

COMMONSENSE ONTOLOGY AND SEMANTICS OF NATURAL LANGUAGE

Johannes Dölling

Summary.

This paper is concerned with the problem of what ontological knowledge underlies the semantic structures of natural language. It discusses some proposals made by workers in philosophy, linguistics, cognitive science and artificial intelligence in the last few years. Section 1 sketches the basic view in regard to the relationship between ontology and semantics. Section 2 points out the importance to analyse the commonsense view of the world. Section 3 outlines semantic investigations which take into consideration the variety of ontological categories reflected by natural language expressions. Section 4 establishes the necessity of a cognitive approach to ontology and semantics. Finally, Section 5 specifies a few consequences concerning the semantic analysis of nominal expressions.

1 Introduction

A basic characteristic of natural language is that it can be used by human beings to reflect and to communicate about the world. This inherent "aboutness" of language is one of the reasons why investigations in this area are necessarily connected with philosophical decisions. Anyone who deals with the semantics of natural language and, therefore, with the explication of the relationship between language and the world, is driven to ask questions like these:

- (1) What is the world which people talk about?
- (2) What categories or fundamental kinds of entities are there and how are they related?

Indeed, such questions lead the inquirer into the field of ontology or metaphysics. To give a serious account of the semantics of natural language, one has to find convincing answers to them. In this paper, I want to say something on the problem of what ontology is implicitly included in natural language and should be described for the purpose of semantic analysis. My first point is that the variety of semantic sorts of expressions indicates that there is not simply one undifferentiated domain of entities in our world. Rather, we conceptually discriminate numerous ontological domains - e.g. things, persons, stuffs, aggregates, configurations, collections, groups, institutions, events and states, as well as pluralities of these - for which the language disposes of various means to refer. In this sense, the semantics has to take into consideration a view of the

world reflected by our commonsense ontology. However, a further aspect of the problem concerns the question of where these ontological structures come from. In response to this, my second point is that the world which natural language is about must not be taken to involve structures independent of man. On the contrary, our world is linked to our cognitive capacity and does not exist in the absence of human creators. In this respect, the semantics has to consider a view of the world which conflicts with *naive* common sense.

I will offer some suggestions on how a cognitively founded analysis of natural language ontology could be realized. In connection with this, I will make a few proposals on how the results of such investigations can be used in the explanation of semantic phenomena.

2 Two views of analytic ontology

Analytic philosophers have paid much attention to ontological problems raised by work in the area of analysis of language. Particularly, using the logical methods invented by Frege, Russell, Wittgenstein, Leśniewski and others, ontology has been transformed into formal ontology. However, a task is to find out how much has been done by analytic philosophers for an illumination of ontology underlying the semantic structures of natural language.

According to Strawson (cf. Strawson 1959), there are at least two conceptions of theoretical ontology obeying principles of analytic philosophy. The first, which Strawson calls *revisionary metaphysics*, tries to identify the "real" constituents of the world, the ones to which we ought to be rationally committed, whether they coincide with our commonsense conceptual scheme or not. The second, which Strawson calls *descriptive metaphysics*, has a different focus. This conception tries to describe the world of our ordinary experience and thought. Although both directions of ontological analysis follow the tradition of Ockham's razor, it is obvious that they differ in the results of their investigations.

For a long time, revisionary metaphysics has dominated the development of analytic philosophy. In correspondence with the idea of reductionism, its advocates have the attitude that philosophical progress is to be measured by the degree to which one can explain away apparent givens in terms of less controversial entities. In this way, Ockham's razor: thou shalt not multiply entities beyond necessity, is interpreted as: thou shalt diminish entities wherever possible. In an extreme case, the results of such an approach are "one-category ontologies". Particularly, many philosophers of this orientation take it for granted that the real structure in the world is physical structure. Accordingly, the ultimate aim of their ontological considerations is directed to something like the categorial commitments of physics, or more specifically, of microphysics.

But such an "ideal" ontology is not all ontology which we need. Instead of this, for the

purposes of everyday life, we use an ontology which accommodates every entity which can be the object of ordinary talking. The world given by this includes, for example, coloured and sounding objects, i.e. objects with qualities being not real for physics. Therefore, developing a picture of commonsense world conforms to Strawson's conception of descriptive metaphysics in not contemplating the effect of revisionary principles. Indeed, such an ontological analysis is directed towards an elucidation of the structure of our commonsense view of the world. Like every other empirical investigation, it must meet the demand of descriptive adequacy.

In the last few years, much work has been done by analytic philosophers in this area. Following the tradition of Aristotle and using, for instance, ideas of Brentano, Meinong and Husserl, several aspects of commonsense ontology are treated formally. I suggest that at the current stage of development investigations directed to certain general structures like part-whole relations are particularly interesting (cf. Smith 1982, Simons 1987). Simons, for example, has shown that previous treatments of the concepts of part and whole, especially the classical mereology of Leśniewski and the calculus of individuals of Leonard and Goodman, are descriptively inadequate on the ground that they are not applicable to most of the entities around us (cf. Simons 1987). Taking into consideration the necessity of distinguishing between several ontological domains in our world, he argues for more differentiation in the categorial analysis. Above all, instead of assuming one mereological relation to be universally valid, Simons inquires into different but analogous part-whole, giving a specific structure to each ontological domain.

A number of recent moves have been made by workers in the field of artificial intelligence in the direction of a formal account of our everyday world (cf. Hobbs/Moore 1985). This enterprise, initiated by McCarthy and Hayes, was furthered by the realization that the development of computer systems exhibiting general intelligence, among other things a language faculty, would require a systematic analysis and formalization of virtually all our commonsense knowledge. Accordingly, researchers interested in such a knowledge representation supply characterizations of entities like space, time, movement, shape, material bodies, stuffs (solids and liquids), events etc. As for as, in comparison with philosophy, these investigations go more into particulars and do not restrict themselves to categorial analysis, they establish specific empirical sciences like commonsense physics, sociology or psychology.

There is agreement among all practitioners of descriptive ontology, whether they work in philosophy or in artificial intelligence, that natural language, i.e. the language we use to talk about the commonsense world, constitutes a very important intermediate of ontology. Therefore, considerations of linguistic usage are an integral part of their investigations. A systematic analysis of meanings, however, is realized in the area of natural language semantics.

3 Ontological presuppositions of semantics of natural language

Model-theoretic semantics is usually regarded as the paradigm of formal semantics of natural language. It involves the construction of set-theoretic models of those things in our world which make up the denotations of natural language expressions. Evidently, decisions about the content and structure of such models are founded on the ontology assumed with respect to the world. In order to minimize the ontological basis, analytic philosophers and linguists working in the field of model-theoretic semantics have commonly restricted themselves to just a few kinds of primitive entities. In particular, Montague's model structures make use of three sets: the set of individuals forming the universe of discourse, the set of possible worlds and the set of moments of time. All other things taken into consideration are modelled by set-theoretic constructs over these ingredients.

However, though the reductionistic strategy of modelling is formally correct and very effective it gives rise to some difficulties. Presupposing a homogenous universe of discourse, its orthodox adherents represent distinctions between categories of individuals by means of discriminations within the hierarchy of sets (or functions). It is easy to show that such an approach is inadequate in several respects. Additionally, there is a multitude of categorial differences reflected in the way we talk about the world that cannot be grasped in this manner. As a result, the semantic analysis fails to correspond with natural language ontology. For such reasons, the principle of ontological parsimony assumed in model-theoretic semantics has increasingly been retracted in favour of the recognition of several categories of individuals.

To illustrate this circumstance, I will briefly touch on the problem of semantic plurality. In the literature, there are two ways to treat plural nominal phrases in a set-theoretic fashion. Traditionally, a definite plural NP like *the linguists* is represented by the same predicate as *linguist* and, therefore, receives the same set of individuals as its denotation, whereas the definite singular NP *the linguist* denotes an individual. This gives rise to the semantic difficulty that many verbal phrases need various interpretations dependent on whether they are combined with an individual-denoting or set-denoting NP. Thus, for example, *walked* has to be represented by a first as well as a second-order predicate. In order to overcome this difficulty, some authors suggest that not only plural but also singular NP's denote sets of individuals. Actually, singleton sets are assumed to be denotations of definite singular NP's. However, both proposals are ontologically defective: abstract entities outside of space and time are turned into bearers of qualities characteristic of concrete objects.

In my view, an adequate approach to the semantics of plurals is to be found in Link

(1983, 1991). The main points of Link's proposal may be summarized as follows: No set-theoretic model is given for pluralities; rather, the universe of discourse is extended so as to comprise ordinary, i.e. atomic, individuals along with plural individuals. Indeed, the universe involves an ordering structure in which pluralities are identified as individual sums. By this means, the interpretations of singular and plural NP's can be treated on a par. There is another important aspect to this approach. Taking into consideration the obvious analogy between plurals and mass terms, a combined theory is supplied for both phenomena. According to this theory, the denotations of mass terms, which represent an intricate problem for semantics (cf. Pelletier/Schubert 1989), are also considered to be individuals inside a specific partial ordering.

Link's assumptions about the ontological basis of plural and mass expressions are on a level with general modifications performed in the methodology of natural language semantics. Indeed, in the last fifteen years, a number of different categories of entities have been admitted to the universe of discourse. Thus, in addition to pluralities and quantities of matter, kinds, stages, situations, events and degrees have all been regarded as individuals in the attempt to capture the ontology involved in natural language (for a review, cf. Bach 1986). In this context, it is significant that intensional entities like properties and propositions have also been introduced into the domain of model structures (cf. Chierchia/Turner 1988). In view of this development, it seems to me that one should abandon all ontological scruples and allow as an individual everything that can be the object of natural language predications. In this way, we end up in a "Platonic universe" which contains several sorts of particulars and universals, of concrete and abstract entities as well as of mental or extramental entities, regardless of whether or not they actually (or potentially) exist.

To sum up, I want to emphasize that the crucial innovation realized by Link is the idea of structuring subdomains by means of various part-whole relations, which, together with the completeness of the corresponding sum operations, create algebraic structures on each of the domains. Supplying axiomatic characterizations for the several subdomains leads us away from extensive set-theoretic modelling and, thus, from a mood of ontological reductionism. Following this idea, Krifka has extended this algebraic approach to further categories of entities, such as events, times, places, types and tokens (cf. Krifka 1989). As a result, he assumes a system of postulates which characterize the lattice structures on the domains and lay the basis for connections between these.

Needless to say, there are points of connection with the research outlined in the second section of this paper. Obviously, projects carried out in several areas of descriptive ontology could supplement each other even though they differ in regard to points conceived of as

important. In particular, it is possible to discern some parallels between investigations initiated by Simons and Link, inasmuch as in both cases different part-whole relations are used to characterize ontological domains. Taking into consideration the variety of semantic phenomena connected with concepts of part and whole (cf. Winston/Chaffin/Herrmann 1987), it becomes apparent that in a certain way natural language semantics, and particularly lexical semantics, depends on a profound analysis of these ontological concepts.

4 The cognitive approach to ontology and semantics

However, what I have said up to this point is not the whole story. If we have to distinguish between the way we conceptualize the world in our everyday life and the way the world really is then it should be legitimate to ask questions like these also:

- (3) What are the basic categories in terms of which the mind conceptualizes the world?
- (4) How is natural language related to this faculty of the mind?

To give an answer to such questions is the mission of cognitive science which deals with the explanation of human mental faculties underlying the acquisition, representation and use of knowledge (cf. e.g. Fodor 1981, Gardner 1987, Dölling/Dölling 1988).

One of the fundamental assumptions of philosophers, psychologists, linguists and computer scientists contributing to cognitive science is that the world as experienced is not independent of the human mind. Instead, contrary to naive common sense belief, this world has to be conceived of as the result of an interaction between environmental input and certain mental principles that impose structure on that input. More generally, in analogy to ideas of Kant, the ontology of the commonsense world is determined by a "categorical framework" which is a part of the genetic inheritance of human beings and, as such, universal. For obvious reasons, parts of the framework are inherently connected with experiences of our bodies and their interplay with the environment. In this context, the question of why it is that just this framework should have arisen as it did, can be answered by reference to the phylogenesis of man.

Although the inventory of the categorical framework and its architecture are still unknown in many respects, one may assume that it provides the basis to subdivide (a) entities in universals (types) and particulars (tokens), (b) universals in kinds, attributes (or properties) and propositions, (c) particulars in objects and situations, (d) situations in events and states, (e) objects in physical and non-physical (social and mental) objects, (f) physical objects in complex and mass objects (g) complex objects in things and configurations, (h) mass objects in stuffs and aggregates, etc. Given its *a priori* character, the knowledge involved in the categorical framework is "epistemically supreme": all propositions which conflict with it are rejected as including a

"category mistake" and, therefore, as being senseless. In this way, ontological (or categorial) knowledge determines what we are prepared to accept as knowledge about the commonsense world at all.

Following a suggestion of cognitive science, I assume that our conceptual activities are organized by a central cognitive system called *conceptual system*. I assume that each of the supposed categories can be identified with a set of principles that forms the basis of a separate conceptual subsystem, which for its part represents an ontological domain. In particular, these conceptual modules contain specific intrinsic and extrinsic constraints which are imposed on all elements of the corresponding domains. Accordingly, strictly speaking, the domains which make up our commonsense world should be conceived of as projections of the afore-mentioned subsystems. More specifically, the entities of the various domains are projected by conceptual structures formed in accordance with the principles and rules of the conceptual system.

Let me briefly look at the place of natural language in this framework. Given that language is a peculiar mental faculty of the human species, it has to be understood likewise as a cognitive module which realizes our linguistic activities in union with other cognitive systems. Obviously, the problem of the meaning of linguistic expressions, or more generally, of the information that language conveys is intimately connected with the conditions of the conceptual system. However, there are currently several proposals for cognitive semantics. Particularly, the approaches proposed by Jackendoff (1983), Lakoff (1987) and Langacker (1987) have gained recognition. In this paper I cannot discuss these proposals in more detail. In accordance with Bierwisch (1988), I will rather make the following assumptions: Every expression is equipped with a semantic structure which fulfils an intermediate function between its morpho-syntactic structure and the conceptual structure assigned to it as a result of its final interpretation. This means that there is no direct relationship between language and the world; rather, natural language expressions are first related to conceptual structures, which for their part are then related to the world. (Compare also the detailed analysis of linguistic expressions referring to spatial objects and their dimensions in (Lang 1990) for an approach adopting these general assumptions.)

Following my previous observations, it should be clear that the level of semantic structure has to obey certain conditions which take their origin at the level of conceptual structure. In particular, the semantic categorization of expressions is not only regularly associated with their syntactic categorization but above all also motivated by the categorial framework embedded in the conceptual system. In this way, the ontological distinctions induced by the latter are reflected in the division into semantic categories. To discover the effects of this

relationship is one of the tasks of semantic investigation.

5 Some aspects of natural language ontology

In this section, I will specify a few points concerning the commonsense ontology underlying the semantics of nominal expressions. To begin with, recall that the syntactic category of nouns is by no means semantically homogenous. Above all, we have to distinguish between proper names and common nouns. Whereas the former carry out the semantic function of individual terms and, therefore, seem to induce no problems for analysis the status of the latter is obviously more intricate. Investigations of mass nouns like *gold* or *cattle* originally showed that such common nouns perform at least two semantic functions. On the one hand, a mass noun is used to refer to a kind, which allows it to be analysed as a proper name of this kind. On the other hand, it has a predicate interpretation under which it denotes the set of objects which belong to the relevant kind. However, in correspondence with recent results (cf. Krifka et al., to appear), this approach can be extended to count nouns like *linguist* or *committee* in the same way. Accordingly, although there are some syntactic peculiarities in languages like English or German, such common nouns cannot only have a use as predicate of objects; rather, they may have a kind-denoting interpretation too. In this sense, every common noun belongs to two separate semantic types of expression: the type of individual expressions and the type of first-order predicates.

Supposing that the two possibilities of interpretation do not reflect a simple semantic ambiguity of the nouns in question, one has to clarify the systematic relationship between them, in particular, the question of whether the referential use or the predicate use of common nouns is semantically primitive and, therefore, lexically determined. In my opinion, this is the point at which the cognitive approach comes into play. Following the considerations discussed above, I suggest that a kind as a projection of conceptual structure is cognitively and hence ontologically prior to objects, which are specimens or instances of this kind. This follows from the fact that in order to identify an object instantiating a kind, we have to relate this object to the corresponding kind, while we can identify a kind independently of whether we refer to a special instance or not. In this context, a predicative sentence such as (5) should be interpreted as (6):

(5) *Jack is a linguist*

(6) "Jack is an instance of the kind of linguists"

Therefore, it seems to be clear that the kind-referring use of common nouns is basic and their use as a predicate of instances (and, being only mentioned, as a predicate of subkinds) of the relevant kind is derived from it.

In accordance with this, the ontology underlying natural language comprises two domains

of entities: the domain of kinds and the domain of objects. Both are projections of conceptual subsystems which involve constraints on the elements of these domains. Semantic postulates, as descriptions of such constraints, characterize the part-whole structure internal to each of the domains, as well as the instantiation relation holding between entities contained within them.

However, semantic restrictions on predicability indicate the necessity of allowing for additional ontological differentiations reflected by common nouns. Following earlier suggestions, we may distinguish between further semantic sorts of nouns. So, for example, nouns like *tree* or *house* refer to things, nouns like *gold* or *water* to stuffs, nouns like *cattle* or *furniture* to aggregates and nouns like *heap* or *stack* to configurations. Since some nouns also reflect the specific area of human beings and their social relationships, it should be evident that the hierarchy of semantic categories exhibits corresponding distinctions too. Accordingly, there are semantic sorts which contain nouns of persons (*linguist, killer*), social groups (*committee, family*) and social institutions (*school, museum*), respectively.

Each of the semantic sorts is associated with an ontological domain which is included in the domain of objects and has a structure imposed by a specific conceptual subsystem. In addition, there are also intermediate relations that hold between objects that are elements of several subdomains. To illustrate this, I mention only the following: the relations of constitution between stuff and things, between things and configurations and between persons and groups, the relations of representation between persons and groups and between persons and institutions as well as the relation of physical containment between institutions and things.

Let me conclude this paper with the remark that many problems related to the topic have been ignored in the foregoing discussion. Some more detailed observations about semantic sorts and their ontological foundations are given by Dölling (1992). However, it should be emphasized that the development of a cognitively oriented natural language ontology is a task which in many respects has still to be done.

References

- Bach, Emmon, 1986, Natural Language Metaphysics. In *Logic, Methodology and Philosophy of Science, Vol. VII*, ed. Ruth Barcan Marcus / Gerald Dorn / Paul Weingartner, Amsterdam, Elsevier, 573-595
- Bierwisch, Manfred, 1988, On the Grammar of Local Prepositions. In *Syntax, Semantik und Lexikon* ed. Manfred Bierwisch / Wolfgang Motsch / Ilse Zimmermann, Berlin, Akademie Verlag, 1-65
- Chierchia, Gennaro / Turner, Raymond, 1988, Semantics and Property Theory. In *Linguistics and Philosophy* 11, 261-302
- Dölling, Johannes, 1992, Flexible Interpretation durch Sortenverschiebung. In *Fügungspotenzen* ed. Ilse Zimmermann / Anatoli Strigin, Berlin, Akademie Verlag, 23-62

- Dölling, Evelyn / Dölling, Johannes, 1988, Logik, natürliche Sprache und Kognition. *Deutsche Zeitschrift für Philosophie* 36, 1003-1012
- Fodor, Jerry A., 1981, *Representations. Philosophical Essays on the Foundations of Cognitive Science*, Cambridge, MIT
- Gardner, Horward, 1987, *The Mind's new Science. A History of the Cognitive Revolution*, New York, Basic Books Publishers
- Hobbs, Jerry R. / Moore, Robert C., ed., 1985, *Formal Theories of the Commonsense World*, Norwood / New Jersey, Alex Publishing Corporation
- Jackendoff, Ray, 1983, *Semantics and Cognition*, Cambridge / London, MIT
- Krifka, Manfred, 1989, *Nominalreferenz und Zeitkonstitution. Zur Semantik von Massentermen, Pluraltermen und Aspektklassen*, München, Fink
- Krifka, Manfred / Carlson, Greg / Chierchia, Gennaro / Link, Godehard / Pelletier, Jeff / ter Meulen, Alice, to appear, Genericity. An Introduction. In *The Generic Book*, ed. Greg Carlson / Jeff Pelletier, Chicago, University of Chicago Press
- Lakoff, George, 1987, *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago, University of Chicago Press
- Lang, Ewald, 1990, Primary Perceptual Space and Inherent Proportion Schema: Two Interacting Categorization Grids Underlying the Conceptualization of Spatial Object. In *Journal of Semantics* 7, 121-141
- Langacker, Ronald, 1987, *Foundations of Cognitive Grammar*. Stanford, Stanford University Press
- Link, Godehard, 1983, The Logical Analysis of Plurals and Mass Terms: A Lattice-theoretical Approach. In *Meaning, Use and Interpretation of Language*, ed. Rainer Bäuerle / Christoph Schwarze / Arnim von Stechow, Berlin/New York, de Gruyter, 303-323
- Link, Godehard, 1991, Plural. In *Semantik*, ed. by Arnim von Stechow / Dieter Wunderlich, Berlin/New York, de Gruyter, 418-440
- Pelletier, Jeff / Schubert, Lenhart, 1989, Mass Expressions. In *Handbook of Philosophical Logic*, Vol. IV, ed. Dov Gabbay / Franz Guenther, Dordrecht/Boston/London, Reidel, 327-408
- Simons, Peter, 1987, *Parts. A Study in Ontology*, Oxford, Clarendon Press
- Smith, Barry, ed., 1982, *Parts and Moments. Studies in Logic and Formal Ontology*, München / Wien, Philosophia Verlag
- Strawson, Peter F., 1959, *Individuals: An Essay in Descriptive Metaphysics*, London, Methuan University Press
- Winston, Morton / Chaffin, Roger / Herrmann, Douglas, 1987, A Taxonomy of Part-Whole Relations. In *Cognitive Science* 11, 417-444