

Information and Philosophy of Language

Hans Kamp, Jeroen Groenendijk, Martin Stokhof

8th March 2005

Starting point is that the chapter should address both philosophical and methodological issues as well as systematic questions concerning the nature and the role of the concept of information in philosophy of language and linguistic semantics. Whenever possible it should do treat both aspects in close connection: we will focus on systematic issues that illustrate a philosophical concern or a methodological point, and on philosophical and methodological issues that have bearing on a systematic question. This interconnection will also allow us to make the necessary choices, since it is obviously impossible to deal with all the relevant issues in the span of a single chapter.

Below we present a short sketch of relevant historical development separately from an introductory overview of systematic issues, but in the final version of the chapter we will try to turn them into a single thread. What follows is a first attempt to sketch the contours of the view that will be developed in the chapter, and as far as the materials that will be covered is nowhere near complete.

1 Meaning and information in historic perspective

To set the stage, we start with a sketch of some important stages in the *historical development* of the concepts information – content – meaning as they have been employed in the study of natural language in modern times. Likewise

in philosophical treatments of language we see a development from ‘rich’ and (fairly) concrete notions of meaning, closely tied to judgements and experience and application, to (rather) abstract notions of meaning and (ultimately) to a view of natural languages as information coding and information transferring devices.¹

A very rough sketch of the various stages in this development distinguishes the following stages. First stage (roughly at the rise of modern philosophy): here meaning is treated in close connection with epistemological concerns; cf., the ‘idea’ theories of meaning of classical empiricism and rationalism; here there are still intimate relations between the concepts of meaning, experience and knowledge, reason and justification, and there is no such thing as a separate philosophy of language.² A case can be made that this is the stage that most closely resembles a ‘common sense theory of meaning’.

The second stage is characterised by the rise of ‘meaning proper’ in the wake of the development of modern logic (Frege, Russell, early Wittgenstein). Here meaning gets more and more dissociated from epistemology (Frege’s anti-psychologism, Wittgenstein’s form of logical atomism), but remains strongly related to ontology. This is (also) due to the philosophical aims of the analysis of meaning: stimulated by the success of the development of new formal languages, the old ideal (Leibniz) of a philosophically transparent (‘ideal’) language (a formal one this time) gained new momentum. This stage also marks the rise of philosophy of language as a separate discipline: the ‘linguistic turn’ put language centre stage, and hence philosophy of language became an important and distinct undertaking. It is interesting to note that in phenomenology (Husserl, Heidegger, Ricoeur, Merleau-Ponty) a similar development took place,

1. This is not unlike Gøran Sundholm’s view on development of logic: from a theory about judgements & reasoning as psychological acts to (ultimately) formal symbol manipulation, although we do not necessarily subscribe to his negative evaluation of that development.

2. Hence Ian Hacking called idea theories ‘nobody’s theory of meaning’: since meaning as such is not as separate concern nobody has a *theory* about it.

but without the strict separation from epistemology that is characteristic for analytic philosophy.

The third stage in this historic sketch that we need to distinguish can be characterised as the stage in which formal semantics arises as a separate discipline: developments in philosophy and in linguistics come together and give rise to the idea that a formal theory of meaning can be developed and applied in the description of actual natural languages. This is the time in which Davidson, Montague, Lewis and others do their seminal work, the heyday of Montague grammar. Important from the perspective of this chapter is that meaning becomes once more a different kind of concept: meanings are first and foremost formal constructs, and theories of meaning are primarily differentiated in terms of the formal machinery one deems necessary for the description of semantic features of natural languages:³ concerns with epistemology or ontology become less and less important as semantics becomes more and more autonomous, and the nature of the concept of meaning reflects this. This stage, that is still of crucial importance for understanding our present ways of thinking, is also one in which a new and problematic relationship with psychology starts: semantics and philosophy of language develop a ‘love-hate’ relationship with the Chomskyan approach to grammar and its close alliance to rationalistic thought and computational psychology. In the wake of the third stage we see increasing attention being paid to ‘flexibility’ and ‘resource dependency’ of natural language and to mechanisms of meaning specification in context (rules for the contextual resolution of certain types of presupposition and anaphora, certain discourse particles, etc.)

The fourth stage is characterised by a growing attention for matters of information proper. Increasingly, language and language use are viewed in terms of information exchange. Stalnaker’s work is an early representative, dynamic the-

3. Slightly exaggerating one might say that we are dealing with an altogether new type of phenomena, here.

ories of meaning (discourse representation theory, file change semantics, update semantic, dynamic semantics) develop this trend further by analysing meaning itself in terms of information change. Stalnaker introduces the common ground in communication as a common resource for speaker and addressee; dynamic semantics focuses on discourse contexts as resources for specification of sentence meaning. This shifts (or blurs) the distinction between semantics and pragmatics, i.e., between what is supposed to be a matter of meaning proper and use. This shift is still continuing. Central issues here is that of information *structure* (topic, focus, presupposition, etc) as a further set of linguistic devices for linking each new sentence in a text or conversation to what went before (or to prepare the ground for what comes next). This increasing focus on information exchange and information change also weakens the link with ontology, that was secured in more traditional formal semantic theories by the central role of reference and truth: truth becomes a limit concept of a more general notion of acceptance. The dynamic turn also further problematises the connection with psychology: cf. the debate between discourse representation theory and dynamic semantics, where the very nature of the concept of meaning is used as an argument in (or is supposed to be determined by) a particular view on the relation between psychology and the theory of meaning.

The current stage of thinking in philosophy of language and semantics about meaning is one of diversity: there seems to be no one dominant conception of meaning. Partly this is due to the fact that while thinking about meaning in the ‘mainstream’ developed along the lines just sketched there have been an number of often quite outspoken and successful ‘counter-currents. The following two examples illustrate also represent some of the systematic issues involved.

The first is a trend toward the development of what is taken to be a more psychologically realistic notion of meaning. Several internal issues in formal semantics (in particular in possible worlds semantics), such as direct reference and logical omniscience, make a psychologically realistic view on the resulting

concept of meaning implausible (if not impossible), and that, in its turn, has given rise to theories that deal with meaning in terms of a distinction between ‘broad content’ and ‘narrow content’, (cf., the debate between externalism and internalism) which would allow us to circumvent these problems. This goes against the general tendency of anti-psychologism. It should be noted that this ‘return to psychology’ has caught on mainly in philosophy of language,⁴ less so in semantics of natural language.

Another example is that of situation semantics and situation theory that one can regard as (among others, since there are also other motivating factors) an attempt to restore relation between meaning and its naturalistic determinants. Viewed from that angle we can look upon situation theory as an attempt to restore the traditional connection between philosophy of language, epistemology and psychology of a particular bend, viz., naturalistic and empiricist psychology. Finally, although current thinking about meaning and information is diverse and in flux, it *is* true that there is a continuing tendency toward the use of notions of meaning and information that are further distanced from what we could call a common sense notion: the rise of quantitative, statistical notions of information in combination with the use of techniques from in (evolutionary) game theory, for example, is a definite trend, that fits in quite nicely with the use of ‘shallow’, non-rule based techniques in NLP, information retrieval, semantic web, and so on.

The end result (for now) might strike some as perhaps somewhat paradoxical: on the one hand all these abstractions have led to success⁵, on the other hand there is a continuing abstraction away from what meaning is as an everyday

4. According to some it has even transformed philosophy of language into a branch of philosophy of mind.

5. Which in fact gives rise to a serious methodological question, viz., what the nature of that success is: *what* is it that current semantics and philosophy of language are successful in? What are the measures of success here? Are the (relatively) theory independent? What do they apply to? And so on.)

concept (Wittgensteinian concerns here). It seems reasonable (and it actually is quite fashionable) to come up with the following conjecture that is supposed to alleviate any worries in this area, viz., that neuro-cognitive research will serve as a bridge. The idea is that if the neuro-cognitive reality of the current ‘information approach’ could be validated by empirical research, then that would *show* that it was the right road after all. Maybe, but there are similar concerns here as there for semantics proper: the relationship between neurological correlates of concepts and these concepts themselves in some cases is actually quite problematic, and it is concepts such as meaning (intention, etc) that seem particularly vulnerable in that respect.

2 The conventional nature of natural language meaning

This section gives the outlines of a treatment of systematic issues, that often have been operative in bringing about the changes and developments outlined in the previous section. The starting point is that natural languages are designed to represent and communicate information. They are able to do so because their well-formed expressions are (normally) meaningful. Well-formed expressions have the meanings they have because of the conventions which govern and define them. These conventions take a number of distinct forms. Understanding how linguistic expressions succeed in carrying information is largely understanding of what these conventions are like. And a close study of the contents and function of these conventions reveals some of the mechanisms and structures that shape our linguistic, cognitive abilities.

2.1 The conventional meaning of words

2.1.1 Words linked to perceptually accessible aspects of the world

The following is a widely accepted assumption: By the time that word meanings are put in place (i.e., acquired by an aspirant member of the speech com-

munity, the perceptually operational concept is already in place. Exactly what this assumption amounts to is a more difficult and more controversial matter. Minimally we should be committed to this: by the time the language learner acquires the word she is already able to represent that some part of the world, which she is capable of independently identifying and re-identifying, has certain features, and thereby retain the information that the world has these features when it is not accessible to her inspection (and then, in case the possibility of inspection is restored, recognise that the representation still fits or else notice a change. More precisely, it is normally certain parts of the world, which the learner can identify and re-identify, which are represented as having the given features, and recognition whether the representation still applies arises only when that part is re-identified and inspected. Representing a part of the world as having given features is to represent it as satisfying a certain concept. It is to such concepts - which the learner already 'has' in the sense of being able to represent parts of the world as satisfying it in the sense alluded to - to which words can then be attached as phonologically defined labels.

Such pairings of concepts and labels are conventional in that in principle any phonological string could serve as label for the given concept. However, the language learner must learn conventional pairings that are already in place. That is, she must pair each given phonologically identified word that she realises to be part of the language she is acquiring with the same concept as those who already know the language. This presupposes equivalence of concepts and concept representation between the language users (whether accomplished or acquiring). Concept equivalence requires at a minimum that members of the community recognise by and large the same parts of the world as satisfying the same concepts. In practice it is not easy for them (or for anyone else) to establish that this is so unless a common language enables them to verify if they do (cf. Wittgenstein's private language argument), though there are at least some concepts for which language-independent evidence is possible that

members of a community share them.

This raises a number of questions:

1. What more can be said about the nature and implications of the cognitive capacity to represent perceptually anchored concepts? Can this capacity take different forms, and what are these? if there are different forms, which are a precondition of acquiring the meaning of the corresponding concept word?
2. Does the acquisition of a concept word really presuppose the ‘having’ of the concept (i.e. the capacity of representing parts of the world as satisfying it)? Or can learning the word (in the sense of learning to use it) lead to the capacity of representing parts of the world as satisfying the concept? And (cf., Whorf) are or could there be perceptually operational concepts which would not have been acquired but for the given word? Or, more strongly even, are there concepts that could not be acquired except by those who share a language in which there is a word for the concept? Do the answers to these questions perhaps depend on the kind of concept we are dealing with? (Psychological research suggests that concepts are more easily acquired when learnt in combination with a vocabulary. And certain scientific concepts actually seem to depend for their acquisition on the presence of an underlying theory.)
3. Does the acquisition of a certain class of concept words force a certain structure on the concepts that were antecedently available, in that the concepts for which the learner now has words become more salient and tend to push other concepts into the background when information is represented once this class of words has been learnt? How much do individual natural languages differ from each other in this respect: How much do the conceptual structures they make salient for their users differ as a function of their different vocabularies? (Whorf again)

2.2 Words for action concepts

A special role is played by concepts that classify actions – one’s own actions and those of other individuals who are recognised as agents. Action classification, i.e., knowing what it is one is doing or going to do, must be present at the earliest stages of cognitive development and activity. (It is a necessary and inalienable part of intentional action of any kind.) Thus the cognitive preconditions for acquiring many concept words will be in place when vocabulary acquisition gets under way.

Questions in this area are:

1. To what extent is classifying intentional actions a matter of perception? To what extent is it a matter of observing the actions one performs oneself to have the intended properties or to produce the intended states or results?
2. How important (or essential) is it to action classification and to learning the corresponding words that the classifier/learner be able to recognise the actions of others as falling under the same concepts as her own?
3. Words that are neither directly linked to perception nor to action. We return to these after having said some things about compositionality.

3 Concept compositionality and the linguistic conventions of word combination

To represent an identifiable part of the world satisfies a certain concept the representation must also contain some element that serves as identifier of this part, and that will enable the representer to tell that the part is once again available for inspection if and when that happens. This entails that the representations of which we have been speaking must at a minimum involve two distinct constituents, the represented concept and a constituent which identifies the part of the world which is represented as satisfying the concept. In many such representations the ‘part of the world’ will be what in a more articulated ontology will

qualify as an ‘individual’. In such cases the representation constituent which serves to identify the ‘part of the world’ can be thought of as a ‘designator’, or ‘name’, which the representer has for the represented part or individual. In such cases the representation as a whole can be thought of as taking the form of an elementary predication, to the effect that the designated individual satisfies or instantiates the concept. This is the most elementary form of compositionality in representation. It is genuinely compositional in that its constituents can be reused in other representations, which either represent some other part of the world as satisfying the same concept or represent the given part as satisfying some other concept.⁶

When designators of individuals or other parts of the world are reusable in other representations, they are available for linguistic labelling: If the individual is publicly available in the sense that it is identifiable and re-identifiable by other members of the community as well, and if identifications can be shared in the sense that several members (e.g. the learner and one or more other community members) can come to share the conviction that they are all (re)identifying the same individual or other sort of world part, then it should be possible for the learner to also acquire a linguistic designator for that individual or world part. And once both a word for the concept and one for the world part are in place, there arises the question how to represent the part as satisfying the concept: can be expressed by combining the two words - is the concept word to be placed before or after the part designator, or are they to be combined in some other way, with some further phonological material to serve as ‘grammatical glue’? Here we encounter the most elementary form of linguistic compositionality, and the ways in which individual languages resolve this question can be seen as involving the most elementary manifestation of linguistic convention at the compositional level.

6. An issue that needs to be addressed here is how this concept of compositionality sits with the distinction that is often made between compositionality and systematicity.

Moderate extensions of this most elementary form of compositionality are possible, and can be expected to arise as a matter of virtual inevitability, in various directions. First, once a way of representing individuals has been put in place and the awareness as been established that some of the represented individuals can move about, thereby giving rise to world parts which vary as regards which individuals they contain as subparts, it should be possible to represent such subparts as involving two or more of these separately representable individuals. Concepts that are satisfied by such world parts can take on the status of relational concepts in that it is only such world parts, with more than one individual as subpart, that can satisfy them, and that it is only in virtue of two or more of the individuals contained in the part that the concept is satisfied. Thus the given cases of concept satisfaction can be represented by combining the concept with designators for the two or more relevant individuals in the satisfying part. This, it is reasonable to assume, is one source for the origin of relational concepts.

Once more we encounter the problem of representational form, and that both at the level of cognition and of language. The matter of the form of representation is particularly for non-symmetric relational concepts. In such cases the designators for the different individuals who jointly satisfy the concept must be represented in a certain order. In the case of language the issue is once again one of linguistic convention: What is the right order for the three or more words needed to express the satisfaction of a relational concept, and what if any is the additional grammatical glue that different human languages require for this purpose?

A second extension consists in representing one and the same world part as simultaneously satisfying two or more concepts. Such representations require some basic form of conjunction, and the same applies to linguistic expressions of such conceptual conjunctions.

A third extension is at least as important as the first two. We have mentioned

concepts which serve to classify the state of the world as it is perceived and concepts for the classification of actions. It is part and parcel of the earliest needs and uses of cognition to combine these concepts: The developing agent must avail herself of hypotheses as to what sorts of effects certain actions have on the situations in which they are performed – hypotheses, in other words, according to which an action of a certain type A, when performed in a situation of type C1, will necessarily or probably result in a new situation of type C2.⁷ The representation of such hypotheses is considerably more complex than those mentioned so far. It will have to consist of a representation of the action type A together with two representations of world parts, as satisfying the concepts C1 and C2, respectively; and these two representations must be distinguished as initial and result state. Such representations display a new dimension of complexity even in the simplest cases where C1 and C2 are non-relational concepts. The new representation also differs from those considered up to now in that it is the representation of a general rule; so the representations of initial state, action and result state of which it is composed must all be ‘schematic’ (in a sense that requires more careful analysis). When such hypothesis representations are to be expressed in language, further questions of linguistic convention arise. And here we reach territory that is more familiar from a linguist’s perspective. Typically, and perhaps universally, action concepts are verbs, while those used to classify individuals and other parts of the world are more commonly adjectives and nouns. The rules (of any particular language) for putting the different expressions that are needed in the expression of such a hypothesis will instantiate what the grammar has to say in general about the combinations of words of these different categories, including additional grammatical ‘wrapping material’ that individual languages may require.

In addition to the schematic representations of such general hypotheses the

7. Thus it is assumed that an adequate representation of such actions does not involve a representation of their effects.

agent should also be able to represent particular episodes as supporting instantiations of them. (Or, alternatively as episodes which contradict the hypothesis.) Such a representation must consist of (a) a particular world part or individual satisfying C1, (b) a particular action as satisfying A; and (c) a particular world part, often the very same as the one mentioned under (a), as satisfying C2.

Such representations represent the initial state and result state as successive in time. This is one way in which the concept of time enters into cognition and into language. And more particularly, it introduces the concept of temporally dependent predication. For especially in those case where the individual or part of the world that is involved in the result state is the same as that involved in the initial state, the representation as a whole must convey that it satisfies C1 at the earlier time of the initial state and C2 at the later time of the result state.

Once this much is in place, the possibilities for introducing further combinatory devices, at the level of cognitive representation and/or that of language, multiply. All such further devices require their own syntax (i.e. representational or grammatical form) and their own ‘semantics’. Especially questions of the semantic information carried by such further representation forms constitute a range of difficult and varied problems, where the indirect connections with observable reality and questions of inferential content (the information of a given representation is to be assessed in terms of the representations that can be inferred from it) will have to play a central role. It is also in this connection that the many different types words will have to be given a closer analysis which are neither designate individual world parts nor denote concepts of such parts, nor concepts of the kinds of actions we have so far talked about.

Along these lines cognition may develop a rich repertoire of structured representations of propositional content, where the propositions say in more or less direct ways what particular past or present parts of the actual world or of alternative fictitious worlds, are like and of how some of these possibilities are

systematically connected. Both because representations of this second - generic or schematic - kind play such an important role in cognition and because some episodic representations involve concepts which are not operationally linked in any direct way to observation, inferential relations between representations and devices for inferring representations from others are, for all we know, a crucial ingredient to identifying what the information is that the more complex representations carry. (Here, there is an important need for the different conceptions of meaning (and especially natural language meaning) - model-theoretic, inferential, prototype-based, ... - will have to be lined up and compared!)