In Brisard, F. (ed.) *Seduction, Community, and Speech.*
A Festschrift for Herman Parret.
Amsterdam, John Benjamins, 2005

Language as pragmatics:
studying meaning with simulated language games

Domenico Parisi
Institute of Cognitive Sciences and Technologies
National Research Council, Rome
parisi@ip.rm.cnr.it

1. Meanings are a "terrible" problem for science

Unlike philosophers, who rely only on conceptual analysis, reasoning, and discussion with other philosophers, both living and dead, scientists must ultimately judge the goodness of their theories and hypotheses by comparing these theories and hypotheses with systematically and objectively observed empirical phenomena. Hence, the possibility of observing, systematically and objectively, reality, not just talking or thinking about reality, is a critical requirement for doing science. Language is both linguistic forms and the meaning of these forms. Linguistic forms by themselves are just sounds that are heard or are produced in speaking or visual images that are read or produced in writing. To become linguistic signs these forms must have a meaning for both speakers/writers and hearers/readers. It is their meaning that allows linguistic forms to play such an important role in the mental and social life of human beings. But meanings are very difficult to observe. Linguistic forms are reasonably well defined entities, easy to perceive, locate, identify, compare, and measure. The meanings of linguistic forms are nothing like that. As argued by Wittgenstein (1953), the meaning of a linguistic form is the role the form plays in the behavior and social interactions, or, more generally, in the life of the individuals who use it. But human behavior is difficult to observe sistematically and to describe and measure and, consequently, the meanings of linguistic forms are difficult, if not impossible, to perceive, locate, identify, compare, and measure. It is not even clear where to begin or where look in order to observe the meaning of a linguistic form. In fact, as objects of scientific inquiry meanings are almost hopeless. A science of language must necessarily be a science of both forms and meanings. But while linguistic forms can be systematically and objectively observed, their meanings cannot. Hence, meaning represents a “terrible” problem for a science of language.

Since the meaning of linguistic forms are difficult to observe and identify, scientists of language make recourse to clearly inadequate substitutes. Meanings are introduced into discourse by simply producing the linguistic form which designates them: “the meaning of ‘dog’”. Nothing is supplied except the form. Or what is supplied is a definition of the form, i.e., a set of other forms that collectively are assumed to have the same meaning as the defined form: “a dog is a mammal with a tail”. But one never sees the meaning, neither of the defined form nor of the defining forms. If one is studying a foreign language, for example in linguistics, the meaning of a form belonging to the foreign language is identified simply by offering the form that translates the foreign form in the native language. The meaning of ‘hund’ in German is ‘dog’. Again, one sees only forms, never meanings. Historical linguistics dealing with dead languages is an even more desperate case since only linguistic forms survive. Another strategy is ostensive definitions: one points out to some particular object in reality to which the linguistic form is used to refer. One identifies the meaning of ‘dog’ by pointing to a dog or to a picture of a dog. Aside from the fact that this strategy can be used only for some linguistic forms but not for all, ostensive definitions have well known intrinsic limitations. When I point to a dog to explain the meaning of ‘dog’, how does the hearer know that I am pointing to the entire animal, and not, for example, to the color of its skin?

It is clear that all these methods function only because they implicitly assume that who is listening to a lecture or reading a scientific article reporting a piece of research on language already knows the meaning of the linguistic forms that are studied in the research. But in this way the researcher is renouncing to study one half of language in that he or she does not provide or report any systematic observation on the <u>meaning</u> of the linguistic forms. Linguists only deal with linguistic forms and they do not have available to them any methodology for observing and describing meanings. Psycholinguists try to remedy to this situation and to capture the meaning of linguistic signs for young children by systematically observing and audio- or video-recording the behavior of children while they are using language. However, clearly most speakers/hearers are <u>not</u> young children. Psycholinguists also do experiments on the linguistic behavior of adults. But the context of use of linguistic forms in laboratory tasks is always very artificial. Participants may be asked to decide whether a succession of sounds or letters is a word or not a word, to choose the picture to which a given word applies among a set of different pictures, or to say whether a sentence is true or false with respect to a given picture. Clearly, in these tasks the meaning of the linguistic forms is presupposed, not observed.

The fact that the science of language does not possess an empirical methodology for studying the meaning of linguistic forms, not only makes the science of language a half-science - a science of linguistics forms but not of their meanings - but is at the origin of a series of distortions and mistaken views about the nature of language and of linguistic meanings. By practically identifying the meaning of linguistic forms with these forms, meanings appear as entities that are well-defined, clearly identifiable, fixed, invariant from individual to individual, with a well-defined internal structure, whereas meanings as components or aspects of behavior are entities with ill-defined boundaries, difficult to identify, always changing, different from individual to individual, very sensitive to the context of use, with an internal structure which is complex and different from case to case. In this way some of the most important properties of language elude the science that should capture and explain them.

2. How to solve the problem

Can the “terrible” problem that meanings represent for a science of language be solved? In this chapter we argue that it can be solved by using a methodology which uses Wittgenstein’s notion of language games and realizes language games as computer simulations. As already noted, for Wittgenstein the meaning of a linguistic form is the role the linguistic form plays in the behavioral interactions of the user of the form with the world, including his or her conspecifics. The linguistic form is part of these interactions, is part of the individual’s life. To study language as “a form of life” Wittgenstein introduces the notion of a language game, a simplified situation in which a user uses a simple language in his or her interactions with a simplified world. By creating and analyzing these simplified scenarios one can observe and study how language is actually used and, therefore, determine what is the actual meaning of linguistic forms. Wittgenstein’s justification for introducing language games was philosophical. He thought that, by observing how linguistic forms are actually used, one can dissolve philosophical problems which, for him, were pseudo-problems that arise from misusing language. But we believe that that the idea of language games can be adapted for the purposes of science, that is, for studying language as an empirical phenomena to be inserted in the unified picture of reality that science is progressively building up.

The main adaptation which is required to use languages games for the purposes of science is to realize language games as computer simulations. What are computer simulations? Computer simulations are theories or models expressed in a novel medium. Traditionally, scientific theories and models are expressed by using either ordinary language or mathematical symbols. Simulations are theories or models expressed as computer programs. When it runs in the computer, the program generates results which are the empirical predictions that are derived from the theory or model incorporated in the simulation. If the results match the observed phenomena, the theory or model is confirmed. Otherwise, it has to be modified or abandoned.

Computer simulations are a new research methodology which is being increasingly adopted in many scientific disciplines because it offers many important advantages.

The first advantage is that simulations force the researcher to formulate his or her theories or models, that is, his or her ideas and hypotheses on what is behind the observed phenomena and explain them, in ways that must necessarily be precise, explicit, detailed, and complete. The reason is that theories or models which are not sufficiently precise, explicit, detailed, and complete cannot be incorporated in a program which actually runs in a computer and produces the expected results.

Another advantage is that by manipulating the various conditions and variables of a simulation the researcher can obtain a large number of detailed results which, as we have already said, are the empirical predictions derived from the theory incorporated in the simulation which can be then compared with the actual phenomena. In this way, not only the dialogue between theories and empirical data, which is critical for science, can be enlarged and strengthened, but the researcher and, especially, his or her colleagues can have a clear idea of the actual empirical content of any given proposed interpretation and explanation.

A third advantage of simulations is that they can reproduce complex situations in which there are many kinds of causes and factors and the observed phenomena are the results of the interactions among many different entities which cannot be predicted or deduced from even a perfect knowledge of the entities and of the rules that govern their interactions. The researcher puts into the computer a large number of entities together with the rules of their interactions, and he or she observes what comes out from these interactions, including changes in the interaction rules themselves. Furthermore, computers have greater memory and processing capacities than those of human researchers. They can store large knowledge bases that belong to various different disciplines and they can run interdisciplinary theories and models that invoke all these different knowledge bases. In this way simulations function as a common language, a ‘lingua franca’, among different disciplines that tend to use different theoretical frameworks and different methods.

All these advantages of simulations are clearly important for studying such a complex phenomenon as human language but they are critical for addressing the problem of how to study the meaning of linguistic forms (Knight, Studdert-Kennedy, and Hurford, J., 2000; Parisi, 1997; Parisi and Cangelosi, 2002). If the meaning of linguistic forms is the role that these forms play in the behavior and interactions of human beings with each other and with the environment, one can simulate a certain number of individuals who live in an environment which contains various types of objects, including other individuals of the same species (conspecifics). Each individual interacts with the different objects and with his or her conspecifics. To interact means that the environment causes inputs in the sensory receptors of an individual and the individual responds by generating movements of his or her body that modify the external environment. Linguistic forms are sensory inputs for hearers produced by the phono-articulatory movements of speakers.

Each individual’s behavior is controlled by a neural network which simulates the individual’s nervous system. A neural network is a set of units (neurons) linked by unidirectional connections (synapses) (Rumelhart and McClelland, 1986). Connections can be excitatory or inhibitory and each connection has a ‘weight’ which simulates the efficacy of the synapse (number of synaptic sites). A neural network includes three layers of units: one layer of input units encodes states and events in the external or internal (bodily) environment; the input units are connected to one layer of internal units (interneurons) that encode transformations of the input encodings; the internal units are connected to one layer of output units that produce either physical movements of the individual’s body and, as a result of these movements, changes in the external environment, or changes inside the body. The neural networks of human beings include many recurrent circuits that produce self-generated inputs which constitute “mental life”, that is, mental images, rememberings, thoughts, etc.

The regularities that are present in the environment and in the interactions of the individual with the environment are captured and represented in the particular synaptic weights of the connections linking the network’s units. It is these weights that cause an individual’s neural network to respond to each input with the appropriate output, that is, with an output which tends to increase the individual’s survival and reproductive chances and his or her happiness. Synaptic weights are generally learned, i.e., they change during the individual's interactions with the environment in such a way that the individual tends to respond to the input with more appropriate output.

Linguistic forms are sounds produced by the phono-articulatory movements of an individual (the speaker) that arrive to and are encoded in the acoustic receptors of another individual (the hearer). What gives meaning to these sounds for an individual is that in the individual's experience (i.e., in the interactions of the individual with the environment) specific sounds tends to co-vary with specific aspects of the individual's experience. Specific sounds co-vary with specific objects seen, touched, or manipulated, with specific properties of these objects, with specific events or processes, with specific actions executed by the individual or seen by the individual as they are executed by another individual. These regularities and co-variations are captured and represented in the individual's synaptic weights and it is these synaptic weights that give to each sound its meaning for the individual.

In computer simulations one can study (observe, analyze, measure, manipulate) both the linguistic forms produced or heard by the individuals and the co-variations of specific forms with specific aspects of the individuals' experience in the environment. In this way, simulations make it possible to overcome the asymmetry between linguistic forms and their meaning from the point of view of the possibility of observing and measuring them which, as we have argued, is the main problem for a science of language. Both can be simulated.

3. Some concrete examples

An individual lives in an environment that contains both edible and poisonous mushrooms (Cangelosi, and Parisi, 1998). To survive and reproduce the individual must eat the edible mushrooms and avoid the poisonous ones. Since each individual mushroom is perceptually different from all other mushrooms, the individual's task is to recognize each encountered mushroom as edible or poisonous and to respond to edible mushrooms by approaching and eating them and to poisonous mushrooms by going away from them. When it encounters a mushroom, the individual can trust its personal capacity to discriminate the two categories of mushrooms or there may be a more expert conspecific who wants to help the individual and tells the individual what is the category of the mushroom. The conspecific produces one particular sound if the mushroom is edible, and another, different sound if the mushroom is poisonous. In another scenario, the individual may see a mushroom but the mushroom is too distant for the individual to be able to perceive sufficiently well the mushroom's perceptual characteristics and recognize its category. Hence, for the individual the only available strategy is to approach all distant mushrooms until it is sufficiently close to a mushroom to be able to recognize its category - which would not be a very energy-efficient strategy. In these circumstances, if a conspecific is closer to the mushroom and tells the individual what is the mushroom's category, the individual would be able to approach only edible mushrooms - with a clear advantage in terms of energy consumed for moving in the environment.

By realizing these simple scenarios as computer simulations one can study in which conditions a simple language like this can emerge given various environmental constraints, how the language can arrive to possess such desired characteristics such as social sharedness and limited homonymy and synonymy, what induces an individual to help another individual by sending informative signals to the other individual, how the neural network of both speakers and hearers is internally organized compared to the neural network of individuals without language.

Are the two signals that are exchanged in the preceding very simple scenario verbs or nouns? Should they be translated using two English nouns (or, more precisely, noun phrases), "edible mushroom" and "poisonous mushroom", or two English verbs, "approach" and "avoid"? The scenario appears to be too simple for deciding one way or the other. In the experience of both the speaker and the hearer the two sounds co-vary with both the type of mushroom and the type of action at the same time. Hence, the two sounds cannot be classified as nouns or as verbs.

But consider a somewhat more complex scenario (Parisi, Cangelosi, and Falcetta, 2002). An individual is endowed with a retina on which one or more different objects can appear at the same time, and with an arm with which the individual can reach and manipulate the objects. When there is a single object in the individual's retina, another individual produces one sound to which the first individual responds by pushing the object away from itself and another sound to which the first individual responds by pulling the object towards itself. If two differently shaped objects are present at the same time in the individual's retina, the conspecific produces one sound in response to which the first individual executes some particular action on a specific one of the two objects and another sound in response to which the same action is executed on the other object. Since in the first case the two sounds co-vary with the two different actions, the two sounds can be interpreted as proto-verbs; in the second case the two sounds co-vary with the objects on which an action is executed and they can be interpreted as proto-nouns. If both a proto-verb and a proto-noun are produced in a sequence by the conspecific so that the hearer knows both which action to execute and on which object to execute the action, we can talk of a proto-sentence. In this case we can study how proto-nouns and proto-verbs emerge in the simulated scenario, which class of words emerges first, what happens, and where, in the neural network of both speakers and hearers when they produce or hear proto-nouns and proto-verbs. For example, it turns out that in the neural network of our artificial organisms, which includes successive layers of internal units, proto-nouns are represented in internal units that are closer to the sensory input units whereas proto-verbs are represented in internal units closer to the motor output units.

The simulated scenarios can be appropriately modified and multiplied to capture other classes of words and meanings. For example, if in the individual’s retina there are many different objects that can vary in both their shape and color, to allow the hearer to indentify the object on which he or she must execute some particular action the speaker must produce not just a proto-noun but both a proto-noun and a proto-adjective, i.e., a proto-noun phrase. (In actual languages nouns tend to co-vary with the shape of objects, adjectives with their other properties. Cf. Landau, Smith, and Jones, 1988; 1992.) Or an object can be identified by its spatial relation to another object. In these circumstances, the speaker must provide the hearer with both a sound that co-varies with the other object and a sound that co-varies with the particular spatial relation of the object to be identified with the other object, e.g., “under”.

In all the simulations that we have described speakers and hearers use linguistic forms that co-vary with aspects of public experience. An object or an action is experienced by both the speaker and the hearer. But language can also be used to talk about private experiences. Mothers can teach their children to produce some specific linguistic form when they have some external, i.e., public, evidence that their children are having some specific private experience, such as pain. In this way individuals can learn to use language to communicate about their mental life with others.

In our simulations individuals have a physical body and they live in a physical environment but these physical aspects are very simplified and are simulated in a computer. In other researches the physical aspects are still simulated but they can be more realistic (Marocco, Cangelosi, and Nolfi, 2003) or, even, the emergence and self-organization of language can be studied in artificial organisms which have a real physical body and live in a real physical environment, that is, in robots. In these conditions one can study the emergence of a language in a collection of robots or between robots and humans (Steels, 2002).

4. Conclusion

Computer simulations of how linguistic forms are used by a population of individuals interacting with the environment and among themselves can be a useful tool to study the meaning of linguistics forms. Human behavior remains a difficult phenomena to observe, analyze, and measure and therefore the meaning of linguistic forms, being part of such behavior, remains a difficult topic for science. However, computer simulations are an important help for formulating theories and models of meaning in explicit, detailed, complete, and consistent ways, for publicly demonstrating the detailed empirical predictions derived from such theories and models, and for generating new empirical predictions to be compared with the actual observed phenomena.

Of course using computer simulations to study how language is used has its costs. Simulations may seem to be appropriate for studying simplified languages in simplified environments but language is very complex and it can be used to talk about very complex objects and events. Our simulations resemble Wittgenstein’s language games, which are simplified models of the complex roles that language plays in our everyday life and are similar to the games “by means of which children learn their native language” and to a “primitive language” (Wittgenstein, 1953, 5).

However, there are two arguments in favor of the method proposed for studying the meaning of linguistic forms. Simulations are theories and as theories they simplify like all theories in science. Scientific theories and models allow us to go beyond descriptions of the variety of observed phenomena and to discover what lies behind these phenomena and explain them just because they simplify and capture what is essential in the observed phenomena. Therefore, that simulations simplify is not in itself an objection. Of course, like all theories, simulations must make the correct simplifications. What can be objected in specific cases is that some particular simulation makes the wrong simplifications, that is, that it leaves out aspects of the actual phenomena which are critical for explaining and understanding the phenomena of interest. But simulations, like all theories, must simplify.

The second argument in favor of our linguistic simulations is that although they resemble Wittgenstein’s language games they are different from Wittgenstein’s language games in one important respect. The simulations that we have described are not just language games; they are physically realized language games. They actually run in a computer or control physical robots. As we have observed, computers have memory and processing resources that go well beyond those possessed by humans in their heads. Given the great memory and processing capacities of computers we can expect that in the future the kinds of linguistic simulations that we have described will simplify less and will make it possibile to study more complex languages in more complex environments.

The approach to the study of language which underlies our simulations is based on the following "genetic epistemology" assumption: if you want to understand X, try to reconstruct how X has become what it is now. Hence, we are forced to start from simple languages in simple organisms in simple environments. Then, we can move on and study how language changes and possibly becomes more complex in evolution, ontogeny, and language change (Parisi and Cangelosi, 2002).

On a more general plan, our simulations lead to a view of language as entirely constituted by pragmatics. Traditionally, one distinguishes three aspects of language: semantics, which is concerned with the meaning of linguistic forms, syntax, which is concerned with the rules for creating linguistic forms made up of more elementary forms, and pragmatics, which studies how language is actually used in context and the purposes of speakers. Wittgenstein’s idea of language games and our implementation of this idea in computer simulations imply a view of language as entirely constituted by pragmatics. Everything in language is language use. Language is the use of linguistic forms in the behavior and social interactions of a particular species of organisms. Semantics and syntax are only abstractions or idealizations from language use. This does not mean that semantics and syntax are nonexistent or unreal. But, as abstractions from language use, semantics and syntax cannot be really understood unless we first understand the uses of language from which they are abstracted.

References

Cangelosi, A., and Parisi, D. The emergence of a ‘language’ in an evolving population of neural networks. Connection Science, 1998, 10, 83-97.

Knight, C., Studdert-Kennedy, M., and Hurford, J. (eds.) The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form. Cambridge, Cambridge University Press, 2000.

Landau, B., Smith, L.B., and Jones, S.S. The importance of shape in early lexical learning. Cognitive Development, 1988, 2, 291-321.

Landau, B., Smith, L.B., and Jones, S.S. Syntactic context and the shape bias in children's and adults' lexical learning". Journal of Memory and Language, 1992, 31, 807-825.

Marocco D., Cangelosi A., Nolfi S. (2003), The emergence of communication is evolutionary robots. Philosophical Transactions of the Royal Society London – A, 2003, 361, 2397-2421.

Parisi, D. An Artificial Life approach to language. Brain and Language, 1997, 59, 121-146.

Parisi, D. and Cangelosi, A. A unified simulation scenario for language development, evolution, and historical change. In Cangelosi, A. and Parisi, D. (eds.) Simulating the Evolution of Language. New York, Springer, 2002.

Parisi, D., Cangelosi, A., and Falcetta, I. Nouns, verbs and language games. Italian Journal of Linguistics, 2002, 14, 99.114.

Rumelhart, D.E. and McClelland, J.L. Parallel Distributed Processing. Explorations in the Microstructure of Cognition. Cambridge, Mass., MIT Press, 1986.

Steels, L. Grounding symbols through evolutionary language games. In Cangelosi, A. and Parisi, D. (eds.) Simulating the evolution of language. New York, Springer, 2002.

Wittgenstein, L. Philosophical Investigations. Oxford, Blackwell, 1953.