

The dynamic nature of meaning

Claudia Arrighi

Philosophy Department, Stanford University, CA, USA
carrighi@stanford.edu

Roberta Ferrario

Laboratorio di Ontologia Applicata, Consiglio Nazionale delle Ricerche, Trento, Italy
roberta.ferrario@loa-cnr.it

Introduction

This paper wants to be a contribution, mainly of philosophical character, to a new current of thought and research in semantics that have been defined the “dynamic turn” in the study of meaning (cf. Peregrin 2003). The central point of this approach to meaning is that there is not such a thing like a precise literal meaning of a word or expression, but meaning is something that gradually evolves from the dynamic processes of communication. This dynamic character is not some kind of secondary aspect of meaning, but it is instead a fundamental feature and we believe that it should receive more consideration into philosophical theories on meaning. What we are going to do is to stress the importance of this dynamic character and to make few steps towards a philosophical approach to meaning more focused on processes that shape meaning, instead of focusing on definitory issues.

Even though the dynamic aspect of meaning may seem obvious, strangely enough it has been mostly neglected in philosophical theories. When dealing with meaning, some of the main currents in philosophy of language tend to focus on defining what meaning *is*, looking for something that can be circumscribed and pinned to a specific word; dynamics doesn't seem to be the main issue in these theories. We can also find a different tradition, the origin of which has been universally connected with Ludwig Wittgenstein's later work; this is engaged in the effort of showing a strong relation between what meaning is and what is the use we make of it, its relation with social interactions. Unfortunately, this approach to meaning still struggles to find a philosophical account that has not been accused of being 'foggy' or mere 'hand-waving'.

On the other hand, if we look outside of the philosophical circle, we can see that recently a lively debate related to meaning has animated computer scientists: the

artificial agents they are dealing with don't always have access to a common and conveniently labeled ontology, they have their own representations of things, and maybe their own names, but they have to 'communicate' things to each other: how can this be done, practically? One of the solutions proposed is to imitate a way in which often humans manage to understand each other, i.e. by negotiating the intended meaning of words and expressions while engaged in conversations. How come that philosophers don't talk about meaning negotiation? It seems thus that the study of the processes of meaning negotiation is one of the points in which the studies in philosophy and in computer science can merge and therefore obtain reciprocal enhancement.

In this paper we would like to show that meaning cannot be defined independently from the practices in which it is used, but it can be more efficiently described as tool used to pursue specific linguistic tasks. And, just as could happen to workshop tools, in the long run it itself is shaped by the many jobs done. We will consider the meaning of a word as a variable, the value of which slowly changes as a consequence of many individual negotiations of reference between speakers. The main idea is that repeated negotiations of a term in a social community *shape the meaning* of such term, especially the meaning attributed to it by each individual agent, but also the one that has been called the "literal" meaning. To accomplish this, we need to reformulate the idea of negotiation taken from AI in a way that can be applied to real human communication; to this end we will refer to the work that Herbert Clark has done in psychology on collaborative processes in using language.

We start in section 1 by giving a brief account of the different approaches that have been historically developed in philosophy on the analysis of meaning and we single out four of these approaches, from which we try to extract interesting hints. Then we give a short survey of what has been done in computer science. In section 2 we look at the insights given by a theory of language use as a joint activity, while in section 3 we present our proposal and in section 4 we sketch some possible future developments.

1. Looking for a definition of meaning

Main theories in the philosophy of language

The idea of a tight connection between meaning of words and the dynamic of social interactions has never been denied, but theories of meaning and reference that rely heavily on this connection have been considered in some kind of opposition to the other

main stream in philosophy of language. This is well summarized by Jaroslav Peregrin (1995) at the beginning of a review of Robert Brandom's book "Making it explicit":

The philosophy of language of the present century can be seen as dominated by two contrary tendencies. The first of them is to take language to be a kind of nomenclature, to take the word-meaning relation to be a basic and irreducible fact (either of the causal kind, or of a specific kind peculiar to 'intentional' mediae of representation). The second is the tendency to view language rather as a toolbox; and to take the 'intentional', representational capacities of words to be parasitic upon their involvement with human activities. According to this second view, 'to have meaning' is to play a certain role within the structure of human conduct and within the social institutions which regulate it. (Peregrin 1995, p. 84)

The second tendency described by Peregrin is the one we are going to follow. In order to do this, we feel compelled to compile a brief overview of the attempts made so far to define meaning, just to sketch the landscape that is framing this contribution. This is a list of what seem to be the essential features of the main traditional theories on the nature of meaning. They can be roughly summarized by the following slogans:

- the meaning of a word is the denoted object;
- meaning is something in the speaker's mind (idea, image, concept, intention);
- meaning is a set of conditions of satisfaction;
- meaning is determined by the use in a linguistic community.

The first position, taken literally, would assume a one-to-one relationship between words' meaning and entities in the world. More sophisticated and elaborated versions of theories of direct reference (cf. Kripke 1972) are able to account for more realistic relations many-to-many (the name "Mary" doesn't apply to just one thing in the world, and a certain object doesn't necessarily bear just one name), both in the case of correspondence of common names and "natural kinds" and in the case of proper names and single objects; nevertheless the central claim of these theories is that the meaning of a word is its denoted object. One of the main problems of this kind of approach has been underlined by the famous Fregean example of the morning and the evening star, that we rephrase here with a more prosaic example in everyday life: the words "entrance" and "exit", used to denote a unique door. In the example, one could be tempted to say that the two words denote the same object, but they mean different things, leaving space for an intermediary between word and object: from here the rise of the Fregean "sense". The merit of this position has been to underline the necessity of a connection between meaning and reference, i.e. between language and the world, but this relationship cannot be easily stated as a simple identity.

The second position in the list stresses the importance of an intermediate component between words and objects, and characterizes it as a subjective point of view on the object denoted. We use words to refer to things that we have in mind, so we could identify the meaning of a word with the mental content representing the object denoted. Let's take again the example of the door: whether this is an entrance or an exit depends on where I am and how that door is represented in my mind in a given moment. Or we could formulate the same idea using an intentional vocabulary: the door is an exit if I indicate it with the intention to indicate a way out from somewhere. But even this position presents hard problems. The main one is that in this way we have lost the connection to things in the world. If meaning is definable with what is inside one's head, how can we exchange thoughts with others speakers? A possible way out is that of assuming that the same word is simply connected to the same set of thoughts in every person. But if we assumed this, then we wouldn't have any misunderstanding in communication (clearly false), and if we don't assume this commonality, how is any communication possible at all given that we can't see directly in each other's mind?

The third position is one that characterizes the meaning of a word as the set of conditions to be satisfied by an object: whatever object satisfies the conditions, is the object denoted. This position has some useful features: it's able to connect the linguistic expression with states of the world, but without a one-to-one relationship; it doesn't need to be subjective, because the conditions can be the same for everyone (we could say "an exit is a door that takes outside with respect to the speaker, whoever he/she is"); then this definition doesn't identify meaning with a kind of entity, but rather with a procedure to check the relevant conditions. This view is really promising, but how to choose the relevant conditions to be satisfied is an open problem. The attempt to build a set of conditions defining the proper use of a word has proven to be hard to accomplish; the aim has been mainly to select a set of conditions individually necessary and jointly sufficient to describe the word, but it is often impossible to reach an agreement on such a set of conditions. The conditions approach could maybe rely on the help of pragmatic procedures to specify what kind of conditions have to be satisfied to accept a certain utterance, from a specific speaker, in a specific context. But it doesn't provide a convincing tool to describe general semantic knowledge.

In the fourth position there is an acknowledgment of the impossibility to set a fixed frame of conditions to specify what a word denotes, and instead it points out that the meaning of an expression consists in the actual use that can be made of that

expression in social interaction. The spirit of this idea was suggested by Wittgenstein in the *Philosophical Investigations* (1958): one of the key paragraphs where Wittgenstein suggests this view is §197, where he, talking about understanding a word, says:

It becomes queer when we are led to think that the future development must in some way already be present in the act of grasping the use and yet isn't present.
– For we say that there isn't any doubt that we understand the word, and on the other hand its meaning lies in its use. (Wittgenstein 1958, §197)

Following what seems to be his train of thought, if we consider the literal meaning of a word as a fixed set of things (rules, conditions, features), then we imply that when we learn that word, we are including in this knowledge any possible future usage of the word. According to Wittgenstein, something is weird in this view, and he promotes the idea that, in the end, to understand a word is the same thing as being able to use it properly, it is a practice; we don't need to be able to describe it in a finite and precise way, and maybe we wouldn't be able to do it (as for the practice of riding a bike). The problem with this view is that it doesn't provide a way to determine the details of a representation of meaning; just saying that the meaning of an expression consists in its use doesn't provide any information to represent this kind of knowledge. What is interesting in this view, however, is exactly what seems to be the source of its vagueness: if we want to determine a meaning, we have to look at the dynamic of linguistic practice, and not just at static rules and definitions. In this way this approach underlines the dynamic nature of language, and its social component. From this we could infer that, when looking for a representation of meaning, we can just represent the basic structures, the guidelines, the processes that we need to start and get involved in the practice of language.

Another potential problem in considering meaning as something subject to social practice is that it would become a complete relativistic notion, different every time, something that cannot be decontextualized, therefore we could never talk about the *literal* meaning of a word. But the possibility of drawing a distinction between literal meaning and speaker's meaning is an important requirement, orthogonal to the positions about what meaning is. This distinction allows us to say that a certain sentence has a meaning, independently of any actual use: if I consider the sentence "the ball is green" without any other information, I can say that it has a meaning. In virtue of this distinction we can talk of an abstract literal meaning of a sentence as opposed to the specific referential function that the sentence is performing in a given context. In this way we can also account for the fact that the speaker is free to use the sentence, in an

appropriate context, to perform an illocutionary act or to indicate a referent different from the literal meaning, but in order to do this we are supposed to know the a-contextual literal meaning. If we want to support a position that relates meaning with language use, we have to rethink this speaker's - literal meaning distinction.

We'll come back later to this issue; before doing that, we'll try to extract the most remarkable features from the approaches just presented.

Harvesting the best from the traditional theories

If we try to get the good qualities from the four positions on meaning that we have mentioned, but at the same time we try to avoid the major problems, we would have to draw a theory of meaning that has the following features:

- it must be able to identify a referent, but must not tie a meaning to a particular referent;
- it must be related with the mental content of the speaker, but it must not be tied to this one either;
- it must take into account the conditions of satisfaction, but doesn't consider them as fixed;
- it must consider important the variability of use of a linguistic expression, but has to clarify the constraints and boundaries of this variation;
- finally, it has to give an account of the intuitive difference between a literal and a speaker's meaning.

This last point is particularly important: even if we are claiming that there is not a fixed meaning for a word, the classical difference between literal and speaker's meaning has to be accounted for.

The draft of a theory that meets all these requirements is twofold: on the one hand, it describes how the private meaning is formed as a mental state of individual agents and, on the other hand, it shows how a public meaning emerges from communication among agents and social practices; more than this, we would like to trace at least the outline of the structure of these "social practices".

Some interesting attempts in this direction have been made in computer science; even if originated by somewhat different motivations, these studies can be taken as inspiration for an analogous analysis in philosophy.

Theories of meaning in computer science

As already noted, theories of meaning in computer science are originated as attempts to give solutions to concrete problems of the everyday practices of storing and managing data. Many of these problems come from the so-called semantic heterogeneity, namely the diversity of meaning attributed to names and concepts by different information sources and users. More concretely, very often there is the need of merging information coming from different sources, that use different criteria to store and classify information.

If we want to merge information or to use information coming from different sources, we need a tool able to create communication among heterogeneous sources, such as databases using different schemas, document repositories using different classification structures, users' file systems etc.

In literature two main approaches have been proposed: the first is based on the creation of a shared model, a kind of "frame of reference" in which the concepts belonging to the different sources should be "translated". Even though this approach has proven useful in restricted or very specific domains, where the different parties seem to have similar goals and needs, it is less effective in open and dynamic environments. In order to deal with these, a second approach has been developed, that doesn't assume the presence of a shared model, but is based on a "peer-to-peer" philosophy. According to this approach, each peer keeps its own schema or conceptualization and they manage to communicate through two complementary processes: a process of *meaning coordination*, which is an attempt to find mappings between the meaning of a collection of expressions, and a process of *meaning negotiation*, that takes place when a direct mapping is not available and has the purpose of solving semantic conflicts among parties.

Now, if we try to draw a parallel with human language, the first, "centralized" approach can be seen as the process of compiling a dictionary or creating the frame of reference needed to account for the different uses of some words. In order to do this we must assume that some sort of common ground is already available, that there is a position from which the different uses can be observed and collected together. Instead the second, "distributed" approach is more similar to the way in which minor divergences in meaning are settled in everyday usage of language within a social community of speakers. We can have any kind of difficulties in understanding each other in many occasions, and often we have to solve the problem on the spot, with no access to an already established common ground. From a philosophical perspective, this is the

situation that is closer to the Quineian problem of *radical translation*, and this is also the dimension of language that is central in our approach, as we discuss in the next section.

2. Language use as a joint activity

The work of the psychologist Herbert Clark presented in his book *Using language* (1996) is one place where the connection between language and social practices has received a detailed and thorough formulation. In what follows, we summarize his reasons to claim that language use is a form of joint action. Then we outline Clark's general idea of joint actions and activities and we formulate an example to illustrate how the properties of an expression to "mean something" and to "refer to something" are properly described as special kinds of joint actions of speaker and listener .

At the very beginning of Herbert Clark's book, *Using language*, the author states his thesis:

Language use is really a form of joint action [...] A joint action is one that is carried out by an ensemble of people acting in coordination with each other. [...] When Fred Astaire and Ginger Rogers waltz, they each move around the ballroom in a special way. But waltzing is different from the sum of their individual actions [...]. Waltzing is the joint action that emerges as Astaire and Rogers do their individual steps in coordination, as a couple. Doing things with language is likewise different from the sum of a speaker speaking and a listener listening. It is the joint action that emerges when speakers and listeners – or writers and readers – perform their individual actions in coordination, as ensembles. (Clark 1996, p.3)

One of the main representatives of this tendency in philosophy of language is Paul Grice (1969), and Clark refers to his ideas a great deal, in particular regarding his concept of speaker's meaning and m-intention.

Grice's m-intention –the heart of speaker's meaning– is a curious type of intention: it is one the speaker cannot discharge without the audience's participation [...] I can discharge my intention to shake a stick, an autonomous action, without anyone else's actions. But I cannot discharge my intention to do my part of our hand shake, a joint action, without you doing your part. (Clark 1996, p. 130)

So, as a result of this character of the m-intention, the act of meaning something is what Clark calls "a participatory act in a joint act". In this way, Clark can formulate the principle he wants to defend:

Signal recognition principle: signaling and recognizing in communicative acts are participatory acts. (Clark 1996, p. 130)

where he's using "signaling" to indicate the speaker's action, and "recognizing" to indicate the listener's action. Together, they are participatory acts which constitute a "communicative act", the joint activity of communication.

In his book and in other articles, Clark takes in particular consideration the act of referring, describing how referring is a collaborative process (cf. Clark and Wilkes-Gibb 1986 for a specific paper on the topic). He highlights the single steps of this process, that resemble very much the steps of a negotiation. We are going to give an example of this process that exemplifies some of the features pointed out by Clark.

In our example, Mary and Bob have a very general common goal, "keeping each other company", that doesn't necessarily need to have a linguistic component. During the pursuit of this goal, Bob comes out with an individual subplan, namely "to exchange opinions on a bike", and decides to present it to Mary as a candidate for a joint project. The presentation of this joint project, and the eventual achievement, can make very good use of language as a tool. Hence, Bob says: "Isn't that bike nice?" (let's call this the utterance *s*). Bob has proposed this utterance as an opening of the sub-project 'exchange opinions on a bike' on the basis of, at least, the following assumptions of common ground:

- i) he and Mary both speak English
- ii) they both have access to the same visual field
- iii) they both are seeing a bike
- iv) they both recognize 'that bike' as salient.

It is now Mary's turn to take up or reject the project. Before that, however, she is engaged in the sub-sub-level joint activity of understanding Bob's utterance. This is what Clark calls principle of joint construal:

Principle of joint construal: for each signal, the speaker and addressees try to create a joint construal of what the speaker is to be taken to mean by it. (Clark 1996, p. 212)

Reaching a joint construal equals the mere linguistic problem of reaching convergence on the referent of an expression. To obtain this result, we need to engage in a specific kind of joint activity: this is the kind of joint activity where part of language is not just a tool, but is the product of the activity. So, what Clark calls "problem of joint construal", we are going to refer to as the "joint reference problem" .

Let's go back to the example: if Bob is using their common ground correctly, the joint reference should be no problem at all, and if Mary goes ahead with a contribution to the "exchange opinions" project, they assume to have a joint referent, until otherwise proven. But in this case Mary has a problem with it, she doesn't know what Bob is referring to with "that bike", and so she proposes to solve this construal problem by asking: "Which bike?". The mistake is in Bob's assumption iv) about common ground, because there is not a clearly salient bike for Mary. Bob realizes this and answers "The green one", uptaking the "reach a joint reference for *s*" project, and proposing a correction. Now it is Mary's turn, and she says "Oh, I see...". In this way she accepts Bob's correction, and complete the "reach joint reference for *s*" project. Now she can go on and give her contribution in the project "exchange opinions on a bike", that has been suspended. She does so by saying "Well, I don't really like that bike", and the second level project can also be considered concluded, unless Bob or Mary disagrees on the exit point, and makes another contribution, or opens another sublevel project and so on.

With this short example we want just to show how the process of *referring* to something is deeply connected with processes of interaction, collaboration, and this is also the main idea in the work by Clark that we have considered. But Clark himself traces precise boundaries to his goal. In Clark and Wilkes-Gibb (1986) it is clearly stated that "our concern is not with semantic reference, but with the speaker's referent", (p. 2) and they provide a distinction between literary model and conversational model of linguistic exchange, claiming that the collaborative process for determining reference is active only in a conversational situation. On the one hand we agree on the peculiarity of tools used for this process during a "face-to-face" conversation (for example, the use of particular intonation of utterances), but on the other we think that the main features of this process can be generalized to give a more extensive account of words' meaning.

3. Our proposal

Starting from these positions, we want to take a further step: we not only want to show that in order to converge on the use of a word a negotiation is needed, we also want to claim that repeated negotiations *shape the meaning* of such word; they shape the 'private' meaning, attributed to the word by each individual agent, but negotiations also shape what has been called the 'literal' meaning, the meaning a word is supposed to *have* independently from context. In the two next subsections we'll show how these two different notions of meaning are characterized in our approach.

Private (or speaker's) meaning

We can say that the speaker's private meaning of a linguistic expression is a mental representation consisting of a variable set of conceptual features that don't represent a specific description that has to be satisfied by a candidate referent for the word, but rather must be considered just as tools to use when we engage in any linguistic exchange.

The linguistic exchange has a twofold purpose: the speakers try to find an agreement on the intended reference for a given word or expression, meshing their individual perspectives, and at the same time they refine their internal representation of the meaning in order to be more successful in future exchanges. The set of conceptual features is variable in two ways: it varies in time for the same speaker, due to the exposure to multiple linguistic interactions, hence to multiple refinement processes; it varies amongst different speakers, because there are hardly two speakers that have been through exactly the same series of refinement processes.

The private meaning is continually reshaped by the negotiation process, that creates new connections between words or concepts, strengthens some of these connections and weakens others etc.

As a possibility to represent private meaning as something shaped by interactions we could assume a flexible semantic network, where words (or concepts) are connected one another when they frequently present themselves together in the experience of the agent.

The basic idea of a semantic network, as conceived in (Quillian 1966), is to represent a semantic field as a graph where the nodes represent words (or concepts, or features), and the links between the nodes represent relationships. The links, or connections, between the nodes can have different nature (similarity, inclusion, converse...), and different strength (more or less similar, for example). The nodes of such a network are activated when the corresponding concept is in use, and this activation spreads to the connected nodes, according to the distance, or the strength, or the nature of the connection. A psychological theory of spreading activation as a model for semantic processing was originally presented in (Collins and Loftus 1975) and has been applied with success to explain psychological phenomena like semantic priming. Implemented examples of semantic networks include WordNet (Miller 1990) or

Semantica, an expansion of SemNet (Fisher 2000). What we are interested in is a flexible kind of semantic network, where nodes and connections can be reshaped as a consequence of use (cf. Slipnet in Mitchell 1993).

Literal (or Public) meaning

What has been called literal meaning, or also public meaning, of a word in a given language is also a set of conceptual features, but it is an abstract set and not the specific mental content of a person. It grows as a generalization from the most common conceptual features representing the speakers' meaning in widespread successful linguistic interactions¹. We are then speaking of an abstraction, a "mean" value extracted from the speakers' usage of that meaning, and being the mean amongst values that vary in time, it also varies in time, even if slower. How is this mean value calculated? In everyday life it is not really calculated, but just estimated according to the best knowledge we have of a language. Compiling dictionaries is a professional performance of this estimation, which tries hard to extrapolate from as wide a basin of language usage as possible.

To sum up, our proposal consists of a treatment of meaning as emerging from processes of communication and negotiation taking place among agents: these processes can reshape the representation of private meaning, and in the long run the abstract representation of the literal meaning. As (Rapaport 2003) puts it:

"We almost *always* fail [...]. Yet we almost always nearly succeed: This is the paradox of communication." (Rapaport 2003, p. 402)

In order to understand how this process of negotiation determines a continuous transformation of the meaning of a word, in what follows we are going to describe a simplified example of negotiation where we can see the changes in the semantic representation.

A simple example of negotiation

Let's take the proper name "Socrates", pointing out the process that can transform someone's representation of such a name. If we are not learning a new word, the process of negotiation starts with a mental representation of the word/expression that

¹ Language (or semantic) games, as described in (Hintikka 1976) are a very interesting logical tool that has been proposed for the formalization of these processes.

has been consolidated through all the previous negotiations, that is what we have called private or speaker's meaning. The most common referent for "Socrates" is the well-known Greek philosopher, so we can imagine that the word "Socrates", inside a semantic network, is connected with words like "proper name, male, person, philosopher, past times, Greece, well-known, maieutic, ..." (for the sake of simplicity, we describe here only simple word-to-word connections). The level of activation in these connections will be higher or lower in relation with the actual context, i.e. the general state of activity in the whole semantic network². The number and the strength of the connections, instead, can be different in relation to my personal history of acquisition of the word; it is the product of a chain of previous linguistic exchanges that I have engaged in while learning and using this name, and maybe this chain could be followed backwards to trace the original source of the name, in this case the philosopher called Socrates³.

The history of learning and usage of the name "Socrates" is even the source of the assumption that the person we are speaking with shares the same information about this name. There are words that we consider shared between anyone who speaks the same language, others that are shared between people of the same region, of the same social group, of the same grade of education, of the same technical background. All this information is related to what we know about the person we talk with, and what we know about the acquisition of a certain word.

So let's come to the conversation between two persons that we will call Massimo and Viola; this is taking place in Massimo's living room. Entering the living room, Viola notices the phonebook half-destroyed on the floor, and asks:

Viola: "My God, who has made this mess?"

Massimo: "Must have been Socrates"

Viola: "Socrates?"

Massimo: "Yes, he has even left a bone in here"

Viola: "Is the name of your dog Socrates?"

² For example, If I'm talking about gardening, the connections leading to "Socrates" in my network are supposedly not touched by the spreading of activation (unless at some point the conversation touches the word "hemlock").

³ In this way, this theory can account for the crucial insight provided by the causal chain theory of reference. But there is a main difference with some of the accounts of the causal chain theory, as the one given by Kripke for example: according to Kripke, the connection between word and reference is transmitted through the causal chain as an constant connection, being a "rigid designator"; in the account of meaning as a dynamic result of negotiation, there is not space for such rigidity. Unsuccessful exchanges that didn't have a chance to be corrected and had wide resonance in a community can lead to permanent deformations of the connection established *ab initio*.

Massimo: "Yes"

Following freely the terminology used by Clark and Wilkes-Gibbs (1986) on the psychological side and Heeman and Hirst (1995) on the computer science side, we can describe this conversation as follows: Viola and Massimo have engaged in a *referring plan*, Viola formulating a question that requires a referent as an answer (whoever has made the mess), Massimo responding with a *presentation* of a referring expression, identifying the author of the mess with the referent of the name "Socrates". Following the general schema given in (Clark and Wilkes-Gibbs 1986), after the *presentation* we need an *acceptance* (explicit or implicit) or a *rejection* from Viola. Asking "Socrates?", Viola is trying to manifest a rejection of the name as something able to pick up a referent. She is following a pragmatic principle of charity: there is no way to make sense of the fact that her referent for Socrates, the dead philosopher, is the one who made a mess in Massimo's living room. An internal revision process is in action: first of all, the association with past times has to be dismissed, because we are dealing with the agent of a recently happened event. Giving up on that means giving up several other connected features, or at least they become not likely (as philosopher, Greece, maieutic...). Let's imagine she is left with these connections: "proper name, male, living being, person". This is still not enough, so Viola asks for an *expansion* of the referring expression, trying to reach the goal of what we called *referring plan*, i.e. to find a common referent.

Now it's Massimo's turn, and he replies "Yes, he has even left a bone in here": Massimo has misunderstood Viola's rejection, in fact instead of thinking that Viola didn't get what the referent of "Socrates" is, he thinks Viola doubts that the actual referent of Socrates (in Massimo's use of the name) is also the author of the mess. Consequently, Massimo presents an *expansion*, to justify why he thinks that the reference of "Socrates" and the reference of "who has made the mess" are the same: "He has even left a bone in here". In Viola's representation system, the idea of an agent that leaves a bone on a floor in a house is connected with a pet, particularly a dog. Now the activation of "dog" helps to retrieve a background information, in fact she knows that Massimo has a puppy. The feature "person" can be deleted from her connections and she builds up a new hypothesis: "proper name, male, living being, dog, Massimo's pet". Viola, once again, has to *repair the referring plan*, as a consequence of her guessing, so she asks: "Is the name of your dog Socrates?". Finally, Massimo accepts this final correction, the negotiation is complete and the referential process is successful. End of negotiation.

The referent of the word "Socrates" has been established, and so now she has extended her application of the word to a new usage and a new referent. This doesn't mean that this new referent has been fixed by the circumstance: any new linguistic act can require a new negotiation.

How do these changes in the speakers' meaning affect the general linguistic meaning of a word? In no relevant way. Actually, if Massimo was not a friend with whom Viola has frequent interactions, but just an occasional acquaintance, maybe the changes affecting her representation of Socrates in this conversation would be destined to weaken and dissolve in time. But saying this we don't mean to subscribe the view that this negotiation process is relevant only for the speaker's meaning. What we are arguing here is that this is also the starting point to define the more general linguistic meaning, that is nothing more than a large abstraction from the single speakers' meaning. We can imagine that these changes in the representation of the name "Socrates" can become relevant on a large scale if, for example, Massimo's dog becomes a movie star like Lassie. We are going to be more explicit about this in the next section, where we are going to sum up what has been said so far.

4. An alternative definition of meaning: conclusion and future issues

Let's summarize the view that we have delineated in the preceding sections. Using the traditional distinction between speaker's meaning and literal meaning, we can say that the private (speaker's) meaning of a linguistic expression is a mental representation consisting of a variable set of conceptual features, compositionally related to the syntactic structure of the expression. This set varies with every exchange in which such word or concept is involved.

What has been called literal meaning, or also public meaning, of an expression in a given language is also a set of conceptual features, but it is not something that can be clearly separated from the speakers' meaning: it emerges as an abstraction from the private meanings attributed by the speakers of a certain community. In common linguistic interactions, the literal meaning is the minimum set of features that we can assume to be shared by an unknown person who speaks our language: even in this case we just perform an heuristic estimation of what can safely be assumed. Any kind of

knowledge of the background of our interlocutor can lead us to extend the set of assumption to be made⁴.

How can such a process be formalized? We need to find an acceptable definition of the representation of the meaning of a word at the beginning of the exchange, for every agent engaged in the interaction. Then we follow the transformations of these representations while the dialogue is adding common information, and we can examine, at the end of the exchange, if the agents have reached a convergent structure, and what kind of structure it is: this will be the referent in that context. The kind of structure that has been the most successful on the total of linguistic interactions will be perceived as the linguistic meaning. The traditional relationship between private meaning and literal meaning is here inverted: we don't obtain the private meaning by applying pragmatic rules to the literal meaning, but rather we obtain a literal meaning when the negotiations about a word amongst speakers reach a large scale agreement⁵.

Now, let's go back to the main question: can this kind of procedure be formalized and added to the grammar defining a language? In our opinion, this could be done by building a flexible semantic network that can represent ideally the main connections between concepts in an average speaker, and then describing the main processes that dynamically act on the network during the negotiation procedure with other speakers.

A concrete possibility to describe such processes could be the dynamic semantics proposed by Groenendijk and Stokhof (1991). In the account given by Veltman (1996), the author stresses the dynamic component of this new kind of semantic stating:

the slogan "You know the meaning of a sentence if you know the conditions under which it is true" is replaced by this one: "You know the meaning of a sentence if you know the change it brings about in the information state of anyone who accepts the news conveyed by it" (Veltman 1996, p. 221).

In computer science several attempts have been already made to describe and reproduce the way in which these processes work; we mention here only the more "semantically biased" approaches: the one based on ontologies (cf. Masolo 2002), and on algorithms of semantic matching (cf. Bouquet 2004 and Giunchiglia 2004).

⁴ Maybe we could think of this process of formulating assumptions before a conversation as the result of a preliminary negotiation process with the context of conversation. In this case the process would take place between a speaker and a given situation, and not among speakers.

⁵ A deeper development of this view can lead us to rethink the role of literal meaning. Outside an artificial and abstract description of language, is the idea of a literal meaning playing any role at all? A development of this view can be found in Sperber and Wilson (1998).

The direction shown by these studies is the one we believe is worth pursuing also in philosophy.

References

- Bouquet, P., Serafini, L. and Zanobini, S., 2004, Coordinating Semantic Peers *Proceedings of 11th AIMS Conference (AIMSA 2004)*, September 2-4, Varna, Bulgaria.
- Clark, H.H. and Wilkes-Gibbs, D., 1986, Referring as a collaborative process, *Cognition* 22: 1-39.
- Clark, H.H., 1996, *Using Language*, Cambridge University Press, New York.
- Collins, A. M. and Loftus, E. F., 1975, A spreading-activation theory of semantic processing, *Psychological Review* 82(6): 407-428.
- Fisher, K. M., 2000, SemNet® Semantic Networking, in: *Mapping Biology Knowledge*, K. M. Fisher, J. H. Wandersee and D. Moody, Kluwer, Dordrecht, Netherlands, pp.143-165.
- Føllesdal, D., 1995, In what sense is language public? In: *On Quine: New Essays*, P. Leonardi, ed., Cambridge University Press, New York & Cambridge.
- Giunchiglia, F., Shvaiko, P. and Yatskevich, M., 2004, Semantic matching, in *Proceedings of the 1st European semantic web symposium (ESWS'04)*", C. Bussler, J. Davies, D. Fensel and R. Stuer, Heraklion, 10-12 May 2004, pp. 61-75.
- Grice, H. P., 1969, Utterer's meaning and intentions, *The Philosophical Review*, 78: 147-177.
- Groenendijk J. and Stokhof M., 1991, Dynamic predicate logic, *Linguistics and Philosophy*, 14: 39-100.
- Heeman, P. and Hirst, G., 1995, Collaborating on referring expressions, *Computational Linguistics* 21 (3): 351-382.
- Hintikka, J., 1976, Language-games, *Acta Philosophica Fennica* XXVIII(1-3): 105-125.
- Kripke, S., 1972, *Naming and Necessity*, Harvard University Press, Cambridge MA.
- Masolo, C., Borgo, S., Gangemi A., Guarino N. and Oltramari, A., (2003), *WonderWeb Deliverable D18. Ontology Library (final)*. Trento, Laboratory for Applied Ontology-ISTC-CNR.
- Miller, G. A., 1990, WordNet: An on-line lexical database, *Inter. Journal of Lexicography* 3(4): 235-312.

- Mitchell, M., 1993, *Analogy-Making as Perception*, Bradford Books/MIT Press, Cambridge.
- Peregrin, J., 1995, Review of "Making it Explicit" by Robert Brandom, *Canadian Philosophical Review* XV: 84-86.
- Peregrin, J. (ed), 2003, *Meaning: The Dynamic Turn*. Elsevier, Amsterdam.
- Quillian, M.R., 1966, *Semantic memory*, Unpublished doctoral dissertation, Carnegie Institute of Technology, reprinted in: *Semantic information processing*, M. Minsky, ed., MIT Press, Cambridge MA, 1969.
- Rapaport, W. J., 2003, What did you mean by that? Misunderstanding, negotiation, and syntactic semantics, *Minds and Machines* 13(3): 397-427.
- Veltman F., 1996, Defaults in update semantics, *Journal of Philosophical Logic*, 25: 221-261.
- Wittgenstein, L., 1958, *Philosophical Investigations*, The Macmillan Company, New York.