Perspectives on meaning and interpretation

Chris Potts, Ling 130a/230a: Introduction to semantics and pragmatics, Winter 2025

1 Overview

The interpretation function is ⟦ ⟧. It connects language to meaning. This raises two fundamental questions:

i. What are meanings?

ii. What is the nature of this connection?

This short reading offers a variety of perspectives on these questions.

2 Partee (1995): Meanings are (partly) aspects of objective reality

In one of our assigned readings, Partee is circumspect concerning the nature of meaning:

> What is meaning?
>
> It is not easy to tackle a question like that head-on; and while it is an important question to keep wrestling with, a total answer is not required in advance of doing fruitful work on semantics, any more than biologists wait for the answer to the still-difficult question "what is life?" before getting down to work. A scientific community just needs *some* clear examples to get started, and then empirical and theoretical advances proceed together, along with further sharpening of key concepts.

This is reasonable, and we will to some extent adopt this perspective ourselves, in that we will develop theories that are agnostic about the precise nature of linguistic meaning. However, it seems unsatisfying to leave this fundamental question completely open.

In an effort to remain agnostic about the nature of meaning, Partee follows "Lewis's advice":

> Not surprisingly, it is philosophers who have provided two particularly useful strategies for thinking productively about the question of what meanings are. The first comes from David Lewis (1970).
>
> *Lewis's Advice*: "In order to say what a meaning is, we may first ask what a meaning does, and then find something that does that." (p. 22)

In practice, this means adopting mathematical constructs as "things that meanings do": sets, functions, and so forth. Our perspective in this class will be that meanings are akin to little computer programs: functions that can interact with each other to create new meanings, and that knowing a language entails knowing how to use these functions. This seems like it's compatible with a lot of different goals one might have for a theory of meaning. For example, it could inform natural language processing systems, computational theories of cognition, and/or purely theoretical work on language.

However, the above passage continues in a way that seems to commit Partee to really substantive claims about the nature of meaning:

> So let's think about what meanings do besides combine in some way to make more meanings. For this, Max Cresswell (1982) has shown how a great deal of mileage can be gotten from a very minimal assumption. Cresswell notes that we have no good a priori conception of what meanings are, but we do know at least one thing about them, which he dubs his "Most Certain Principle."
>
> *Cresswell's "Most Certain Principle"*: "For two sentences α and β, if [in some possible situation—BHP] α is true and β is false, α and β must have different meanings." (p. 69)
>
> If we follow these two strategic pieces of advice, they lead rather inevitably to the idea that truth-conditions are at least one fundamental part of what should go into the notion of the "meaning" of a sentence (not necessarily all, by any means). And while truth-conditions may at first look much too austere to make up a very large part of what meanings should be, it turns out to be surprisingly nontrivial to assign meanings to the lexical items and principles for combining meanings of syntactically structured parts so as to eventually arrive at relatively correct truth-conditions for sentences.

Cresswell's "Most Certain Principle" says that truth conditions are important aspects of meaning. More specifically, it says that a difference in truth conditions entails a difference in meaning. What, then, are the "situations" referred to here, and what is "truth"? It sure seems like these are statements about an external, objective reality. In other words, the answer to the question "What are meanings?" is that they are aspects of an objective reality.

One brief digression: importantly, Cresswell's Principle says "if different truth conditions, then different meanings". It does *not* say "if different meanings, then different truth conditions". In other words, it allows for meaning contrasts that are not reflected in truth conditions. So there could be more to meaning than the objective reality that defines truth. But objective reality is an aspect of meaning on the Partee/Cresswell view.

Our second major question is: What is the nature of the connection between language and meaning that our interpretation function $[\![\ ]\!]$ establishes? It seems to me that Partee does not resolve this question. Rather, she focuses on arguing for particular connections between language and objective reality. We will look in detail at the connections she makes for adjectives.

3 Extensions and intensions

Fashioning a theory of meaning in terms of objective reality immediately runs into a problem that Partee confronts later in her article. Consider, for example, the fact that, in our world, an animal has a heart if and only if it has at least one kidney. This means that *animal with a heart* and *animal with a kidney* pick out the same pieces of our objective reality. However, intuitively, these two phrases have different meanings. To capture this, semanticists distinguish between the *extension* of a phrase and its *intension*. Partee explains:

The idea, which traces back to Frege (1892) and was further developed through the work of such philosophers as Carnap (1956), Hintikka (1969), Kripke (1963), and Montague (1970), is that every noun expresses a *property*, which we will call its *intension*; that property, together with the facts in a given state of affairs, determines what set the noun happens to denote (as its *extension*) in the given state of affairs. The *intension* comes much closer than the *extension* to what we ordinarily think of as the meaning of the noun; the intension is more like a characterization of what something would have to be like to count as instance of that noun. The nouns *unicorn* and *centaur* both have (presumably) the same extension in the actual world; namely, the empty set: there are not any of either. But they do not have the same meaning, and that correlates with the fact that there are fictional or mythical states of affairs where the two nouns have different extensions.

Intensions and extensions can be modeled using the notion of *possible world* (possible situation or state of affairs, possible way things might be), a notion that may be approached from various angles (see the collection in Allén 1989). Linguists working on formal semantics tend to view possible worlds as a formal tool for illuminating a certain kind of semantic structure, without necessarily taking a stand on the many deep philosophical issues that can be raised about them. It is worth noting, however, that some such notion is probably essential for an understanding of some very basic aspects of human cognition. Evidence for conceptualization of "other possible worlds" can be seen even at a prelinguistic level in any child or animal that can show surprise, since surprise signals mismatch between a perceived state of affairs and an expected state of affairs. The notion of alternative possible worlds should therefore be understood not as a matter of science fiction but as a fundamental part of the ability to think about past, future, and ways things might be or might have been.

To a first approximation, we can take the extension of the predicate *surgeon* at a time t in a possible world w to be the set of things that have the property of being a surgeon in w at t. More generally, the extension of a predicate in a given state of affairs is, by definition, the set of all those things of which the predicate is true in that state of affairs. This set is a reflection of what the predicate means; for, given the way things are, it is the meaning of the predicate that determines which things belong to the set and which do not. But the extension is also a reflection of the facts in the state of affairs or possible world; the meaning and the facts jointly determine what the extension happens to be. Two predicates may therefore differ in meaning and yet have the same extensions; but if they differ in meaning they should differ in intension. Or more accurately, if they differ in truth-conditional aspects of meaning, they should differ in intension. Frege (1892) notes that there are things like "tone" or "emotional affect" that might also be ingredients of meaning in the broadest sense that have no effect on determining extension and are therefore not reflected in intension. Two terms differing only in "tone" or "connotation" or the like might therefore have the same intension but not be considered to have quite the same meaning.

This is still a realist position that is grounded in truth conditions, but the truth conditions are now defined in terms of possible worlds.

Important: Partee *almost* commits to something much stronger that Cresswell's Principle here, in saying "if they differ in meaning they should differ in intension", but she catches herself ("if they differ

in truth-conditional aspects of meaning") and goes on to acknowledge that there may be aspects of meaning that go beyond intensions.

4 Jackendoff (1996): Meanings are mental constructs

Jackendoff is an important critic of the realist position that Partee seems to adopt by following Cresswell's Advice. He calls Partee's sort of approach "E-semantics" and constrasts it with the theory of "I-semantics", which focuses on people's knowledge of language. In the following passage, he provides a high-level overview of this perspective. As you read, note in particular what he is saying about how these two views relate to each other.

1 The Basic Stance of I-Semantics

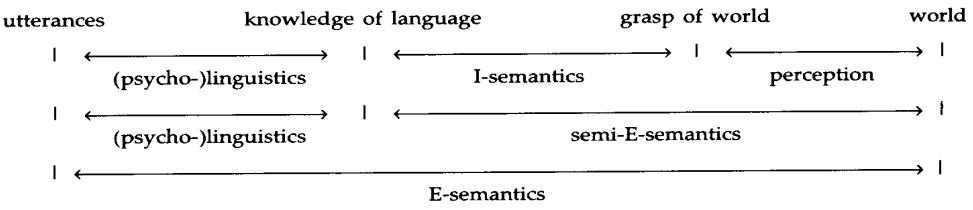
Chomsky (1986b) distinguishes between two ways to treat language as an object of inquiry. The study of "externalized language" or *E-language* treats language as an external artifact used by human beings, and seeks to characterize its properties as part of the external world with which humans interact. By contrast, the study of "internalized language" or *I-language* treats language as a body of knowledge within the minds/brains of speakers, and seeks to characterize its properties within the context of a more general theory of psychology. Since at least the time of *Aspects of the Theory of Syntax* (1965), the goal of generative grammar has explicitly been an account of I-language.

Clearly, an account of how humans actually use language in the external world eventually requires an account of their grasp of the principles of language, i.e. a theory of I-language. Thus a theory of I-language is ultimately necessary, whether or not one desires an account of E-language as well. (Chomsky has in fact argued further that E-language is not a rewarding subject of scientific inquiry at all.)

Semantics, the theory of the relation between language and the world, suffers from the same duality: one can examine it either as an abstract relation external to language users (*E-semantics*) or as a body of knowledge within the brain/mind of the language user (*I-semantics*). Again, an account of how humans use language requires an account of how they *grasp* the relationship between language and the world, so ultimately a theory of I-semantics is desirable.

More specifically, if the theories of syntax and phonology are situated within the stance of I-language, semantic theory may be concerned either with the relation of language users' knowledge of language to the external world (I'll call this "semi-E-semantics"), or with the relation of language users' knowledge of language to their *grasp* or *understanding* of the world. In the latter

Figure 20.1



case, a full theory then requires as well an account of the relation between the world and the language user's grasp of it, which falls under the standard topic of "perception and cognition." Figure 20.1 schematizes the relationship of these different approaches.

Most research in semantics either is inexplicit about its stance or else professes E-semantics. However, a growing body of work in several somewhat independent traditions has explicitly adopted the stance of I-semantics, aspiring to study the relation between human language and human concepts, and to use language as a tool for exploring the structure of human cognition. The present chapter discusses the consequences of adopting such a stance and some of the more prominent results and disputes that have emerged.

The following passage from Jackendoff (1996) describes both what I-semantic meanings are and what the semantic interpretation function does:

4 The Nature of Truth and Reference in I-Semantics – Preliminaries

In order to treat semantics as an issue about the structure of the human organism, it is necessary to be careful about basic goals of the enterprise. In particular, the traditional preoccupation with explicating the notion of the truth or falsity of a sentence must be re-evaluated. For there is no longer a direct relation between an utterance and the world that renders the utterance true or false; there is instead the sequence of three relations diagrammed in the upper line of Figure 20.1.

As a consequence, a definition of truth for I-semantics parallels the definition of grammaticality for I-linguistics. In I-linguistics, the statement (7a), which appears to be about sentences abstracted away from speakers, is always taken as an abbreviation for statement (7b), which puts grammaticality squarely in the mind of the language user.

(7) a. String S is a grammatical sentence of language L.
b. A speaker of language L judges string S grammatical (subject to limitations of memory and processing, and under an idealization of uniformity among speakers).

Similarly, the traditional Tarskian definition of truth (8a) must be reinterpreted in I-semantics as an issue of judgment, as in (8b).

(8) a. Sentence S of language L is true iff conditions $C_1, \ldots, C_n$ obtain in the world.
b. A speaker of language L judges sentence S true iff conditions $C_1, \ldots, C_n$ obtain in his or her construal of the world (subject to limitations of memory and processing, and under an idealization of uniformity among speakers).

That is, truth is no longer regarded as a relation between a sentence and the world, but rather as a relation between a sentence and a speaker's construal of the world. Parallel reinterpretations must be adopted for logical relations such as entailment, presupposition, and so forth.

This reinterpretation places a crucial burden on I-semantic theory. It is no longer possible simply to characterize "the world" logically or set-theoretically, as is frequent in formal semantics. Rather, it is an empirical problem to determine what sorts of entities inhabit the world as humans construe it. These entities may or may not be characterizable in standard logical or set-theoretic terms, and in fact they prove not to be, as will be seen below. Furthermore, truth-conditions must be stated in the vocabulary of human construal of the world.

The claim of this approach is that when people communicate linguistically, they do not communicate about the world plain and simple, but about the world as humanly understood. The entities to which speakers refer are not entities in "the world" plain and simple, but rather entities available in the human construal of the world. These include physical objects and events, illusory objects such as virtual contours, fictional objects such as Santa Claus, social constructs such as marriages and university degrees, mental constructs such as intentions and beliefs, and theoretical constructs such as numbers and logical operators. From the point of view of I-semantics, they all have equally robust status. (This position is worked out in detail by Jackendoff 1983 and Lakoff 1987.)

It is a further empirical question how these different sorts of entities are related to the "real" real world, or alternatively how "meanings in the head" come to be "meaningful." We return to these issues, which are still controversial, at the end of the chapter.

5 Lewis (1975): Meaning as social convention

In a paper called 'Languages and language', the Lewis of "Lewis's Advice" develops a particular answer to the question of what $[\![\ ]\!]$ is actually doing. The paper opens by describing the view that Partee explores:

I. THESIS

What is a language? Something which assigns meanings to certain strings of types of sounds or of marks. It could therefore be a function, a set of ordered pairs of strings and meanings. The entities in the domain of the function are certain finite sequences of types of vocal sounds, or of types of inscribable marks; if σ is in the domain of a language £, let us call σ a *sentence* of £. The entities in the range of the function are meanings; if σ is a sentence of £, let us call £(σ) the *meaning of* σ in £. What could a meaning of a sentence be? Something which, when combined with factual information about the world—or factual information about *any* possible world—yields a truth-value. It could therefore be a function from worlds to truth-values—or more simply, a set of worlds. We can say that a sentence σ is true in a language £ at a world w if and only if w belongs to the set of worlds £(σ). We can say that σ is true in £ (without mentioning a world) if and only if our actual world belongs to £(σ). We can say that σ is *analytic* in £ if and only if every possible world belongs to £(σ). And so on, in the obvious way.

The above passage is, if anything, even more explicit than Partee is about the nature of meanings and the nature of semantic interpretation. However, having offered this "thesis", Lewis develops an "antithesis" that is radically different. It focuses on interpretation as a complex social convention and seems to imply that we can replace "meaning" with specific behavioral responses:

II. ANTITHESIS

What is language? A social phenomenon which is part of the natural history of human beings; a sphere of human action, wherein people utter strings of vocal sounds, or inscribe strings of marks, and wherein people respond by thought or action to the sounds or marks which they observe to have been so produced.

This verbal activity is, for the most part, rational. He who produces certain sounds or marks does so for a reason. He knows that someone else, upon hearing his sounds or seeing his marks, is apt to form a certain belief or act in a certain way. He wants, for some reason, to bring about that belief or action. Thus his beliefs and desires give him a reason to produce the sounds or marks, and he does. He who responds to the sounds or marks in a certain way also does so for a reason. He knows how the production of sounds or marks depends upon the producer's state of mind. When he observes the sounds or marks, he is therefore in a position to infer something about the producer's state of mind. He can probably also infer something about the conditions which caused that state of mind. He may merely come to believe these conclusions, or he may act upon them in accordance with his other beliefs and his desires.

Not only do both have reasons for thinking and acting as they do; they know something about each other, so each is in a position to replicate the other's reasons. Each one's replication of the other's reasons forms part of his own reason for thinking and acting as he does; and each is in a position to replicate the other's replication of his own reasons. Therefore the Gricean mechanism[1] operates: *X* intends to bring about a response on the part of *Y* by getting *Y* to recognize that *X* intends to bring about that response; *Y* does recognize *X*'s intention, and is thereby given some sort of reason to respond just as *X* intended him to.

Within any suitable population, various regularities can be found in this rational verbal activity. There are regularities whereby the production of sounds or marks depends upon various aspects of the state of mind of the producer. There are regularities whereby various aspects of responses to sounds or marks depend upon the sounds or marks to which one is responding. Some of these regularities are accidental. Others can be explained, and different ones can be explained in very different ways.

Some of them can be explained as conventions of the population in which they prevail. Conventions are regularities in action, or in action and belief, which are arbitrary but perpetuate themselves because they serve some sort of common interest. Past conformity breeds future conformity because it gives one a reason to go on conforming; but there is some alternative regularity which could have served instead, and would have perpetuated itself in the same way if only it had got started.

The "Gricean mechanism" referred to here is from Grice 1989, which develops a theory of linguistic meaning that is grounded in recognizing speaker intentions.

Lewis ultimately brings these together in a "synthesis": the interpretation function captures a complex set of normative conventions about how to use language, and he proposes in addition some behavioral conventions – e.g., produce truthful utterances for which you have evidence, interpret according to the conventions.

References

Allén, Sture. 1989. *Possible worlds in humanities, arts, and sciences: Proceedings of Nobel symposium*. Walter de Gruyter.

Carnap, Rudolf. 1956. *Meaning and necessity*. Chicago: Chicago University Press.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1986. *Knowledge of language*. New York: Praeger.

Cresswell, M. J. 1982. The autonomy of semantics. In Stanley Peters & Esa Saarinen (eds.), *Processes, beliefs, and questions*, 69–86. Reidel.

Frege, Gottlob. 1892/1980. On sense and reference. In Peter Geach & Max Black (eds.), *Translations from the philosophical writings of Gottlob Frege*, 56–78. Oxford: Blackwell 3rd edn.

Grice, H. Paul. 1989. Meaning. In *Studies in the way of words*, chap. 14, 213–223. Cambridge, MA: Harvard University Press. Reprinted from *The Philosophical Review* 66:377-88, 1957.

Hintikka, Jaakko. 1969. *Models for modalities*. Dordrecht: Reidel.

Jackendoff, Ray S. 1996. Semantics and cognition. In Shalom Lappin (ed.), *The handbook of contemporary semantic theory*, 539–559. Oxford: Blackwell Publishers.

Kripke, Saul. 1969. Semantical considerations on modal logic. *Acta Philosophica Fennica* 16(83–94).

Lewis, David. 1970. General semantics. *Synthese* 22(1). 18–67. doi:10.1007/BF00413598.

Lewis, David. 1975. Languages and language. In Keith Gunderson (ed.), *Minnesota studies in the philosophy of science*, vol. VII, 3–35. Minneapolis: University of Minnesota Press. Reprinted in Lewis 1983, 163–188. Page references are to the reprinting.

Lewis, David. 1983. *Philosophical papers*, vol. 1. New York: Oxford University Press.

Montague, Richard. 1970. English as a formal language. In Bruno Visentini et al. (ed.), *Linguaggi nella società e nella tecnica*, 189–224. Milan: Edizioni di Communità. Reprinted in Montague (1974), 188–221.

Montague, Richard. 1974. *Formal philosophy: Selected papers of Richard Montague*. New Haven, CT: Yale University Press.

Partee, Barbara H. 1995. Lexical semantics and compositionality. In Lila R. Gleitman & Mark Liberman (eds.), *Invitation to cognitive science*, vol. 1, 311–360. Cambridge, MA: MIT Press.