally equipped to handle. These phenomena include all instances of non-literal use (i.e. metaphoric, ironic, hyperbolic, etc. use), propositional attitudes (Hausser 1979c), non-declarative sentence moods (Hausser 1980b), context-dependent expressions (Hausser 1979b), and the lexical analysis of word meaning (Hausser 1981, 1982a); but also traditional paradoxes such as the Sorites paradox to be analysed in section 8 below.

The strategy in contemporary model-theoretic semantics (represented for example by Barwise, Kamp, and Stalnaker) is to enrich the standard theory to handle some of the above natural language phenomena without changing the basic assumption of the standard approach (which is schematically indicated in (11a) below). But the standard approach as conceived by Carnap and Quine was never really intended for the analysis of natural language. Rather, its purpose was the representation of scientific truth (e.g. physics). The purpose of our alternative approach, on the other hand, is to provide a formal framework (i.e. the SID) to analyse *natural tanguage communication.* If we make additional assumptions as to the accuracy of perception and the literal interpretation of sentences, the SID-approach is suited as well as the standard approach to model 'scientific truth'. If these additional assumptions are dropped, on the other hand, our approach extends naturally into a general theory of communication.

It is the purpose of the SID-approach to preserve and extend the formal achievements of model-theoretic semantics in the analysis of natural language, as it culminates with Montague. Specifically, we maintain the definition of the classical logical operators, though we tend more towards operational definitions than toward the usual meta-language definitions. The reason is twofold:

(i) The words in the meta-language definitions of some theories are often quite loose, whereas operational definitions (which will be discussed in section 6) seem to be more objective and of unquestionable precision.

(ii) Operations are more in tune with our idea of synthesizing models which are intended to serve as the icon of the literal meaning of surface expessions.

As explained in section 3 above, the real change brought about by our proposal concerns the definition of the so-called 'unanalyzed logical constants'. Instead of fixing their meaning some pre-existing model-theoretic reality, we specify the meaning of elementary constants in terms of matching sense data. Assuming the efforts in Artificial Intelligence to electronically simulate perception are successful, there is the obvious possibility to operationalize the definition of all the elementary constants of the logical language, too. The SID is the framework of a model-theortically interpreted logic *all symbols of which* are defined in terms of operational definitions.

While the change brought about by our approach maintains most of the formal semantic analysis of, e.g., PTQ it amounts to a new paradigm as far as the model-theoretic ontology is concerned.

(11a) PARADIGM I

surface expression
| denotation=reference
model-theoretic reality

(standard approach)

(11b) PARADIGM II

surface expression
| denotation
synthesized model
⋮ use (reference)
synthesized context
|
outside reality—— ——

(our alternative approach)

The standard approach is based on a single, unified notion of meaning (fusing the literal meaning of expressions and the speaker meaning) and a single, unified notion of truth (fusing semantic truth and pragmatic truth) by describing the speech act situation from the view point of an omniscient outside observer (god?), who provides the situations relative to which the truth-value of formulas may be determined. Our alternative approach, on the other hand, describes the speech-act situation from the view point of the *speaker/hearer*. We thus distinguish between the outside reality, on the one hand, and how it is perceived by the speaker, on the other. While the standard approach is interested solely in modeling valid inferences, our alternative approach aims at modeling the general principles of communication. However, our approach may be narrowed to modeling valid inferences only, if certain highly restrictive assumptions regarding the nature of pragmatics, or language use, are adopted.

What is the most general principle of communication? In sections 1 and 2 we concluded that the most fundamental principle of natural language communication is that the literal meaning of expressions is used by the speaker. We formulated this principle in the formulas (1) and (4). This presupposes, of course, that natural language expressions have literal meanings, an assumption some linguists do not share (cf. section 8 for further discussion). However, that a certain literal meaning can be put to different uses in varying utterance situations, thus creating different speaker meanings (meaning^2) can be observed at the earliest stages of language acquistion.

For example, a child can use the word "bow wow" to mean^2 different things, such as expressed by the following paraphrases: "I am afraid of the dog.", "Look, there is a dog!", "I want the dog.", and even "Play dog for me!". The standard approach would have to assign different lexical readings for each of these interpretations of "bow wow". We, on the other hand, assume (i) a literal reading of "bow wow" corresponding to the concept or icon **dog** and (ii) pragmatic principles for using this literal meaning relative to a given utterance context.

The terms *meaning*[1], *literal meaning*, *concept*, and *icon* — as associated with certain language symbols — are different names for the same thing, expressing different aspects. The most neutral term is *meaning*[1], but depending on the subject of discussion, the terms *literal meaning*, *concept*, and *icon* will be used. For example, concepts such as (denoted by the words) **red** or **four-legged**[7)] are not only present in man, but also in lower animals. The specifc language ability of man consists in the *fixing of certain concepts to certain symbols* — thus creating icons. When we talk about the development of literal meanings in the process of language acquisition we talk about the development of complex concepts in their attachment to certain symbols.

For example, in language acquisition the literal meaning of "bow wow" in the sense of **dog** is preceded by an earlier stage where the child calls everything from a cow to a cat a bow wow. The abstraction behind this (or rather the consequent application of an abstract concept or icon) is no different from the behaviour of an adult, who calls different fruits (apples), some red, some green, some yellow, some large, some small, and a lot of them rotten, all "apples". The generalzed (and thus "incorrect") use of "bow wow" in the earlier phase is simply due to an icon that has not yet been differentiated to the degree found in the later phase.

The pragmatic principles guiding the use of expressions relative to a context are formalized in our approach in terms of matching the token-model and the context-model.[8)] The most basic (or most straightforward) type of use is the so-called *literal use*. The literal use of a meaning[1] relative to a context is represented in our approach as a complete match between the two models (i.e. the token model is properly embedded in the context-model).

It is the literal use which the standard approach takes as the paradigmatic case, fusing literal meaning (meaning[1]) of expressions and literal use by treating as a direct relation between expressions and the model-theoretic reality. The expression "This car is red" in (10), for example, is evaluated as *true* by the standard approach, whereas our alternative approach evaluates it as "pragmatically true under the literal use interpretation". In short, our alternative approach captures as a *special* case both the Aristotelian notion of truth underlying the standard approach (cf. Tarski 1944) and the prototype of utterance situation (literal use) analysed by the standard approach.

The consequences of our alternative approach, as represented by the SID, may be summarized as follows:

(12i) Since the literal meaning of the 'token representation' in the SID is characterized in terms of a *synthesized* model, where the basic sets A, I, and J (cf. Montague 1974, chapter 8) of the semantic space cannot possibly contain any real objects, but must be interpreted as consisting of purely abstract memory spaces in the SID, the *ontological objections* justly raised against the standard approach of model-theoretic semantics do not apply (cf. 5. for further discussion).

(12ii) Since we distinguish between *denotation* (i.e. the relation between the token-representation and its synthesized meaning) and *reference* (i.e. part of the relation between the token-model and the context-model), semantics and pragmatics are effectively separated.

(12iii) By reinterpreting the model-structure as a semantic space, which assigns partially defined denotations to the elementary constants in the elementary translation of tokens, we provide the basis for a viable theory of lexical meaning.

(12iv) At the same time we create the need, and the room, for a *coherent notion of context*, defined as a model theoretic representation of what the speaker/hearer perceives and remembers at an utterance moment under consideration.

(12v) By distinguishing between the formal context and reality, we are able to describe cases of *perception or memory error*. See Hausser (1979c) where such a case, namely the 'man with the martini'-example of Donnellan (1966), is analyzed.

(12vi) By distinguishing between the real token and the token representaion in the SID, we are adle to describe cases of *acoustic misunderstanding* and *high-level speech errors* (e.g. spoonerisms).

5. Sense Data and the Operational Definition of Elementary Constants

As indicated in (10) and (11) in section 4 above, our approach distinguishes between the *external reality* and the *internal context*. The external reality is what the universe is really like at a given moment, and as such a largely unknown arrangement of various kinds of facts. It is surely an important philosophical and scientific problem to find out what the external reality is. But for the purpose of analyzing communication in natural language it is sufficient to represent only the internal context, i.e. what the speaker/hearer knows or believes at the utterance or interpretation moment, whereby we treat the difference between knowledge and believe simply in terms of different degrees of subjective certainty (which is in turn dependent on the believe-systems fashionable in the society of which the speaker/hearer is a part).[9)]

The assumption by the standard approach to model-theoretic semantics, according to which the model-structure is regarded as a representation of the external reality is not only linguistically unsound because of the fusion of meaning^1 and meaning^2, but also philosophically unpractical. The obvious reason is that the external reality presents an infinity of facts, some known, some unknown, some present, some past, and some still in the future. If the model structure is assumed to represent the external reality, it must represent an actual infinity of facts. Indeed, the icardinality of this infinity is likely to be aleph one. For example, assuming that the real numbers are part of the totality of facts (an assumption at the essence of the so-called realist position), the individuals of the model-structure must be assumed to constitute a recursively enumerable set.

On the SID-approach, on the other hand, the number of elements is definitely *finite*, just as the cells or neurons in a human brain. But there are algorithms that allow to compute arbitrarily large numbers, just as in a pocket calculator (though in a pocket calculator the size of the numbers is limited by the number of digits on the display).

The assumption of an aleph-zero infinity of facts in a traditional model structure means that such a model structure can not be implemented on a computer, only finite sections of it. This shows that the ontology of standard model theory is overly idealistic. The assumption of an aleph-one infinity of facts, on the other hand, cannot even be described by any of the logical languages (where no limitation is imposed on the number and the size of sentences), since logical languages can only describe a countable infinity of facts.

But even if we assume a finite representation on the standard model theroretic approach, where would one begin with the description of the external reality? Since the model-structure is supposed to represent the *true* state of reality (just considering the real world, leaving possible worlds aside for the moment), one would presumably start with those facts which are felt to be most firmly established, that is, one would start with the most elementary (and most advanced) principles of modern physics, chemistry, biology, etc. in order to treat as many facts as possible as systematically derived from an assumed initial state. This tendency to adopt scientific concepts for the definition of such terms as gold, water, tiger, etc. is quite evident in, e.g. Kripke (1972) or Putnam (1975). While questions concerning the true nature of gold or other features of the universe are legitimate and important in philosophy, they contribute very little to the mysteries of human communication.

On the other hand, with what kind of facts should we start when constructing the SID-internal model structure called semantic space? Here the answer is not based on the current notion of which facts are most securely established (e.g. physics, biology, astronomy, etc.), but rather on the anthropomorph structure of the SID. Just as the external reality is assumed to be basically physical (as witnessed by the efforts of the so-called double aspect theory to correlate 'private sensations' such as pain to measurable brain states), the SID-internal semantic space (as well as the context-model and the token model realized in it) is essentially psychological or neurological in nature. Therefore, the natural starting point to describe the semantic space is the *sense data* of the SID.

Our SID-approach may seem "unrealistic" or "solipsistic", etc. to someone raised in the present philosophical main stream, but it is by no means a new or exotic point of view. Rather, it has been articulated throughout the course of Western philosphy and dominated the so-called English empiricism of Hobbes, Locke, Brekely, Hume, and later Mills. But as much as we have in common with these authors and these authors have in common with each other, as many differences exist, often of a sharp and principled nature. While it would be very interesting to compare in detail the various empiricist and sensualist theories with the SID-approach, this cannot be accomplished in the limited space of the present paper.

In order to avoid certain misunderstandings, however, let me emphasize that our approach recognizes the external reality of the world as much as the SID-internal reality of "ideas", or rather perceptions, concepts, and algorithms. When we construct the SID we can clearly see that there is an inside as well as an outside mental reality of the SID and the external reality which is perceived by the SID. The view that all the "knowledge" of an individual is ultimately based on its sense data and mental operations is in no way incompatible with the assumption of the existence of an external reality. Consider the following passage from Locke's "Essay concerning human understanding" (1690):

> "...simple ideas are not fictions of our fancies, but the natural and regular productions of things without us, really operating upon us; and so carry with them all the conformity which is intended: or which our state requires: for they represent to us things under those appearances which they are fitted to produce in us: whereby we are enabled to distinguish the sorts of particular substances, to discern the states they are in, and so to take them for our necessities, and apply them to our uses."[10]

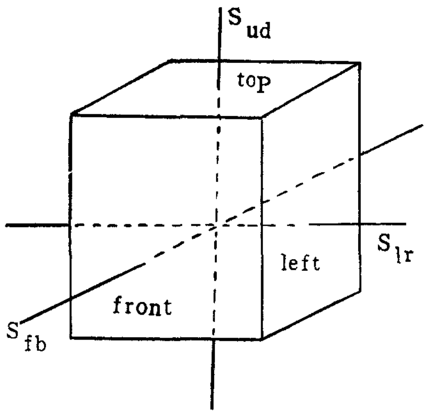
After these more general considerations let us turn now to the main theme of this section, namely the structure of the SID-internal semantic space in which both the token-models and the context-models are synthesized. Since the SID is a theoretical framework designed to simulate human communication, we assume that the SID has an anthropomorph structure. In particular, we assume that the semantic space of the SID is based on the following ***constituting parameters***:

(i) a time axis, defined in terms of an internal oscillator (e.g. the pulse or heartbeat of the SID or some other 'internal clock'), and

(ii) a three-dimensional space, defined in terms of the three-dimensional body structure of the SID, whereby the center of the SID is identified with the center of the spatial coordinate system of the internal context.

The 'now-moment' or 'present state' of the SID is called the ***zero-index*** and defined as the quadruple (t_0, s_{lr}, s_{fb}, s_{ud}), where t_0 is the present moment, s_{lr} is the left-right axis, s_{fb} is the from-back axis, and s_{ud} is the up-down axis of the SID, as illustrated in (13):

(13)



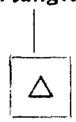
It is in terms of these parameters that the formal meaning of elementary constants corresponding to our notions of ***in front of***, ***behind***, ***above***, ***below***, ***to the left of***, ***to the right of***, and other elementary spatial notions are to be defined. Once these elementary constants are equipped with the indicated formal meanings, they serve two functions within the semantic space of the SID: (i) they are used non-verbally in the construction of the context-model, which reflects what the SID perceives and remembers at a given zero-index, and (ii) they are used verbally in that they occur in the elementary translations of certain surface expressions and thus function in the constructions of token-models.

In addition to the constituting parameters of time and space, the SID is equipped with a number of ***sensory input parameters***, such as ***vision*** ("the special sense by which the qualities of an object (such as color, luminosity, shape, and size) constituting its appearance are perceived and which is mediated by the eye"), *hearing* ("the special sense by which noises and tones are perceived as stimuli"), *feeling* ("the one of the basic pysical senses of which the skin contains the chief end organs and of which the sensation of touch and temperature are characteristic"), *smelling* ("the special sense concerned with the perception of odor or scent through stimuli affecting the olfactory nerves"), *tasting* ("the one of the special senses that perceives and distinguishes sweet, sour, bitter, or salty of a dissolved substance and is mediated by the taste buds on the tongue").[11] It is on the basis of these sense data that the elementary constants corresponding to our notions of color, such as **red**, **green**, or **blue**, of shades such as **light** or **dark**, of shapes such as **round**, **triangular**, **rectangular**, of size, such as **large** or **small**, of sounds, such as **loud**, **high**, or **low**, of feeling, such as **hard**, **soft**, **hot**, or **cold**, of smelling such as **fresh** or **foul**, and **tasting**, such as **sweet**, **sour**, **bitter**, or **salty**, receive their semantic characterization.

In addition to external perception mentioned above, there is also internal perception such as hunger and thirst. Furthemore, besides perception Locke recognizes another mode of mental activity, which he calls *volition*. It is obvious that much more could (and should) be said on the different kinds of mental activity, but for reasons of time and space we must concentrate on a few principled issues concerning the use and interpretation of language. Anticipating some of the discussion in section 8, let us consider the verbal and the non-verbal use of mathematical concepts like triangle or hexagonal within the semantic space of the SID.

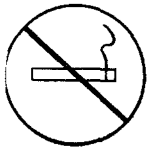
In its iconic use, the elementary concept *triangle* occurs as the denotation of the elementary *constant* **triangle**, which in turn occurs in the translation of the surface word **triangle**. Speaking in a simplified manner, we may imagine the icon as a little picture representing the concept:

triangle′

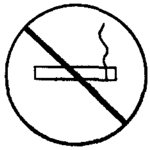


There are also instances where the literal meaning of a complex surface expression can be represented adequately by a pictorial icon, as illustrated in (ii):

(ii) surface: *no smoking cigarette*

synthesized model: 

The advantage of this kind of represention of the iconic content is that such pictures are immediately interpretable by normal participants of our culture. A somewhat confusing property from a theoretical point of view, however, is that pictorial representations of iconic content have a special property: the symbol and its iconic content (or literal meaning) may be treated as identical, as indicated in (iii):

(iii) surface = synthesized model: 

In (iii), the synthesized model is itself used as the surface.

In what sense are cases like (iii) *special* (in comparison to symbols of written or spoken language)? They constitute an exception to de Saussure's principle of the 'arbitrariness of the sign'. But once we take this special property of pictures into account, they are well suited to explain what is meant by 'using a literal meaning relative to a context in order to achieve a certain speaker meaning (meaning^2)'. Note, for example, that the iconic content of (ii) or (iii) is normally *not* used literally. The sign is intended to prohibit not only the smoking of cigarettes, but of cigars and pipes as well. In fact, it is often intended to prohibit any kind of open fire.

The pragmatic principles (principles of using symbols) that lead from the restricted iconic content (meaning^1) of *no smoking cigarettte* to the speaker meaning (meaning^2) paraphrasable as *the smoking of cigarettes, cigars, pipes, as well as the presence of any other kind of open fire is prohibited in the environment of this sign* are of the same nature as the pragmatic principles that get us from the literal meaning of *Can you pass the salt?* (meaning^1) to the speaker meaning (meaning^2) paraphrasable as *I request that you pass the salt to me.*

If this is true, why don't we use pictures like (i-iii) to represent the iconic content of natural language expressions, instead of set-theoretically defined synthesized models? The reason is simply that pictures constitute a rather limited code, whereas the set- or function-theoretic structures of synthesized models are practically limitless in their expressive power. Pictures are not suited, for example, to represent any non-visual notions such as sounds, tastes, or feelings like **pain** or **hope**. And even in a case like (i) above, the picture representation is incapable to account for such important categorial (and semantic) distinctions as between **triangle** and **triangular**. Pictures are unsuited, furthermore, to express the meaning of functional terms like **and**, **or**, or even **not**. This holds at least for the initial, naive stage of a picture language. Once the picture symbolism is enriched with adstract symbols (such as the representation of **no** in (ii) and (iii)!), on the other hand, we move away from the immediately obvious representation of iconic content towards a hieroglyphic code, which obeys de Saussure's principle of the arbitrariness of symbols (regarding the meaning[1] denoted or encoded) to a larger and larger degree.

It is important to note, that we used the immediate interpretability of certain picture symbols only to illustrate that literal meanings (= iconic content), which is of a rare clarity in pictorial icons, may be *used*, thus resulting in a new type of meaning, called the speaker meaning. This does not mean, however, that the representation of iconic content of natural language expressions must in some way be based on immediately interpretable icons. Rather, the requirement on a formal representation of literal meaning (such as in form of synthesized models) is that it be based on *elementary* operations and concepts (such as x is element of the set S) of a most general nature. The structures of set theory may or may not be regarded as immediately interpretable icons, but of relevance is only the fact that these structures are easily definable and explainable in terms of a natural meta-language and may be easily operationalized on a computer. Further advantages of using model-theoretic structures (rather than pictures) for encoding the literal meaning of natural surface expressions are that in this way (i) the composition of meaning in the surface expressions is formally described, and (ii) the logical properties of the surface expressions, such as entailments, are accounted for.

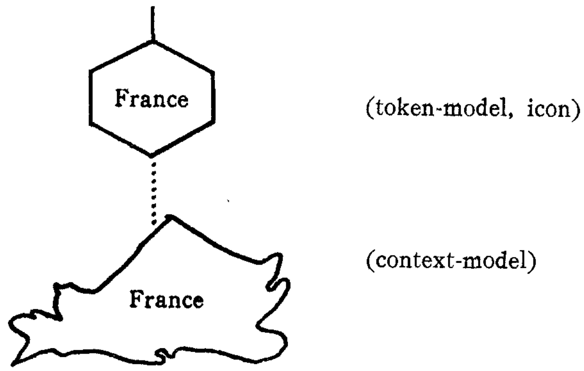
A further question arising with the literal meaning of a word like **triangle** is whether the icon denoted by it represents a particular triangle or something that is universal to all triangles. The pictorial representation in (i) above would suggest the former interpretation, thus giving cause to the objection (raised by Berkely against Locke) that the concept of a triangle should cover all kinds of triangles (scalene, isoceles,etc.) of different sizes. Thus the pictorial representation of the elementary concept *triangle* is as problematic philosophically and linguistically as it is conceptually simple.

A more complicated and at the same time more adequate way to represent the icon *triangle* would be in form of algorithms or operations which allow to transform a triangle into all kinds of trianglar shapes and sizes, as familiar from computer displays. We may take this algorithmic construction as the icon of *triangle′* just as the function denoted by $\hat{P}_1\hat{P}_0\hat{P}[P_0(P)\wedge P_1(P)]$ is taken as the ic on of the surface word **and** T $\downarrow$ (T $\uparrow$ T). If we abopt this second way of representing the icon of a triangle, we may surely say that the SID has a general sense of the mathematical concept in question.

Let us turn now to the non-iconic use of concepts like triangle or hexagon, i.e. the use of these concepts in the construction of context-models within the semantic space. In addition to the 'perfect' triangles provided by the algorithmic construction (designed specifically for iconic purposes). the semantic space contains all kinds of other, roughly triangular shapes, just like a television screen designed to match and interpret electro-magnetic signals may be said to "contain" all the shapes representable on it. Thus of the shapes used in the semantic space, only a subset serves both in the construction of token-models and the of context-models. The imperfect, e.g., triangles (like those without perfectly straight lines), on the other hand, are used only in the construction of context models.

Consider now the famous example by J.L. Austin: "France is hexagonal." The point of our distinction between literal meaning, context, and use is that in order to make this sentence true (false) the shape of France must not be included (excluded) among the 'meanings' of hexagonal. Rather, we define the literal meaning of ***hexagonal*** in terms of a perfect icon. The interpretation (and evaluation) of the sentence is then done in terms of matching the icon with the context-model, as indicated below:

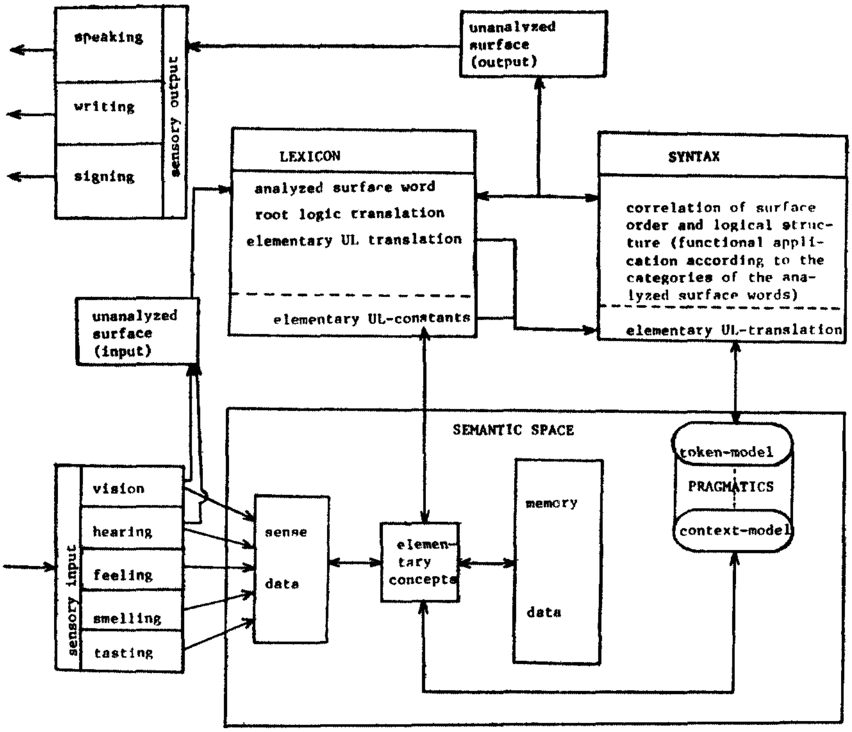
France is hexagonal.



If we are liberal in our evaluation of the correspondence between the two models, the sentence will be taken as pragmatically true (sort of), otherwise as false (see sections 7 and 8 for a discussion of semantic and pragmatic truth, as well as further examples).

After this very brief discussion of mathematical concepts like triangle (which according to Descartes are based on some kind of inner vision of internal perception) and their use, let us return to elementary concepts based on external perception. the interaction between the sense data, the semantic space, the lexicon, the token-model and the context-model inside the SID is schematically represented in (14):

(14)



In (14), as in (10), two kinds of recognition are distinguished: *symbol recognition*, represented by the loop from vision/hearing over the unanalyzed surface and the lexicon to the token model, and *context-recognition*, represented by the loop from vision/ hearing/etc. to the context-model. Thereby the token-model and the context-model are synthesized in the same semantic space.

Let us consider now a situation as in (10), where the SID is in the hearer state, it sees a situation with a red car, hears the token 'This car is red.', and interprets the token relative to the perceived context. In the first place, how does the SID recognize the context? In order to recognize the car as such, the SID must have learned to classify this type of object. That is, in its semantic space there must already exist certain concepts which can be applied to match the visual input pattern. The concept which matches best is taken as what the SID recognizes, and constitutes part of the momentary context-model. Similarily with the property *red*: the visual organ of the SID registers that the external object in focus reflects light at a frequency corresponding to red in the electromagnetic spectrum (roughly speaking). In order for the SID to be conscious that the car is red and not, say, purple, the semantic space of the SID must have formed a concept which is defined to match the particular interval of the colorspectrum which is called red.

The relation between the sensory input and the semantic space is reciprocal. The *meaning* of an elementary constant like RED is ultimately defined as a certain type of visual input (beyond the function theoretic interrelations between constants defined in the semantic space). The particular type of visual input, on the other hand, becomes *conscious* only in terms of the presence of a concept which is defined to match it. Two people can have the same sensory input, but recognize completely different things. Take for example a satellite picture of some part of the earth, which is studied by two people, one trained to recognize missile silos, the other an at least equally intelligent layman. On the level of the sense data, both perceive precisely the same array of colored dots, but whereas the layman sees nothing but colored dots, the trained expert recognizes missile sites.

Thus the difference in interpretation is not due to different sense data in the two individuals (though this is a possibility too, e.g. in a case of color blindness), but due to a difference in the higher order concepts developed by the two individuals to interpret this kind of sense data. Generally speaking, the concepts in the semantic space of the SID may be divided into four types, namely

(i) elementary concepts defined in terms of matching elementary sense data (e.g. RED);

(ii) elementary concepts defined in terms of elementary logical operations (e.g. logical negation);

(iii) complex concepts based on elementary sense data concepts (e.g. CAR);

(iv) complex concepts of a purely logical nature (e.g. TRIANGLE).

We assume that only elementary concepts (i.e. type (i) and (ii)) are innate, whereas complex concepts are acquired. Thus differences in non-iconic (i.e., contextual) as well as verbal interpretation always arise in connection with the individual build-up of complex concepts.

For technical reasons, we assume that the concepts of the semantic space are denoted by formulas of the universal logic UL,[12] whereby elementary concepts are denoted by elementary formulas and complex concepts are denoted by complex formulas of UL. Consider for example the root **kill′** (which has achieved a paradigmatic status within semantic decomposition as it originated with McCawley 1968). **kill′** is defined as the abbreviation of a complex UL-formula, which, for the sake of illustration, may be stated as follows:

$$\hat{m}\hat{n}[\mathrm{CAUSE}(m,\ \hat{Q}Vz[Q(z)=\mathrm{BECOME}(n,\ \hat{P}[P(n)=\sim\mathrm{ALIVE}(n)])])]$$

This formula denotes a complex concept, built up from the elementary concepts denoted by the elementary constants CAUSE, BECOME, ALIVE (which we regard for the sake of the argument as sense data concepts) and from logical operators. The point is that the meaning of a complex concept is not defined in terms of wholesale matching with a complex sensory input, but rather compositionally. Thus, the SID may never have seen anything corresponding to **kill′**, yet be able to recognize the process when seen for the first time. All that is necessary for the SID in order to understand the meaning of **kill′** is to understand the logical operators and constants of the complex UL-formula which **kill′** is defined to abbreviate. This complex UL-formula may pe acquired in a purely verbal way, rather than through an initial experience (either in real life or on television). Thus, when the SID is for the first time in a situation where it recognizes that an individual has the property ALIVE at point t and the property$\sim$ALIVE at point t' (where furthermore some agent is involved in the transition), the SID will infer that what was going on must have been a *killing*. This inference is based directly on the (verbally acquired) meaning structure of the complex concept denoted by **kill′**.

Whether a level of UL-formulas denoting the context-models needs to be assumed, or whether the synthesized context-models suffice without a level of UL-formulas, is a question which may at some point acquire some principled importance. At present, however, we assume the level of UL-formulas simply as a convenient way to formally describe the context-models. It may turn out to be practical, furthermore, to store context-situations in the memory of the SID in form of UL-formulas rather than as complete context-models. When a certain memory is recalled, the denotation of the respective formula may then be synthesized again in the semantic space.

After this brief discussion of the context-recognition loop indicated in (14), let us turn to the symbol recognition loop. Symbols are normally perceived by means of vision or hearing. However, other senses may be trained to recognize symbols, such as feeling (touch), which is utilized in reading braille, a system of writing for the blind, that uses raised dots. The first step of symbol recognition is to extract the ***unanalyzed surface*** from the sensory input. We assume that the unanalyzed surface forms of words (e.g. *John*, *reads*, *a*, etc.) are stored in the lexicon. Recognition of an unanalyzed surface form consists in matching suitable surface forms provided by the lexicon with the sense data suspected of being symbols. This process is similar to the context-recognition where elementary sense data are matched with elementary concepts provided by the semantic space.

The next step of symbol recognition is to relate the unanalyzed surface to correponding analyzed surfaces. In a 'vertically refined' lexicon (cf. Hausser 1981), each unanalyzed surface word (e.g. **bank**) is related to all its so-called homonymical variants, i.e. molecules with a surface identical to the unanalyzed surface in question. By replacing the unanalyzed surface words of a given token by homonymical variants a number of molecule sequences results. Those of the molecule sequences which may be combined into well-formed orthogonal trees represent a reading to the unanalyzed susface.

Once a token has been correlated to an orthogonal tree, the interpretation runs as described in Hausser (1982). That is, the unanalyzed translation of the orthogonal tree is replaced by an analyzed translation, which is then lambda-reduced to a meaning^1-formula. The roots in the meaning^1-formula are finally replaced by corresponding UL-formulas. The process that leads from an unanalyzed surface to a UL-formula expressing its literal meaning (or one of its literal meanings) is thus based (i) on two lexical replacements, namely (a) the unanalyzed surface/molecule replacement (where the molecule is a homonymical variant of the unanalyzed surface) and (b) the root/UL-formula replacement, (ii) on the orthogonal composition of molecules, and (iii) on the lambda-reduction of analyzed translations. Each of these steps leading from an unanalzyed surface to a corresponding UL-formula is logically well-defined, whereby the lexical replacements are constrained by the type-structure inherent in the system.

Let us assume now that an external token, e.g. **This car is red.** has been recognized by the SID and been translated into a UL-formula, e.g. $\vee x[car'(x) \wedge RED(x)]$. This formula consists of two types of symbols, (i) logical symbols, and (ii) elementary constants (the unanalyzed root **car'** would, of course, have to be further decomposed). Now, the meaning assignment to this UL-formula is completely traditional as far as the logical symbols are concerned: they are defined in terms of the usual meta-language definitions (which will be scrutinized in sections 6 and 7)). The only difference with the standard model-theoretic approach concerns the definition of the constants: rather than fixing the denotation of, e.g., RED in a standard model structure (by specifying the extension at every index), we assume an *operational definition* of this constant in terms of matching sense data.

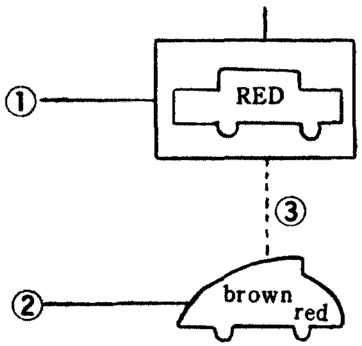
The concept denoted by the UL-constant RED is thus defined just like the constants building up the context model. But whereas in the context-model various different RED-concepts may occur (in order to accommodate various perceptions in the neighborhood of red), the UL-translation of *tokens* uses only one special *prototype* of RED. This prototype of RED is a concept defined like those other RED-concepts in the semantic space, but with the additional property that it doubles as an *icon* in the sense that it is denoted by an elementary constant occuring in a token translation. Which of the RED-concepts of the semantic space is chosen to double as an icon depends on what is regarded as prototypically red by the SID and its speech community.[13)]

The above considerations are once more summarized in the schematic SID-interpretation

(15):

(15)

This car is red. SYNTAX



SEMANTICS Token model, representing the literal meaning of token, consisting of prototypical caricon and prototypical red-icon.

USE

CONTEXT Context-model, representing visual input of the SID, analyzed on higher level as some kind of car with some kind of red color.

On the whole, our SID-based theory of communication presumes three types of matching: (1) the matching of sense data which helps building up the context model, (2) the matching of sense data and symbols, which is called symbol recognition if the SID is in the hearer state and articulation if the SID is in the speaker state, and (3) the matching of token model and context-model, which is called meaning^2-interpretation if the SID is in the hearer state and meaning^1-verbalization if the SID is in the speaker state.

These three types of matching are of different degrees of theoretical interest. Whereas (1) and (2) may be regarded as more or less technical problems within artificial intelligence, (3) constitutes the theoretical field of pragmatics. Ultimately, however, all the parts of the SID-construction interact and the particular implementation of one component will have consequences on the other components with which it interacts.

The by far the best developed component of the SID is the mapping from the SID-internal token representation to the translation formula of intensional logic representing the literal meaning of the token. The reason is that this mapping is essentially covered by the formalism called Montague grammar. However, Montague grammar as originally conceived does not distinguish between literal meaning and use, or between semantics and pragmatics. Rather, it employs model-theoretic semantics in the traditional way criticized in section 1 and 2 above. Let us discuss therefore in the following section in what sense the semantic essentials of Montague grammar are preserved in our framework. We will do this by analyzing the principles of Tarski-semantics, on which Montague grammar is based. We will be especially interested in the role of semantic truth and the object-/meta-language distinction within the framework of the SID.

6. Object-Language, Meta-Language and Tarski's T-Condition

Before we turn to the rather formal discussion of so-called Tarski-semantics, let us summarize the main points developed so far. The concern of this paper is the analysis of meaning as transmitted by means of natural language symbols. Our first move was to separate the literal meaning structurally encoded in natural surface expressions (meaning^1) and the speaker meanig defined as the use of an expression relative to a context (meaning^2). The theoretical consequence from this hypothesis was our proposal to (i) analyze and represent the literal meaning of expressions in terms of abstract, set- or function-theoretic structures, called token-models which are regarded as icons, and (ii) to analyze and represent the context in terms of model-structures reflecting what the speaker perceives and remembers, and (iii) to formalize the theory of use (pragmatics) in terms of matching the token-model with the context-model.

The element of use in our meaning-analysis requires that there be a language user, for which reason it is only natural to locate the token- and the context-models inside a speaker/hearer, called the SID. This move has numerous advantages, especially with regards to the definition of context-variables (cf. Hausser 1979b), the semantic analysis of space and time (cf.section 5), and the definition of elementary constants in terms of matching sense data.

In order to reassure those logically minded readers, who find the content of the previous pages too far away from the beaten path of mathematical logic as established in the last ninety years, I would like to show in the present and the following section that our analysis of meaning^1 in terms of an ontologically reinterpreted Montague grammar not only obeys the principles of Tarski semantics (Tarski 1936, 1944), but constitutes a refined version of it in the sense we analyze expressions of the object-language English in terms of a *strictly formal meta-language*, defined as an extension of intensional logic. On the other hand, our reananalysis of Tarski will also bring to light some deep seated differences of outlook, the discussion of which is hoped to be beneficial in that it clarifies the specific properties of the two approaches.

While Tarki aims at defining truth (via reference to the meanings of a meta-language), our interest will be the definition of meaning (via reference to truth). However, due to the interrelatedness of meaning and truth in declarative sentences, this difference of interest or emphasis will not affect our discussion of the basic questions at issue, at least not initially. Consider (16) and (17) below, where two standard instances of

Tarski's so-called T-condition are given:

(16) ⌜Schnee ist weiss.⌝ is true if, and only if, snow is white.

(17) ⌜Snow is white.⌝ is ture if, and only if, snow is white.

In these clauses, the meaning of an object-language expression (which is German in (16) and English in (17)) is defined in terms of a so-called meta-language (which happens to be English in both (16) and (17)). The crucial point of the T-condition is that the object-language expression has a double status, expressed in terms of a double categorization. On the level of the object-language *Schnee ist weiss.*$_t$ or *Snow is white.*$_t$ is of category t, i.e. the expressions in question are sentences. On the level of the meta-language, on the other hand, the object-language expressions are treated as unanalyzed names, whereby a meta-language name of an object-language expression X is written as ⌜X⌝. Thus on the level of the meta-language, the expressions ⌜Schnee ist weiss.$_t$⌝$_e$ or ⌜Snow is white.$_t$⌝$_e$ are of category e, where e is the category of names.

Intuitively, (16) characterizes the meaning of an object-language expression in terms of translation, while (17) is a case of paraphrase (though of a tautological nature). Yet both (16) and (17) are instances of the same general schema (18) of the T-condition:

(18) X is true if, and only if, p.

How is this to be explained? Tarski specifies that a meta-language must contain (i) names of the object-language sentences, i.e. X in (18), (ii) the predicate *is true* (which we will define as a function from meta-language e-expressions to meta-language t-expressions), (iii) the operator *if and only if* (which we will define as the traditional logical operator $\leftrightarrow$), and (iv) the object-language sentences, i.e. p in (18).

Taking the relation between X and p as well as the categorial nature of X into account, we may replace (18) by the more explicit version (19):

(19) ⌜S_t⌝$_e$ is true if, and only if, S_t.

(19) represents the "paraphrase-version" of (18), which is exemplified by (17). Another possibility to define a meta-language is to treat the p in (18) not as identical with the sentence named by X, but rather as a translation of the sentence named by X. This possibility, which is mentioned but not further persued by Tarski, is formalized in (20).

(20) ⌜S_t⌝$_e$ is true if, S'_t.

(20) represents the "translation-version" of (18), which is exemplified by (16). In (20) we find the meta-language name ⌜S_t⌝$_e$ of the object-language sentence S and the standard translation S'_t of the object-language sentence S. Note that the notion of a

standard translation has a precise meaning within Montague grammar, which provides a formal algorithm for translating analysed surface expressions of English into formulas of intensional logic.

The purpose of Tarski in stating his theory is to avoid certain paradoxes, such as the 'liar paradox'. Let us see now whether our categorially refined versions (19) and (20) of the T-condition (18) are adequate in the sense that the construal of the pertinent paradoxes is not possible. After all, Tarski shows that so-called semantically closed languages (which are inconsistent) obey the T-condition as stated in (18) just as a proper (that is, consistent) meta-language. We proceed by quoting that passage of Tarski (1944), where he states his particular version of the liar paradox. In addition to Tarski's original text (printed in italic letters) I will add some comments (stated in normal print). The numbering of Tarki's examples will be adjusted to the present context.

To obtain this antinomy in a perspicuous form, consider the following senteence:

(21) The sentence printed in this paper on p. 347, l. 31, is not true.

For the sake of clarity, let us assume that the sentence referred to in (21) consists of the words *Paris is the capital of England.* However, nothing hinges on this assumption. *For brevity we shall replace the sentence just stated* (i.e. (21)) *by the letter, 's'.*

According to our convention concerning the adequate usage of the term "true" (i.e. (18)), *we assert the following equivalence of the form (T):*

(22) 's' is true if, and only if, the sentence printed in this paper on p. 347, l. 31, is not true.

On the other hand, keeping in mind the meaning of the symbol 's', we establish empirically the following fact:

(23) 's' is identical with the sentence printed in this paper on p. 347, l. 31.

Here we must pause for a moment. It may be hard to believe, but the move which is supposed to lead to clause (23) is not correct. Remember that Tarski uses 's' to abbreviate the sentence "The sentence printed in this paper on p. 347, *l*.31, is not true." Therefore, keeping in mind the meaning of the symbol 's', (23) cannot be empirically established. Rather, all that can be established is (23′):

(23′) 's' is identical with 'The sentence printed in this paper on p. 341, 1. 31, is not true.'

Now, by the familiar law from the theory of identity (Leibniz law) it follows from (23) (though certainly not from (23′)) *that we may replace in (22) the expression "the sentence printed in this paper on p. 347, l. 31" by "'s'". We thus obtain what follows:*

(24) *'s' is true if, and only if, 's' is not true.*

In this way we have arrived at an obvious contradiction.

Unfortunately, Tarski's conclusion is based on the incorrect identity clause (23). If we apply the correct identity as stated in (23′), we obtain (24′):

(24′) $\ulcorner s_t \urcorner_e$ is true if, and only if, s_t.

And (24′), in contrast to (24), is not contradictory. Rather, (24) constitutes the condition T in its standard (and categorially proper) form. This result is rather confusing. We set out to show that in a meta-language defined according to the categorially refined version (19) of the T-condition (18) certain paradoxes cannot be construed. But to our embarrasment, we found that Tarski made a mistake in construing his paradox which has nothing to do with the pertinent categorial distinctions between the meta- and the object-language level.

At this point, I would like to emphasize that I find myself in complete agreement with the general thrust of Tarski's theory. Furthermore, I take it as established that there are other examples, involving self-reference or impredicativity, which do lead to genuine inconsistencies if the distinction between the meta-language and the object-language advocated by Tarski is not obeyed. What the crucial formal properties of a proper (i.e. consistent) meta-language should be, however, is unfortunately still unclear, because Tarski is rather vague on this point. For him, a meta-language is consistent if it is "essentially richer" than the object language. Yet there is no strict definition of the notion "essential richness", only the guess or suggestion that a meta-language is essentially richer than the odject-language if the meta-language contains higher variables than the object-language. Yet, how the presence of higher variables in the meta-language is supposed to insure consistency (i.e. prevent construal of paradoxes) is not explained.

In order to see the crucial difference between a semantically closed language (which is inconsistent) and a proper meta-language let us have a look at Tarski's definition of a semantically closed language.

The inconsistency of semantically closed languages. If we now analyse the assumptions which led to the antınomy of the liar (see above for our annotated quotation of the passage in question), *we notice the following*:

(I) We have implicitly assumed that the language in which in the antinomy is constructed contains, in addition to its expressions, also the names of these expressions, as well as semantic terms such as the term "true" referring to sentences of this language; we have also assumed that atl sentences which determine the adequate usage of this term can be asserted in the language. A language with these properties will be

called "semantically closed".

(II) We have assumed that in this language the ordinary laws of logic hold.…
Since every language which satisfies both of these assumptions is inconsistent, we must reject at least one of them.

It would be superflous to stress here the consequences of rejecting the assumption (II), that is, changing our logic (suposing this were possible) even in its more elementary and fundamental parts. We thus consider only the possibility of rejecting assumption (II). Accordingly, we decide not to use any language which is semantically closed in the sense given.

Doesn't a semantically closed language as defined (I) look very much like a regular meta-language? Indeed, but there is one crucial difference of a strictly categorial nature. Note that it says in (I) "...semantic terms such as the term "true" referring to sentences (!) of this language...". In a proper meta-language, the predicate *is true* never refers to a meta-language sentence, but only to meta-language names of a certain kind. Therefore, I conclude that the crucial property of a proper (i.e. consistent) meta-language resides in the fact that the predicate *is true* is defined as a function from meta-language names to meta-language sentences, whereas in semantically closed language, the predicate *is true* is characteristically defined as a function from sentences to sentences. A moment of reflection will hopefully suffice to convince the initiated reader that it is this categorial feature of a proper meta-language which prevents the kind of illegitimate substitution which would otherwise render the paradoxes in question. The property of "essential richness", on the other hand, whatever it may be, is of no consequence for the definition of a proper meta-language.

Before we turn next to the definition of the formal meta-language for our object-language of categorially analyzed English let us first discuss the purpose of such a construction. Up to now, meta-language and object-language are used in the following constellation: the object-language is a formal language, e.g. predicate calculus, whereby the meaning of the elementary symbols and the rules for combining symbols are stated in terms of an informal meta-language, usually the native language of the logician. As an example of such a definition consider (25), which gives the semantics of the standard operator '$\wedge$' of propositional calculus :

(25) $\ulcorner A \wedge B_t \urcorner_e$ is true if, and only if, $\ulcorner A_t \urcorner_e$ is true and $\ulcorner B_t \urcorner_e$ is true.

While (25) corresponds more or less to the form of the original T-condition (18), there are also meta-language definitions like (26):

(26) $\ulcorner x^2 \urcorner$ denotes a function with the natural numbers as domain and range which

takes the square of x (or x·x) as its value, for any natural number x.

Though (26) does not mention truth, it is as good a defintion as (25) because it uses only meta-language terms with a clear mathematical meaning. Putting it in another way: (25) and (26) are equally good meta-language definitions of object-language functors because they can be equally well operationalized (in form of computer algorithms).

(27) and (28) given below, on the other hand, are examples of bad meta-language definitions:

(27) $\ulcorner \mathbf{red}'(x)_t \urcorner_e$ is true (relative to a model @ and an index i, j, g) if, and only if, $\ulcorner \mathbf{x}_e \urcorner_e$ has the property of being red (in that model at that index).

(28) $\ulcorner f(x) \urcorner$ takes a motor cycle as value if, and only if, x is a large potatoe.

The point is that a mota-language definition is only as good as the notions used in it. (27) and (28) are bad definitions because they cannot be operationalized. Though (27) can be made formally explicit by listing the extension of **red′** at each index of the model, this would contribute nothing to analysis of the intuitive word meaning **red**.

We have thus arrived again at our criticism concerning the meaning assignment to elementary constants as customary in standard model theoretic semantics (e.g. PTQ), though from a different angle. Our earlier criticism in section 2 concerned the fusion of the lexical and the referential aspect of meaning, resulting in the inability to characterize the use aspect of natural language in referential semantics and the inability to characterize the literal meaning of expressions in discourse semantics. The present discussion complements this criticism by pointing out that the meaning assignment to elementary constants in standard referntial semantics (as exemplified in (27)) is not suceptible to a non-trivial operationalizing. Our alternative approach based on the SID, on the other hand, defines the meaning of elementary constants in terms of matching sense date (see section 5 above). It thus clearly aims at operationalizing the meaning assignment to elementary constants, even though substantial advances in the field of artificial intelligence (concerning the electronic simulation of perception) will be required.

Definitions like (28) are even worse than those like (27). Unfortunately, definitions like (28) are not as uncommon as one would like to think. The literature is full of undefined functions which take fantastic worlds as their domain and no less fantastic heavens as their value, or vice versa.

The above considerations regarding good versus bad meta-language definitions of object-language expressions provide us with one good reason for constructing a *formal* meta-language, rather than using an informal one. The value of using a formal meta-language is that in this way we insure that only clear (i.e. well-defined) notions are used in the meta-language definition of object-language expressions. Of course, the formal meta-language needs itself to be defined in terms of a known language (e.g. the native language of the logician). But if we require of our formal-language that it can be completely operationalized, no reference needs to be made any more to the native intuitions.

Note, incidentally, that the notion of operational definitions is not new, though sometimes misunderstood. Consider the following pertinent passage by Putnam:

> Attempts in the physical sciences to *literally* (Putnam's emphasis, R.H.) specify operational definitions for terms have notoriously failed; and there is no reason the attempt should succeed in linguistics when it failed in physics. Sometimes Quine's arguments against the possibility of a theory of meaning seem to reduce to the demand for operational definitions in linguistics; when this is the case the arguments should be ignored. (Putnam 1975, p. 251)

Putnam's inference from physics to linguistics may appeal to the untutored emotions of some, but it is void of argument or reason. In any case, that there exist successful operational definitions in logic (and thus in logically based linguistics) is witnessed by any old cash register. Since Putnam's remark on Quine is not supported by a reference (to make sure Quine's arguments will be properly ignored?), I will not speculate here about Quine's motivation for demanding operational definitions. Suffice it to say that operational definitions are essential for our overall approach to meaning. The most basic reason is that once the surface interpretation system, comprising a formalized natural surface language (object language) and an operationalized meta-language has been implemented as part of the SID, the infinite recursion of meta-languages inherent in Tarski's system will be of no further consequence for the processing of meaning by the SID. That is, the SID's may continue to communicate with each other even if the native meta-meta-language used to build them (and specifically to define their formal meta-language) is suddenly forgotten and extinct.

But what is the operationalized meta-language of the SID? So far, the surface interpretation system of the SID consists of no more and no less than systematic translation of categorially analyzed English into intensional logic. These formal achievements of Montague grammar, however, provide the main ingredients for constructing a formal meta-language for categorially analyzed English, as will be shown in the following section.

7. Definition of a Formal Meta-Language

In this section we will define a formal meta-language for a fragment of natural language, namely the categorially analyzed surface expressions of English as generated by the PTQ-surface grammar or the system presented in Hausser (1980b, 1981c, 1982). But before we can go ahead with the definition, we must decide whether the formal meta-language should be modelled after the paraphrase- or the translation-version of Tarski's T-condition. Consider once more the categorially refined versions (19) and (20) of the T-condition (18).

(19) *Paraphrase-version*

$\ulcorner S_t \urcorner_e$ is true if, and only if, S_t.

(20) *Translation-version*

$\ulcorner S_t \urcorner_e$ is true if, and only if, S'_t.

Thereby, S_t is a sentence of the object-language. Furthermore, in (19) S_t is also a sentence of the meta-language. And S'_t in (20) is the standard translation of S_t into the meta-language.

As far as the avoidance of paradoxes is concerned, (19) and (20) do equally well. This is due to the categorial nature of the terms in these clauses, as was explained in the previous section 6 in our comparison of proper (i.e. consistent) meta-languages and (inconsistent) semantically closed languages. In as much as both definitions (19) and (20) refer to truth, on the other hand, they are equally limited. After all, truth-values are only one type of possible denotation, and ultimately we need a meta-language which permits the definition of meaning not only of sentential operators (e.g. (26)), but of other types of functions as well (e.g. (27)). For the moment, however, let us postpone the impending generalization of the truth-condition to a denotation condition (which is presupposed in the analysis of syntactic mood in Hausser 1978a, for example) until after the construction of our formal meta-language.

As it is our purpose to provide a formal meta-language for the analysis of the literal meaning of analyzed English surface structures, let us define our meta-language according to the translation version (20) of the T-condition rather than the paraphrase version (19). The reason for this move is that most of the ingredients for such a formal meta-language are already provided by Montague grammar. They are (i) the object-language E_{CAT} of categorially analysed English, as defined in PTQ or Hausser (1980b, 1981c, 1982), (ii) the translation language IL of intensional logic, as defined in PTQ or Hausser (1978b, 1979c), and (iii) the formal algorithm for translating the object-language into the intensional logic, as defined in the systems mentioned in (i). If we extend the formal translation language IL of intensional logic by adding (a) a formal truth-predicate and (b) the names of the well-formed odject-language expressions of E_{CAT}, then we arrive quite naturally at the formal meta-language IL^{meta} of E_{CAT}, which is formally defined below.

(29) *Definition of IL^{meta}*

The basic expressions of IL^{meta} :

(i) If α is an IL-constant of type A, then α is an IL^{meta} constant of type A.

(ii) If x is an IL-variable of type A, then x is an IL^{meta} variable of type A.

(iii) If S is a wellformed expression of E_{CAT} (where E_{CAT} is the categorial surface grammar of English) of category B, then $\ulcorner S_B \urcorner$ is a basic constant of type e of IL^{meta}.

The rules of IL^{meta} :

(iva) If R^{syn} is a syntactic rule of IL, then R^{syn} is a syntactic rule of IL^{meta} (with the possible change that each occurrence of ME_A ($A \in TYPE$) is replaced in each rule $ME_A^{IL-meta}$).

(ivb) If R^{sem} is a semantic rule of IL, then R^{sem} is a semantic rule of IL^{meta} (with the possible change that each occurrence of ME_A ($A \in TYPE$) is replaced in each rule by $ME_A^{IL-meta}$).

(va) If $\ulcorner S_t \urcorner \in ME_e^{IL-meta}$, then T $\ulcorner S_t \urcorner \in ME_t^{IL-meta}$.

(vb) If $\ulcorner S_t \urcorner \in ME_e^{IL-meta}$, then $T\ulcorner S_t \urcorner @^{,i,j,g}$ is 1 if, and onlyi f, $S' @^{,i,j,g}$ is 1, whereby S′ is the standard IL-translation of S.

The operator T defined in (v) is read as *is true*. The definition of the truth-operator in (vb) is compatible with the PTQ-system, where the logical recursion is mostly defined over extensions. See (v*) below, where a strictly intensional definition of this operator is given, which would be compatible with the intensional logic defined in Hausser (1978b, 1979b).

(va*) If $\ulcorner S_t \urcorner \in ME_{s,e}^{IL-meta}$, then $T\ulcorner S_t \urcorner \in ME_{s,t}^{IL-meta}$.

(vb*) If $\ulcorner S_t \urcorner \in ME_{\langle s,e \rangle}^{IL-meta}$, then $T\ulcorner S_t \urcorner @^{,i,j,g}$ is that function from $I \times J$ to $D_{t,A,I,J}$ such that $T\ulcorner S_t \lrcorner$ is 1 if, and only if, $S' @^{,i,j,g}$ is 1, where S′ is the standard IL-translation of S.

In a strictly intensional meta-language, which uses clause (v*) rather than (v), clause (iii) has to be adjusted in the sense that $\ulcorner S_B \urcorner$ is a constant of type $\langle s,e \rangle$ rather than e.

We may prove now that any IL^{meta} statement instantiating the T-condition (20) is a tautology (as desired). Consider for example (30):

(30) $T\ulcorner$John walks.$_t\urcorner_e \leftrightarrow$ walk′ (j′)

(30) is like (20), except that *is true* is replaced by the operator T defined in (29v) above, S_t of (20) is instantiated as *John walks.$_t$*, *if, and only if* is replaced by $\leftrightarrow$ (defined as in IL), S'_t is instantiated as *walk′(j′)*.

(31) *Proof that* (30) *is an IL^{meta} tautology*

$T\ulcorner$John walks.$_t\urcorner_e \leftrightarrow$walk′(j′)$@^{,i,j,g}$ is 1 if, and only if, if $T\ulcorner$John walks.$_t\urcorner_e @^{,i,j,g}$ is 1, then walk′(j′) $@^{,i,j,g}$ is 1 and if $T\ulcorner$John walks.$_t\urcorner_e$ $@^{,i,j,g}$ is 0, then walk′(j′) $@^{,i,j,g}$ is 0. $T\ulcorner$John walks.$_t\urcorner_e$ $@^{,i,j,g}$ is 1 if, and only if, walk′(j′) $@^{,i,j,g}$, i.e. the standard IL-translation of *Jhon walks.$_t$* is 1. And $T\ulcorner$John walks.$_t\urcorner_e$ $@^{,i,j,g}$ is 0 if, and only if, walk′(j′)$@^{,i,j,g}$, i.e. the standard IL-translation of *John walks.$_t$*, is 0.

Q.E.D.

Strictly speaking, we must use the tree (or indexed bracketing) corresponding to *John walks.$_t$* in (30) and (31), because S_t in $T\ulcorner S\urcorner_t$ in (29iii) is defined as an *analyzed* (and thus disambiguated) surface expression of E_{CAT}.

It should be quite obvious now how to generalize the T-operator defined in (29v) to a denotation- or D-operator. Consider (32).

(32) *Definition of the D-operator*

(va^D) If $\ulcorner S_B\urcorner \in ME_e^{IL-meta}$, then $D\ulcorner S_B\urcorner \in ME_{B'}^{IL-meta}$, where B′ is the type corresponding to E_{CAT}-category B.

(vb^D) If $\ulcorner S_B\urcorner \in ME_e^{IL-meta}$, then D $\ulcorner S_B\urcorner @^{,i,j,g}$ is epuivalent with $S'_{B'} @^{,i,j,g}$, where $S'_{B'}$ is the standard IL-translation of S_B.

Let us call a meta-language which uses the D-operator instead of the T-operator IL^{D-meta}. Here we have again the choice between the PTQ and a strictly intensional version. The definition of the strictly intensional version of the D-operator is analogous to (v*) and left as an exercise to the reader.

Based on definition (32) we may now generalize the T-condition (20) into a D-condition. we repeat T-condition (20) for convenience.

(20) $\ulcorner S_t\urcorner_e$ is true if, and only if S'_t.

In IL^{meta} (20) has the form stated in (33):

(33) $T\ulcorner S_t\urcorner_e \leftrightarrow S'_t$.

The D-condition corresponding to (20) is given in (34).

(34) $\ulcorner S_B\urcorner_e$ has the same denotation as $S'_{B'}$.

In IL^{Dmeta} (34) has the form stated in (35):

(35) $D\ulcorner S_B\urcorner_{B'}=S'_{B'}$.

Again, any instantiation of (35) is an IL^{D-meta} tautology. The proof is analogous to (31). IL^{D-meta} is a formal meta-language which allows to formalize not only definitions like (25) (which characterizes a sentential operator in terms of truth), but also definitions such as (26) (which characterizes a function with the natural numbers as domain and range, assuming we have an operational definition of the sqaring function, formalized in the meta-language IL^{D-meta}).

Let us consider now in what respects our approach to meaning and truth differs from that of Tarski (1944). Formally, IL^{meta} as defined in (29) above is a straightforward implementation of Tarski's T-condition (18). One difference with Tarski, however, is that we implemented (18) in form of the translation version (20), rather than the paraphrase version (19). But this does not affect Tarski's goal, which was a characterization of truth (relative to a meta-language) and the elimination of the liar's paradox. We found a mistake in Tarski's construction of this paradox, but as far as the categorial structure of IL^{meta} is concerned (specifically the domain-range structure of the T-operator), we are in complete agreement with Tarski.

A more pronounced departure from Tarski is constituted by the definition of IL^{D-meta} (cf. (32)), even though IL^{D-meta} is technically a completely straightforward generalization of IL^{meta} : just as the truth-predicate T is a special case of the denotation-predicate D, IL^{meta} is a special case of IL^{D-meta}. Philosophically, however, the step from IL^{D-meta} to IL^{D-meta} is considerable, because it turns the attention away from truth (which was Tarski's concern) towards meaning.

We do not regard this as a disadvantage of IL^{D-meta}. After all, already in section 4 the heretic proposal was made to employ two notions of truth, namely semantic truth (used to characterize the literal meaning of certain types of expressions, namely those corresponding to category t) and pragmatic truth. We motivated these two notions of truth in terms of the necessity to distinguish between the literal meaning of expressions ($meaning^1$) and the speaker meaning ($meaning^2$), defined as the use of expressions relative to a context (cf. the discussion of ironic use in section 4). As further support for our two notions of truth one may point to the fact that Tarski's stand on truth is far from being unambiguous. In the introductory part of his paper he characterizes truth in the Aristotelian sense as correspondence between what is said and what is. In the more formal part, on the other hand, Tarski deals with truth as something defined relative to another language (i.e. the meta-language). In our approach, these two definitions of truth are not treated as different aspects of the same thing, but rather as our two different notions of truth. Truth defined relative to the meta-language constitutes our notion of semantic truth, whereas truth defined in terms of correspondence between the token-model and the context-model constitutes our notion of pragmatic truth.

In order to further clarify our specific approach to truth and meaning, consider once more the top half of our SID-schema.

(36)

object-language E_{CAT} >

—meta-language IL^{D-meta}

translation-language IL>

(token model)

What is the nature of semantic truth in this set up? Take for example the object-language token *John doesn't walk.*, which is canonically analyzed in (37).

(37)

John —— $John_T$

doesn't —— $doesn't_{inf \downarrow (T \uparrow t)}$

walk —— $walk_{inf}$

. —— $t \uparrow t$

$\hat{P}[P]\ ((\hat{P}\hat{P}[\sim P(P)](walk'))\ (\hat{P}P(j')))$

meaning¹-formula: $\sim[walk'(j')]$

Consider now the semanitc interpretation of the lambda-reduced analyzed translation in (37), called meaning¹-formula. According to the standard definition of the negation operator, $\sim walk'(j')^{index}$ is 1 (true) relative to an index (consisting of a model-structure, a point of reference, and a variable assignment function), iff $walk'(j')^{index}$ is 0 (false). According to the definition of functional application, furthermore, $walk'(j')^{index}$ is 0, iff j'^{inDex} is not an element of $walk'^{index}$. In other words, the semantic truth values 0 and 1 are completely abstract entities. The logical operators are defined in terms of 1 and 0 just as much as 1 and 0 are defined in terms of the logical operators. These considerations regarding the nature of semantic truth-values apply not only to the translation language IL, but just as much to the formal meta-language IL^{D-meta}, since IL^{D-meta} is defined as a straightforward extension of IL.

What we have said here about the semantic truth values is not particular to our approach, but applies to Montague's system as well. After all, Montague defines the truth-values 1 and 0 purely abstractly as the sets $\{\phi\}$ aod ϕ contained in the power set over A of the model-structure $@=(A,I,J,\leq,F)$. Our definition of IL^{D-meta} (or perhaps rather IL^{D-meta}) is likewise compatible with Montague's system. Where we differ from Montague in particular, and the standard model-theoretic approach in general, on the other hand, concerns the status of the formal model structure, relative to which IL and IL^{D-meta} are interpreted. While the standard approach takes the model-structure as a representation of actual and potential reality given in advance, we defined the model-structure as part of a semantic space on the basis of which models satisfying the formula under interpretation are synthesized. As pointed out before, this switch from the interpretive to the synthesizing mode of model-theory does not change the the formal nature of the definitions of logical operators and the semantic truth values used in these definitions. Intuitively, on the other hand, the switch in question leads to a new paradigm of model-theoretic semantics (cf. (11), section 4) in that it makes room for the distinction between semantic and pragmatic truth, etc.

Now, the whole formal system indicated in (36), consisting of the object-language E_{CAT}, the translation language IL, and the formal metalanguage IL^{D-meta} (whereby IL and IL^{D-meta} define the meanings of E_{CAT}), is defined in terms of a common super meta-language, called SID-meta, which is simply natural English. This leads to the question: how much is gained by the semantic analysis of the object-language E_{CAT} in terms of IL^{D-meta} and how much of the semantic analysis is simply transferred to the unanalyzed notions of SID-meta? Or put differently, is the system indicated in (36) subject to the potentially infinite recursion of meta-languages inherent in Tarski's original conception?

In section 6 (25—28) it was determined that the quality of a metalanguage definition depends on the clarity of the notions used in it. We suggested furthermore that the only good meta-language definitions are those which can be *operationalized*, i.e. can be implemented in form of a computer algorithm (this position is inspired by, and may be regarded as a variant of, the so-called Church-hypothesis (cf. Kneale & Kneale 1962, p. 733 f)). The value of a formal, operational meta-language may now be formulated as the following lemma:

(38) *Lemma*:

A meta-language definition of an object-language term is guaranteed to be operational if the meta-language itself is an operationally defined formal language.

Due to the relation between a formal translation language (e.g. IL) and a formal meta-language (e.g. IL^{D-meta}), defined for one object-language (e.g. E_{CAT}), we may complement lemma (38) with lemma (39):

(39) *Lemma*:

If for any well-formed object-language expression Y there exists a formal translation in an operational translation language, then there exists an operational formal metalanguage definition for each Y of the object-language.

This is proven by the construction defined in (29) and (30) above, where the operational translation language is extended into a formal and operational meta-language.

As an example illustrating the point of the above lemmata consider the following analysis of the object-language expression **and** $_{T\downarrow(T\uparrow T)}$ of E_{CAT}.

(40)

$$\begin{array}{c} \textit{and}_{\;T\downarrow(T\uparrow T)} \\ | \\ \hat{P}_1\hat{P}_2\hat{P}[P_1(P) \wedge P_2(P)] \end{array}$$

The IL-translation in the molecule (40) consists solely of logical operators and variables (but no constants). It constitutes a completely operational definition of the meaning of **and** $_{T\downarrow(T\uparrow T)}$ in the sense that '$\wedge$', functional application, and lambda-abstraction of the IL-translation formula are all operationally definable. This standard translation analysis of **and** $_{T\downarrow(T\uparrow T)}$ in form of a ***molecule*** (a term introduced in Hausser 1980c, 1981a, 1982a) is equivalent to the following meta-language definition in $IL^{D\text{-}meta}$:

(41) $D(\ulcorner \text{and}_{T\downarrow(T\uparrow T)} \urcorner e)\langle s,\ f(T\downarrow(T\uparrow T))\rangle = \hat{P}_1\hat{P}_2\hat{P}[P_1(P) \wedge P_2(P)]_{\langle s,\ f(T\downarrow(T\uparrow T))\rangle}$

(41) constitutes an $IL^{D\text{-}meta}$ tautology, which may be proven analogously to proof (31).

However, IL as defined in PTQ consists not only of logical operators and variables, but also of so-called logical constants. As pointed out in section 6, IL and thus $IL^{D\text{-}meta}$ will be operational languages only to the degree that we manage to operationalize the definitions of the logical constants they contain. To this purpose we proposed (i) that each constant be replaced by a formula of equal type consisting of operators, variables, and ***elementary*** constants (e.g. **kill′** is replaced by CAUSE(BECOME(~ALIVE))) and (ii) that all the elementary constants be defined in terms of matching sense data.

It is, of course, no small task to define all the IL-constants like **walk′**, **talk′**, **man′**, **car′**, etc. in terms of formulas consisting only of elementary constants in combination with the standard operators and variables, and one might question, whether this program of semantic analysis has any realistic chance of realization. But Rome wasn't built in one day, and for the moment we may draw comfort from the consideration that the definition of such constants as **pain′**, which is regarded as extremely problematic within a nominalistic interpretation of standard model-theory, poses no problem at all in our approach. The problem for the nominalistic approach is to provide subjective sensations such as pain with an 'objective' (i.e. externally measurable) counterpart by identifying the sensations with brain states. This is the so-called double aspect theory, so-named because it deals with the relation of 'private sensations' with 'public' measurement data, a task that raises numerous philosophical problems. For a properly equipped SID, on the other hand, the sensation of pain is real enough, so that the elementary constant PAIN may be defined directly in terms of matching a certain type of sense data.

The above considerations show that in order to provide operational definitions for elementary constants we have to consider the SID as a whole (specifically elementary sense data and matching concepts). Note however that once an elementary constant like RED is defined in terms of a concept (matching a certain type of sense data), the constant (in combination with the concept denoted) may be used *independently* of the actual presence of a suitable perception. Furthermore, as pointed out in section 3, the elementary constants (and concepts) of the semantic space may be used in this perception-independent way both in the token-model (when triggered by language) and the context-model (when triggered by either memory or language).

As pointed out already in section 6, an SID-system (as indicated in (36)) breaks the infinite recursion of meta-languages, provided that the translation language IL and the correlated meta-language $IL^{D\text{-}meta}$ are completely operationalized in the manner described above. We require a higher meta-language, i.e. natural English or SID-meta, only in order to (i) define E_{CAT}, IL, and $IL^{D\text{-}meta}$, and (ii) to accomplish the actual implementation of these definitions in form of a computer model. But once this system is put into operation, complete with electronic perception, elementary concepts (denoted by elementary constants and defined in terms of matching elementary sense data), complex concepts (built up from elementary concepts and the standard logical operations), etc., our meaning analysis is independent of the higher meta-language of natural English. The more 'reasonable' and 'natural' the SID behaves in actual discourse, the more we can trust the adequacy of the formal analysis of the literal meaning of E_{CAT} in terms of $IL^{D\text{-}meta}$ in particular and the theory of communication in general, which are implemented in the SID. And conversely, each 'aberration' in the behavior of the SID will provide definite clues for improving the theory.

8. Pragmatic Truth and the Resolution of the Sorites Paradox

In the previous two sections we determined the completely abstract nature of semantic truth within our approach. The semantic truth-values (and the logical operators defined in terms of them) are used for the formal description of the literal meaning (meaning[1]) of the expressions of the object-language E_{CAT}. The literal meaning of an E_{CAT} expression is represented by the synthesized model (or more precisely, class of synthesized models) satisfying the standard IL-translation (or rather UL-translation) of the expression. The advantage of using synthesized models (interpreted as formal icons) rather than traditional models (interpreted as reality) is that our approach results in the separation of literal meaning of expressions and the context with respect to which these meanings are used.

Only if we have both, an icon representing the literal meaning of the expression and a context representing the utterance situation, do we arrive at a satisfactory representation of the traditional, Aristotelian notion of truth, that is the notion of truth in terms of *correspondence*. As proof compare once more our schematic representation of the two model-theoretic paradigms under discussion:

(11a) PARADIGM I

surface expression
| denotation=reference
model-theoretic reality

(standard approach)

(11b) PARADIGM II

surface expression
| dentotation
synthesized token model
⋮ use (reference)
synthesized context model
|
outside reality——

(our alternative approach)

The standard model-theoretic approach (Paradigm I) cannot be accepted as a satisfactory formalization of the Aristotelian notion of truth, according to which a sentence is true if it (or more properly, its meaning) corresponds with the situation referred to. The reason is quite obvious: the standard approach fails to provide anything that could correspond, because the elementary meanings and the elementary objects of the situation are *identified*.[14] The same objection applies, of course, to the so-called picture theory of Wittgenstein's Tractatus, where logical proper names stand for the things they name. Any theory that identifies the denotation of elementary logical expressions with objects of reality fails to provide a context-independent characterization of the literal meaning of expressions. And without a theory of expression-meanings it is impossible to define a correspondence between the expression-meanings and the state of affairs referred to.

Since the paradigm I theories, which erroneously claim to formalize the notion of truth in terms of correspondence, characteristically fail to distinguish between semantics and pragmatics, we are not contradicting anyone when truth defined in terms of correspondence reappears in our approach as the notion of ***pragmatic truth*** (in contrast to the semantic truth values discussed in sections 6 and 7, which are defined as ϕ and $\{\phi\}$). Furthermore, we find ourselves in agreement with J.L. Austin's observation that utterances are made for many other purposes than just making true or false statements.

The problem with Austin's approach, on the other hand, is that he takes the ***utterance*** as the smallest unit of analysis, treating even the literal meaning as an (locutionary) *act*. He thus fails to account for the *use* of the literal meaning of *expressions*, as was pointed out in section 2 above. Austin's complaint about the "craze of either being true or false" (Austin 1971, p. 20), applies only to the discourse semantic version of truth-conditional semantics (cf. section 2), but not to the sentence semantic version. The use of truth-or denotation-conditional semantics for representing the literal meaning of *expressions* (in terms of synthesized models, for example), is not affected by Austin's criticism, because it involves the truth-conditions of sentences, not of statements.

Let us consider now a number of examples which illustrate the necessity to distinguish between the literal meaning of expressions (compositionally encoded in the natural surface and formally represented by a synthesized model satisfying its translation into intensional logic) and the use of the expression relative to a context. The first example is (41), where a pragmatically true ironic statement is analyzed in terms of the (simplified) SID-schema:

(41) token/token-model: *That's real nice wheather today.*
(pragmatic correspondence)
context-model: *The wheather is terrible today.*

As mentioned in section 2 above, (41) constitutes a pragmatically true statement, if the correspondence is properly understood as an instance of ironic use. Thus even in the case of true statements, the correspondence between the literal meaning of the token and the context may be quite varied. Only in statements where the token is used literally does the token model and the context model agree in the sense that there is a proper embedding of the token model in the context model.

Next consider a brief passage from W. Somerset Maugham, which illustrates how the literal meaning of words may be used metaphorically:

(42) Roy's repertory was extensive and his scent for the word of the minute unerring; it peppered his speech, but aptly, and he used it each time with a sort of bright eagerness, as though his fertile brain had just minted it.

According to our analysis, when one reads this sentence one takes the literal meanings of the words and derives from it a speaker meaning (represented as the SID-internal context) which is quite different from the literal meaning (represented as the SID-internal token model). Just consider the literal meaning and the meaning[2] of such phrases like "scent for the word," "word of the minute," "to pepper one's speech," "to mint words," or "fertile brain". I take it that there would be hardly any disagreement on what is meant[2] by this passage. The meaning[2] of "to pepper one's speech," for example, is to be explained in terms of an analogy, namely to putting pepper, i.e. a certain kind of spice, into a dish.

Nevertheless, grasping the distinction between the literal meaning, on the one hand, and different uses, on the other, requires a certain amount of abstraction because any use of natural language involves by definition an element of use. Therefore, the closest we come in most cases to isolating the literal meaning of an expression is to isolate its literal use. From the literal use we may infer the iconic content (i.e. the literal meaning) of the expression, since literal use is defined as that type of use where the iconic content of the expression matches the context completely, at least in statements.

Since the element of use is always involved in natural language, the existence of literal meanings has been vigorously denied by those sympathetic to speech act theory (which proceeds on a holistic notion of meaning, without a distinction between iconic content and use). Many exponents of this school accept only the surface form of the symbol and its various uses, whereby the uses are regarded as the various meanings. This general view is attributed to the later Wittgenstein, who said "For a large class of cases though not for all in which we employ the word 'meaning' it can (!) be defined thus: the meaning of a word is its use in the language." We don't deny this and we agree that in most cases there is a whole spectrum of meanings (corresponding to different uses). Only, we call these meanings speaker meanings (meanings[2]) and assume in addition the existence of the literal meaning (meaning[1]) of the expression or symbol used. Note that the distinction between the icon (literal meaning) and its use can be found also in Wittgenstein's *Philosophical Investigations*, though only as a footnote:

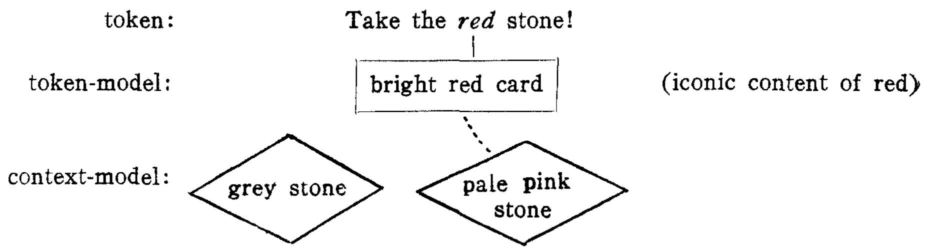
> Let us imagine a picture, representing a boxer in a certain fighting position. This picture can be used to inform someone how he should stand, how he should carry himself; or which position he should not adopt; or how a certain man stood there and there, or etc. etc. One could call this picture (chemically speaking) a sentence radical. Frege presumably interpreted the term "Annahme" ('assumption') in a similar way.

(Wittgenstein, *Philosophical Investigations*, Note to paragraph 22, translated after the Suhrkamp edition edited by Anscombe and Rhees.)

In order to give further support to our claim that meanings^2 are best explained in terms of the iconic content (meaning^1) of the symbol and its use, let us consider once more the word **red**, the iconic content of which is rather clear. It may be defined operationally in terms of matching a certain wavelength of the electro-magnetic spectrum, or, in a more practical way reminiscent of the later Wittgenstein, in terms of little card of bright red color.

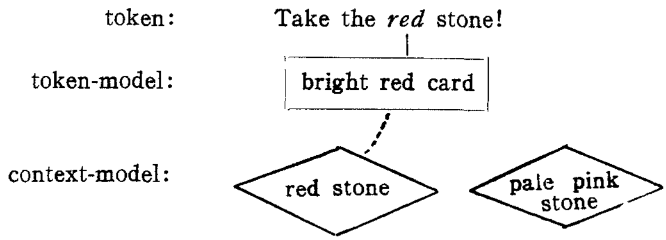
Now, the fact that we can say unambiguously: "Take the red stone!" in a situation where one stone is grey and the other pale pink has been taken as proof that the meaning of **red** is vague. But this claim does not hold water. Rather, in the situation described, the fixed iconic content of **red** is used to pick out the stone the color of which *matches best.*

(43)



If we change to a context where the grey stone is replaced by dark red stone, the pale pink stone cease to be the one that matches the bright red card best. Thus we have a situation as indicated in (44):

(44)



So what happens to be the "red stone" in (43) turns out to be the "non-red stone" in (44). The point is that it is not the meaning of red that is vague or which changes, but rather it is the context which changes and thus the instances of best match. Thus examples like (43) and (44) are no counterexamples to the assumption of a fixed meaning^1 defined as the iconic content of the symbol. To the contrary, they may be taken as support for this assumption.

Next let us consider the so-called Sorites paradox or paradox of the heap, which has recently received considerable attention (Kamp 1981, Kindt 1982, Pinkal 1982). The paradox is described as follows. One grain of sand does not form a heap. If I add one grain, I still don't have a heap. If grains don't form a heap, then adding an n+t[th] grain will not result in a heap. Yet at some point, when enough grains are added, we arrive at something that is undeniably a heap.

The recent proposals to resolve this paradox all accept it as a semantic paradox and thus stay within the traditional framework of semantics. But the price paid for these different kinds of so-called semantics of vagueness is considerable. Kamp arrives at a notion of semantic inference which is so far removed from the traditional notion that he himself has doubts as to whether his system may still be called a logic. Kindt, on the other hand, proposes to incorporate the heavy machinery of mathematical typology into formal semantics, whereas Pinkal's approach of 'precisification' constitutes a sophisticated development of the method of supervaluations.

These proposals have in common that they accept the premises which lead to the paradox. But when we look at another antique paradox, that of Achilles and the turtle, which today is regarded as solved, we see that one acceptable resolution of a paradox is to revise its premises in an intuitively convincing way. The moment we accept that a heap is to be defined in terms of a certain number of grains (e.g. 1 grain: no heap; 100, 000 grains, properly arranged: heap) we are trapped in the paradox. Because now comes the inevitable question: how many grains exactly make the difference between a heap and a non-heap?

But a heap (as an icon or literal meaning) is not to be defined in terms of a certain number of grains, not even upper or lower limits on this number. Rather, the icon of a heap is a prototype involving (i) a certain form (cone-like), a certain subsistence (loosely packed smaller parts, the size of which is in a certain proportion to the size of the whole heap), and a certain size of the heap in comparison to the rest of the context. When two people fly in 10, 000 feet altitude over a farm and A says to B: "That heap wasn't there yesterday.", pointing to what looks like a speck on the ground, he violates the proper use of the icon **heap**, even if it should turn out later that the speck on the ground is indeed a proper heap of sand. The speaker A may be *construed* to be in the right in a narrow, pseudo-scientific or pseudo-semantic sense, but that does not mean that A communicated in a natural and reasonable way.

Of course, if A were to say to B: "Do you see that tiny speck down there? That must be a heap of sand. I don't think it was there yesterday." the situation would be different. In the second case, the speaker A introduces a *context-change*. A leads B from point of view I (at 10, 000 feet altitude) to point of view II (at the ground level close by). In the second (imagined) context the speck in question may well be a proper heap. It is of no consequence that B cannot *verify* A's conjecture. All that is required is that B is a cooperative partner in this communication in the sense that B is willing to *provide* a context which accommodates the icon **heap** (on the literal interpretation intended by A).

The difference between the semantic and the pragmatic approach to the Sorites paradox may now be summarized as follows. The semantic approach presumes a model-theoretic reality which provides varied samples of heaps and non-heaps, starting from one cone grain and going up to a 100, 000 grains, say. The supposed problem is to find a semantic definition of the logical constant **heap**, such that heap(x) is evaluated 0 (false) if x denotes only one grain, **heap**(x) is evaluated 1 (true) if x denotes the 100, 000 grains, and which furthermore assigns the right truth-values in the critical transition from non-heap to heap. However, no semantic theory can fulfill the last desideratum, because the transition from non-heap to heap is intuitively unclear in a non-trivial sense.

On the pragmatic approach, on the other hand, there is no attempt to characterize the transition from non-heap to heap in the semantics. Rather, the icon **heap** is semantically defined in a fixed way as a prototype, just as the icon **red** was defined in terms of a little red card. The question of whether something is properly referred to as a heap or not is left to the pragmatic process of matching the icon with the conext. What counts as a proper heap in one context, may be a definite non-heap in another context (just by changing the relative proportions of the objects relative to each other and relative to the context-frame). This is similar to our example of a pale pink stone, which turned out to be the red species in (43) and the non-red species in (44). A further possibility (never even discussed in the semantic approaches) is the metaphoric use of the icon **heap**, such as when an old car is referred to as a "heap of scrap." In this case, the icon evokes an imagined future state of disintegration which is felt to be so immediately pending as to justify this manner of speaking.

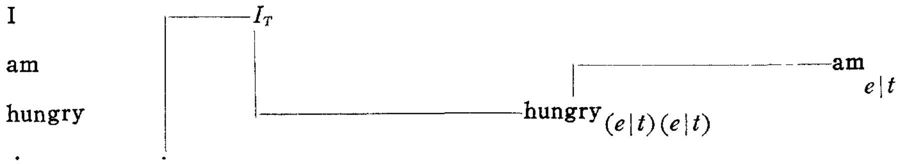
Notes

(1) The basic idea of the SID and the related reinterpretation of model-theoretic ontology was first formulated in Pittsburgh, April 1979, in an abstract to the 1979 LSA summer school conference. The theoretical concept of the SID was presented in talks at Stanford Philosophy Colloquium, January 21, 1980; the Harvard Linguistics Circle, February 19, the Stanford 1980; the Third Amsterdam Conference on Formal Grammar, Ma-

rch 25, 1980, and other places. The sections 1 and 2 of the present paper overlap with the corresponding sectons of the Amsterdam paper (Hausser 1980a). The basic idea of extending the formal translation language of intensional logic into a formal meta-language (sections 6 and 7) was first presented in the spring of 1979 at an informal working group of the Philosophy Department, University of Pittsburgh, led by Nuel Belnap and Richmond Thomason. I also discussed it with Ian Hacking of the Stanford Philosophy Department in the summer of 1979. The treatment of vagueness and of the Sorites paradox in section 8 was inspired by a workshop on altenative semantics, organized by Manfred Pinkal and Dieter Wunderlich, in Schloss Mickeln, March 1982. The substance of the present paper was presented as a lecture at the '82 Seoul Workshop on Formal Grammar, January 6, 1982, in Seoul, Korea. I would like to thank the organizers and participants of this conference, as much of the later additions and developments were prompted by the stimulating discussions at this workshop. I would also like to thank the president of the Linguistic Society of Korea, Professor In-Seok Yang, for his offer to publish this paper and his patience in receiving the manuscript. The final draft was read by my collegue, Claudia Gerstner. This resulted in a number of corrections and improvements for which I am grateful.

(2) A further question is whether (6) should be regarded as false (0) or as undefined (#). If we regard (6) as analytically or necessarily false, then there exist no models satisfying (6) itself, only models satisfying its negation. If we regard (6) as undefined, then there are no models satisfying either (6) or its negation. The fact that contradictory expressions like 'square circle' can be *used* in meaningful ways raises the general question of how the literal meaning of such expressions should be analyzed in a model-theoretic framework. In our particular approach, the most straightforward solution would be to synthesize contradictory models. The-technical nature of such models satisfying contradictory expressions will be discussed at another occasion.

(3) Some readers may wonder about the context-independent literal meaning of sentences containing indexicals such as **I am hungry**. The coordinates approach (briefly discussed here in section 2), e.g. Montague (1974, chapter 3, Lewis 1972) comes to the counterintuitive conclusion that the meaning of such an indexical expression *varies* from context to context. Thus the meaning of *I am hungry* is supposed to vary from speaker to speaker. However, indexicals may be assigned a context-independent literal meaning if we adopt a treatment based on context variables, first proposed in Hausser (1974). Consider for example the following canonical analysis of **I sm hungry**:



$$\hat{p}[p]\ (\hat{P}\wedge x\varepsilon[\Gamma_1(x)]^1P(x)\ \ (\hat{P}\hat{x}[P(x)\wedge\text{hungry}'(x)]\ \ (\hat{y}[Vz[z=y]])))$$

lambda-reduction :

$$\hat{p}[p]\ (\hat{P}\wedge x\varepsilon[\Gamma_1(x)]^1P(x)\ \ (\hat{P}\hat{x}[P(x)\wedge\text{hungry}'(x)]\ \ (\hat{y}[Vz[z=y]])))$$

$$\hat{p}[p](\hat{P}\wedge x\varepsilon[\Gamma_1(x)]^1P(x)\ \ (\hat{x}[\hat{y}[Vz[z=y]](x)\wedge\text{hungry}'(x)]))$$

$$\hat{p}[p]\ (\hat{P}\wedge x\varepsilon[\Gamma_1(x)]^1P(x)\ \ (\hat{x}[Vz[z=x\wedge\text{hungry}'(x)]]))$$

$$\hat{p}[p](\wedge x\varepsilon[\Gamma_1(x)]^1x[Vz\,[z=x\wedge\text{hungry}'(x)]]$$

$$\wedge x\varepsilon[\Gamma_1(x)]^1x[Vz[z=x\wedge\text{hungry}'(x)]]$$

meaning[1]-formula: $\wedge x\varepsilon[\Gamma_1(x)]^1\ Vz[z=x\wedge\text{hungry}'(x)]$

In the meaning[1]-formula (as well as in the analyzed surface translations and the steps of the lambda-reduction), the context-dependency-aspect of first person is represented by the context-variable $\Gamma_{1e|t}$ in the translation (and its remainders) of *I*. The meaning of the context-variable is specified in terms of its so-called ***reconstruction condition.*** The reconstruction-condition of Γ_1 says that Γ_1(x) is 1 if and only if x has the property of being the speaker. (specification of such a property finds an obvious solution in a model-theory that is interpreted as part of a speaker-simulation device). By treating the characteristic property of an indexical in terms of its 'context-dependency aspect' (expressed by the context-variable), we may specify the meaning of an indexical independent of the specific use-context of a token. For detailed analyses of anaphoric and indexical pronouns see Hausser (1979a, 1979b).

(4) For a critique of Hintikka's analysis of questions see Hausser (1978a).

(5) See for example Wunderlich (1976), where literal meanings are defined in terms of standard contexts.

(6) It has been argued in Hausser (1979a), (1979b) that the so-called anaphoric pronouns like **he, she, it, himself, herself, itself,** etc. do not constitute a well-defined natural class. Instead. the anaphoric interpretation of **he, she, it,** etc. should be treated as a special case of the indexical interpretation. which is common to all personal pronouns

I, you, he, she, it, etc.. The anaphoric interpretation is characterized in terms of an 'internal reconstruction' of the context-variable (cf. footnote 3), i.e. the value for the context-variable is taken from the sentence, whereas the indexical interpretaion is characterized by an 'external reconstruction' (the value for the context-variable is taken from the context). Pronouns which have only indexical interpretations are characterized by context-variables defined only for external reconstruction, whereas pronouns like **himself, herself,** etc., which allow only anaphoric interpretations of a certain restricted kind, are characterized by context-variables defined only for internal reconstructruction.

(7) This seems to be illustrated by the common phenomenon that young dogs get upset at the sight of a one-legged man.

(8) These pragmatic principles have not been fully discovered yet and consequently not been presented in terms of a formal theory. Rather, up to now we have made due with the analysis of certain examples: "Could you pass the salt?" (Interrogative used for polite request), "It's really nice weather today." (ironic statement), "The old fox left the room." (metaphoric reference), etc. However, that these general principles of use do *exist*, and are therefore open to discovery, may be inferred from two kinds of facts (beyond the plausibility of our analysis of certain speech acts in terms of literal meaning and its use relative to a specified context). For one, the principles of language use are present very early in language acquisition, as illustrated by our "bow-wow" these principles may even be observed in chimpanzees like Washoe, who have mastered example. And secondly, the use of a sign language. It is therefore a fair assumption that the pragmatic principles in question are of a very basic and general nature.

(9) As natural this proposal seems in the context of the SID-framework, as radical are the revisions it implies for modal and epistemic logics based on possible worlds. I did not envision these consequences when I wrote 'A constructive approach to intensional contexts'. I plan to write an appendix to this paper which proposes to treat possible worlds as points on a SID-subjective certainty x desire scale, rather than possible states of the external universe. In this way, the formal notion of an intension as a function from (reinterpreted) possible worlds may hopefully be preserved.

(10) The problem with the empiricists in general, and Locke in particular, is their rather wide use of the term *idea*. Much of what is said here about the semantic space of the SID may be read as an attempt to rephrase and specify Lock's notion of 'simple' and 'complex ideas' in the particular terminology of the SID-framework.

(11) The bracketed descriptions of vision, hearing, etc., are rough quotations from Websters New Collegiate Dictionary, edition 1976.

(12) In accordance with Hausser (1980c) and (1982), we distinguish three levels of logical analysis, namely (i) the surface logic SL, (ii) the root logic RL, and (iii) the universal logic UL. For example, the SL-translation of the surface word ***kills*** $T \downarrow (T \uparrow t)$ is ***kills***^{0}, the RL-translation is $\hat{P}_1\hat{P}_0[P_1\hat{x}[P_1\hat{y}[\text{kill}'(x,y)]]]$, where **kill′** is a so-called root, a constant of a type corresponding to the category $e|(e|t)$, and the UL-translation is $\hat{P}_1\hat{P}_0[P_0\hat{x}[P_1\hat{y}[\hat{m}\hat{n}[\text{CAUSE}(m,\ Vz[\hat{o}(z)=\text{BECOME}(n,\ \hat{P}[P(n)=-\text{ALIVE}(n)[)]])]\ (x, y)]]]$. The UL-translation is the most basic level of translation in that all roots have been replaced by UL-formulas which consist only of explicitly defined operators and constants. The SL-, RL-, and UL-translations of a surface expression are provided by the lexicon, whereby they must be defined such that (i) they are of the same semantic type and (ii) semantically equivalent. However, the three levels of logical analysis just mentioned are restricted to the token-recognition loop of (14). On the level of context-recognition, which is presently under discussion, concepts are represented only in terms of the most basic level of logical formulas, i.e. formulas of UL.

(13) Note that our notion of a ***proto***type is distinct from Putnam's (1975) notion of a ***stereo***type. Whereas Putnam's stereotypes are characteristically ***public*** and not in the head, our prototypes are regarded as something in the head of the speaker or SID. Prototypes are also public to the degree that different speakers of the same speech community agree on what is, e.g., "really red".

(14) Thus the problem arises again with the definition of elementary constants in the standard approach, though the present question of properly defining ***correpondence*** in order to capture the Aristotelian notion of truth presents the preblem from a rather different angle.

References

Anderson, J.R. (1976): *Language, memory, and thought*, Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Austin, J.L. (1962): *How to do things with words*, Clarendon Press, Oxford.

Aristotle: *Aristotles Metapysics*, a revised text with introduction and commentary by W.D. Ross, Oxford 1949.

Barwise, J. (1981): "Scenes and other situations", *Journal of Philosophy* LXXVIII, 7, pp. 369—397.

Berkeley, G. (1710): "A treatise concerning the principles of human knowledge", reprinted in *The Empiricists*, Anchor Books, Garden City, New York (1974).

Carnap, R. (1947): *Meaning and necessity, a study in semantics and modal logic*, University of Chicago Press, Chicago.

Davidson, D. (1967): "Truth and meaning", *Synthese VII*, pp. 304—323.

Descartes, R.: *Philosophical writings*, a selection translated and edited by E. Anscombe and P.T. Geach, with an introduction by A. Koyre. London, 1954.

Frege, G. (1982): "Über Sinn und Bedeutung", *Zeitschrift für Philosophie und Philosophische Kritik* 100, pp. 25—50.

Friedman, J., Moran B., and Warren, D.: (1978) "Dynamic Interpretations", *Computer Studies in Formal Linguistics* No. 16, Department of Computer and Communication Sciences, University of Michigan, Ann Arbor.

Grice, H.P. (1957): "Meaning", *Philosophical Review*, pp. 377—388.

Groenendijk, J. and Stokhof, M. (1975): "Modality and conversational information", *Theoretical Linguistics* Vol. 2, No. 1/2, pp. 61—112.

Hacking, L. (1975): *Why does language matter to philosophy*? Cambridge University Press, Cambridge.

Hausser, R.R.(1974): *Quantification in an extended Montague grammar* (Ph.D. dissertation, The University of Texas), Ann Arbor, Michigan (University Microfilms).

Hausser, R.R. (1976): "Presuppositions in Montague Grammar", in *Theoretical Linguistics* 3, 3.

_______(1978a): "Surface Compositionality and the Semantics of Mood", in J. Groenendijk and M. Stokhof (eds.), *Amsterdam Papers in Formal Grammar*, Vol. 2, University of Amsterdam (1978); reprinted in: J.R. Searle, F. Kiefer and M. Bierwisch (eds.), *Speech Act Theory aud Pragmatics*, Dordrecht : Reidel Publishing Company (1980).

_______(1978b): "Chapter I, Surface Compositional Semantics: Methodological Deliberations", Pittsburgh 1978, revised: Stanford 1979 and München 1980. Unpublished manuscript, 123 pages.

_______(1979a): "How do pronouns denote?" in Heny, F. & Schnelle, H. (eds.), *Syntax and Semantics*, Vol. 10, Academic Press, New York.

_______(1979b): "A new treatment of context in model theory", appeared in *Sull' Anafora*, atti del seminario Accademia della Crusca, 14—16 dicembre 1978, Firenze, presso L'Accademia della Crusca MCMLXXXI.

_______(1979c): "A constructive approach to intensional contexts", to appear in *Linguistic Journal of Korea* (1982).

_______(1980a): "The Place of Pragmatics in Model-Theory", in Groenendijk et al (eds.), *Formal Methods in the Study of Language*, Mathematical Center Tracts 135, The University of Amsterdam.

_______(1980b): "The Syntax and Semantics of English Mood", to appear in *Question and Answers*, KVAL-series Linguistic Calculus, Dordrecht: Reidel Publishing Company.

_______(1980c): "Chapter III : Categorial Surface Syntax", München, unpublished manuscript, 68 pages.

_______(1981): "Chapter IV : The Structure of the Lexicon", München, unpublished manuscript, 67 pages.

_______(1982): "A surface compositional categorial syntax", to appear in the proceedings of the '82 *Seoul Workshop on Formal Grammar*, Seoul, Korea.

Hintikka, J. (1976): *The semantics of questions and the questions of semantics*, Acta Philosophica Fennica, Vol. 28, No. 4, North-Holland Publishing Company, Amsterdam.

Hume, D. (1748): "An enquiry concerning human understanding", reprinted in *The Empiricists*, Anchor Books, Garden City, New York (1974).

Kamp, J.A.W. (1981a): "A theory of truth and semantic representation", in Groenendijk, J., Janssen, T., and Stokhof, M. (eds.), *Formal Methods in the Study of Language*, Part 1, Mathematical Center Tracts 135, Amsterdam 1981, pp. 277—322.

______ (1981b): "The heap paradox", unpublished.

Karttunen, L. and Peters, S. (1978): "Conversational implicature", in Dineen, D.A. and Choon-Kyu Oh (eds.) *Presupposition: Syntax, and Semantics*, Vol. 11, Academic Press, New York, N.Y.

Kindt, W. (1982): "Vagueness as a topological problem", to appear in Ballmer, T. and Pinkal, M. (eds.), *Approaching Vagueness*, North-Holland, Amsterdam.

Klemke, E.D. (ed.) (1971): *Essays on Wittgenstein*, University of Illinois Press, Urbana, Chicago, London.

Kneale, W. and Kneale, M. (1962): *The development of logic*, Oxford University Press, Ely House, London.

Kripke, S. (1962): "Semantical considerations on modal logic", in *Acta Philosophica Fennica* 16, pp. 83—94, reprinted in Linsky, L. (ed.) *Reference and Modality*, Oxford University Press, London 1971.

______ (1972): "Naming and Necessity", in Davidson, D. and Harman, G. (eds.) *Semantics of Natural Language*, 253—355, Synthese Library, Reidel, Dordrecht and Boston.

Lewis, D. (1972): "General semantics", in Davidson, D. and Harman, G. (eds.) *Semantics of Natural Language*, pp. 169—218, Synthese Library, Reidel,

Linsky, L. (ed.) (1972): *Semantics and the philosophy of language*, University of Illinois Press, Urbana, Chicago, London.

Locke, J. (1690): "An essay concerning human understanding", reprinted in *The Empiricists*, Anchor Books, Garden City, New York (1974).

Meinong, A. (1904): "Über Gegenstandstheorie", in Meinong, A. (ed.) *Untersuchungen zur Gegenstandstheorie und Psychologie*, Johann Ambrosius Barth, Leipzig.

Mills, J.S. (1950): *Philosophy of scientific method*, edited with an introduction by Ernest Nagel, Hafner Press, New York, London.

McCawley, J.D. (1968): "Lexical insertion in a transformational grammar without deep structure", in *CLS* 4, pp. 71—80.

Montague, R. (1974): *Formal philosophy, the papers of Richard Montague*, edited and with an introduction by Richmond Thomason, Yale University Press, New Haven.

Pinkal, M. (1981): "Konsistenz und Kontextwechsel: das Sorites-Paradox", unpublished.

Putnam, H. (1975): *Mind, Language, and Reality*, Philosophical papers, Vol. 2, Cambridge University Press, Cambridge, New York, London, Melbourne.

Quine, W.v.O. (1960): *Word and Object*, MIT Press, Cambridge, Mass.

Russell, B. (1905) : "On Denoting", in *Mind* 14.

Saussure, F. de (1916): *Cours de Linguistique Generale*, Paris, Payot. English translation: *A course in General Linguistics*, New York, Philosophical Library.

Searle, J. (1969): *Speech Acts*, Cambridge University Press, London.

______ (ed.): *Philosophy of Language*, Oxford University Press, London.

Stalnaker, R.C. (1970): "Pragmatics", in *Synthese* 22.

_______(1978): "Assertion", in Cole, P. (ed.), *Syntax and Semantics*, Vol. 9, Academic Press, New York, pp. 315—332.

Tarski, A. (1935): "Der Wahrheitsbegriff in den formalisierten Sprachen", *Studia Philosophica*, Vol. 1, pp. 261—405.

_______(1944): "The semantic conception of truth", in *Philosophy and Phenomenological Research*, Vol. 4, reprinted in Linsky, L. (ed.) *Semantics and the Philosophy of Language*, University of Illinois Press.

Wittgenstein, L. (1960): *Tractatus logico-philosophicus*, Suhrkamp Verlag, Frankfurt.

_______(1967): *Philosophische Untersuchungen*, Suhrkamp Verlag, Frankfurt.

Wunderlich, D. (1976): *Studien zur Sprechakttheory*, Suhrkamp Verlag, Frankfurt.