Encyclopedia of Language & Linguistics (Second Edition), 2006, Pages 564-579

Jain S, Osherson D N, Royer J S & Kumar Sharma A (1999). *Systems that learn* (2nd edn.). Cambridge, MA: MIT Press.
Lightfoot D (1991). *How to set parameters: arguments from language change.* Cambridge, MA: MIT Press.
Morgan J L (1986). *From simple input to complex grammar.* Cambridge, MA: MIT Press.
Osherson D N, Stob M & Weinstein S (1985). *Systems that learn.* Cambridge, MA: MIT Press.
Solomonoff R J (1978). 'Complexity-based induction systems: comparisons and convergence theorems.' *IEEE Transactions on Information Theory 24,* 422–432.
Wexler K & Culicover P (1980). *Formal principles of language acquisition.* Cambridge, MA: MIT Press.

Formal Semantics

G Chierchia, Universita degli Studi di Milano-Bicocca, Milan, Italy

Introduction

Semantics, in its most general form, is the study of how a system of signs or symbols (i.e., a language of some sort) carries information about the world. One can think of a language as constituted by a lexicon (an inventory of morphemes or words) and a combinatorial apparatus according to which complex expressions, including, in particular, sentences, can be built up. Semantics deals with the procedures that enable users of a language to attach an interpretation to its arrays of symbols. Formal semantics studies such procedures through formally explicit mathematical means.

The history of semantics is nearly as long and complex as the history of human thought; witness, e.g., the early debates on the natural vs. conventional character of language among the pre-Socratic philosophers. The history of formal semantics is nearly as daunting as it is intertwined with the development of logic. In its modern incarnation, it is customary to locate its inception in the work of logicians such as Frege, Russell, and Tarski. A particularly important and relatively recent turning point is constituted by the encounter of this logico-philosophical tradition with structural and generative approaches to the study of human languages, especially (though by no means exclusively) those influenced by N. Chomsky. The merger of these two lines of research (one brewing within logic, the other within linguistics), has led formal semantics to become a central protagonist in the empirical study of natural language. The research paradigm that has emerged has proven to be quite fruitful, both in terms of breadth and depth of results and in terms of the role it is playing in the investigation of human cognition. The present work reviews some of the basic assumptions of modern formal semantics of natural language and illustrates its workings through a couple of examples, with no pretence of completeness.

Semantics vs. Lexicography

One of the traditional ideas about semantics is that it deals with the meaning of words. The main task of semantics is perceived as the compilation of dictionaries (semantics as lexicography). To this, people often add the task of investigating the history of words. Such a history can teach us about cultural development. One might even hope to arrive at the true meaning of a word through its history. Compiling dictionaries or reconstructing how particular words have changed over time are worthy tasks; but they are not what formal semantics is about. Lexicography, philology, and related disciplines vs. semantics as conceived here constitute complementary enterprises. They all, of course, deal with language. But the main goal of semantics is to investigate how we can effortlessly understand a potential infinity of expressions (words, phrases, sentences). To do that, we have to go beyond the level of single words.

It may be of use to point to the kind of considerations that have led semantics to move the main focus of investigation away from single word meanings and their development. For one thing, it can be doubted that word histories shed light on how words are synchronically (i.e., at a given point in time) understood and used. People use words effectively in total ignorance of their history (a point forcefully made by one of the founding fathers of modern linguistics, namely F. de Saussure). To make this point more vividly, take the word *money.* An important word indeed; where does it come from? What does its history reveal about the true meaning of money? It comes from Latin *moneta*, the past participle feminine of the verb *moneo* 'to warn/to advise.' *Moneta* was one of the canonical attributes of the Roman goddess Juno; *Juno moneta* is 'the one who advises.' What has Juno to do with money? Is it perhaps that her capacity to advise extends to finances? No. It so happens that in ancient Rome, the mint was right

Encyclopedia of Language & Linguistics (Second Edition), 2006, Pages 564-579

next to the temple of Juno. So people metonymically transferred Juno's attribute to what was coming out of the mint. A fascinating historical fact that tells us something as to how word meanings may evolve; but it reveals no deep link between money and the capacity to advise. This example is not meant to downplay the interest of historical investigations on word meanings; it is just an illustration of how linguistic history affects only marginally the way in which a community actually understands its lexicon.

There is a second kind of consideration suggesting that the scope of semantics cannot be confined to the study of word meanings. Do words in isolation have clearly identifiable meanings? Take any simple word, say the concrete, singular, common noun *dog*. What does it mean? Some possible candidates are: the dog-kind, the concept of dog, the class of individual dogs. ... And the list can go on. How do we choose among these possibilities? Note, moreover, that all these hypotheses attempt to analyze the meaning of the word *dog* by tacking onto it notions (kind, concept, class ...) that are in and of themselves in need of explication. If we left it at that, we wouldn't go far. Looking at dictionary definitions is no big help either. If we look up the entry for *dog*, typically we will find something like:

(1) A highly variable carnivorous domesticated mammal (Canis familiaris) prob. descended from the common wolf.

Indeed, if someone doesn't know the meaning of the word *dog* and knows what *carnivorous* and *mammal* mean, then (1) may be of some practical help. But clearly to understand (1), we must rely on our understanding of whole phrases and the words occurring in them. Words which, in turn, need a definition to be understood. And so on, in a loop. This problem is sometimes called the problem of the circularity of the lexicon. To put it differently, (1) is of help only if the capacity to use and interpret language is already taken for granted. But it is precisely such capacity that we want to study.

The limitation of a purely word-based perspective on the investigation of meaning is now widely recognized. Frege summarized it in a nice motto: "only in the context of a sentence do words have meaning." His insight is that complete sentences are linguistic units that can sort of stand on their own (more so than any other linguistic units). They can, as it were, express self-contained thoughts. We are more likely, therefore, to arrive at the meaning of single words (and of phrases in between words and complete sentences) via a process of abstraction from the contribution that words make to sentence meaning, rather than the other way around. This is so because sentence meaning is somehow more readily accessible (being, as it were, more complete) than the meaning of words in isolation.

These are some reasons, then, why the perspective of modern semantics is so different from and complementary to lexicography and philology; such perspective is much more directly tied to the investigation of the universal laws of language (language universals) and of the psychological mechanisms underlying such laws. Understanding the function, use, etc., of a single word presupposes a whole, complex cognitive apparatus. It is, therefore, an arrival point more than a starting point. It seems thus reasonable to start by asking what it is to understand a sentence.

The main thesis we wish to put forth is that to understand a sentence involves understanding its relations to the other sentences of the language. Each sentence carries information. Such information will be related to that of other sentences while being unrelated to that of yet others. In communicating, we rely on our spontaneous (and unconscious) knowledge of these relations.

The Notion of Synonymy and Its Problems

Imagine watching a Batman movie in which the caped hero fights the Riddler, one of his eternal foes. The Riddler has scattered around five riddles with clues to his evil plans. Batman has managed to find and solve four of them. We could report this situation in any of the following ways:

(2a) Batman has found all of the five clues but one.
(2b) Batman has found four out of the five clues.
(2c) Four of the five clues have been found by Batman.

These sentences are good paraphrases of each other. One might say that they have roughly the same information content; or that they describe the same state of affairs; or that they are (nearly) synonymous. (I will be using these modes of speaking interchangeably.) To put it differently, English speakers know that there is a tight connection between what the sentences in (2a), (2b), and (2c) mean. This is a kind of knowledge they have *a priori,* i.e., regardless of what actually goes on. Just by looking at (2a) vs., say, (2b) and grasping what they convey, we immediately see that they have roughly the same informational content.

This is what we mean when we say that understanding a sentence involves understanding which other sentences count as good paraphrases and which don't. Thus, knowing a language is to know which sentences in that language count as synonymous. Semantics is (among other things) the study of synonymy. Two synonymous sentences (and, more generally, two synonymous expressions) can always

Encyclopedia of Language & Linguistics (Second Edition), 2006, Pages 564-579

be used interchangeably. This last informal characterization can be turned into a precise definition along the following lines.

(3a) Suppose one utters any complex expression α containing a subexpression A. If one can replace in A α with a different expression β, without changing the overall communicative import of A, then α and β are synonymous.
(3b) α is synonymous with β = in the utterance of any expression A containing α, α can be replaced with β without changing the communicative import of the utterance (*salva significatione*).

For example, in uttering (2a) (our A), we can replace the subcomponent that comes after *Batman has found__* namely *all of the five clues but one* (our α) with *four out of the five clues* (our β) and convey exactly the same information. Hence, these two expressions must be synonymous (and, in fact, so are the whole sentences).

This looks promising. It paves the way for the following setup for semantics. Speakers have intuitions of whether two expressions can be replaced with each other while keeping information content unchanged. For any two sentences α and β, they spontaneously know whether they can be substituted for each other (i.e., whether β can be used to paraphrase α). Because the sentences of a language are potentially infinite, it is impossible for speakers to memorize synonymous sentences one by one (for that clearly exceeds what our memory can do). Hence, they must recognize synonymy by rule, by following an algorithm of some sort. The task of semantics, then, becomes characterizing such an algorithm.

There is a problem, however. Sameness of communicative import is a more or less thing, much like translation. In many contexts, even sentences as close as those in (3a) and (3b) could not be replaced felicitously with each other. Here is a simple example. The discourse in (4a) is natural and coherent. The one in (4b) much less so:

(4a) Batman has found all of the five clues but one, which is pinned on his back.
(4b) ?? Batman has found four out of the five clues, which is pinned on his back.
(modeled after a famous example by B. Partee)

Clearly in (4a) we cannot replace *Batman has found all of the five clues but one* with *Batman has found four out of the five clues* while keeping unaltered the overall communicative effect. This means that if we define synonymy as in (3a) and (3b), then (2a) and (2b) cannot be regarded as synonymous after all. Yet they clearly share a significant part of their informational content. What *is* it that they share?

In fact, it has been argued that if (3a) and (3b) are how we define synonymy, then there simply are no two sentences that qualify as such. Here is a classical argument that purports to show this (based on Mates (1950)). Take the following two sentences:

(5a) Billy has a dog.
(5b) Billy has a highly variable carnivorous domesticated mammal prob. descended from the common wolf.

Are these two sentences synonymous? Hardly. They are clearly semantically related. But they surely do not have the same communicative import. Nor can one replace the other in every context. For example, (5a) could describe a true state of affairs, while (5b) might not:

(6a) Molly believes that Billy has a dog.
(6b) Molly believes that Billy has a highly variable carnivorous domesticated mammal prob. descended from the common wolf.

This shows that in contexts like *Molly believes that __* we cannot simply replace a word with its dictionary definition. And if dictionary definitions don't license synonymy, then what does?

The problem can be couched in the following terms. Any normal speaker of English perceives a strong semantic connection among the sentences in (2a), (2b), and (2c), or (4a) and (4b). So strong that one might feel tempted to talk about synonymy. Yet when we try to make the notion of synonymy precise, we run into serious problems. Such a notion appears to be elusive and graded (a more or less thing); so much so that people have been skeptical about the possibility of investigating synonymy through precise, formal means.

A fundamental breakthrough has been identifying relatively precise criteria for assessing semantic relations. The point is that perfect synonymy simply does not exist. No two sentences can be always replaced with each other. The notion of synonymy has to be deconstructed into a series of more basic semantic relations. We need to find a reliable source for classifying such relations, and, we will argue, such a source lies in the notions of truth and reference. Consider the sentences in (2a), (2b), and (2c) again. Assume that the noun phrase *the five clues* in (2a) and (2b) refer to the *same* clues (i.e., we are talking about a particular episode in a particular story). Then, could it possibly happen that say (2a) is true and (2b) false? Evidently not: no one in his right mind could assert (2a) while simultaneously contending that (2b) is false. If (2a) is true, (2b) also must be true. And, in fact, vice versa: if (2b) is true, then (2a) also must be. When this

Encyclopedia of Language & Linguistics (Second Edition), 2006, Pages 564-579

happens, i.e., when two sentences are true in the same set of circumstances, we say that they have the same truth conditions.

Notice that sameness of truth conditions does not coincide with or somehow require sameness of communicative import (too elusive a notion), nor substitutivity in any context whatsoever (a condition too difficult to attain). Our proposal is to replace such exceedingly demanding notions with a series of truth-based notions, while keeping the same general setup we sketched in connection with synonymy: for any pair of sentences, speakers have intuition about whether they are true under the same conditions or not. They can judge whether they are true in the same (real or hypothetical) circumstances or not. Because the sentences of our language are infinite, this capacity must be somehow based on a computational resource. Speakers must be able to compare the truth-conditions associated with sentences via an algorithm of some sort. The task of semantics is to characterize such an algorithm. The basic notion changes (synonymy is replaced with sameness of truth conditions), but the setup of the problem stays the same.

Truth and Semantic Competence

Let us elaborate on the proposal sketched at the end of the previous section. Information is transmitted from one agent to another (the 'illocutionary agents') in concrete communicative situations ('speech acts'). No two such situations are alike. And consequently, no two pieces of information that are transmitted through them are alike. In *Groundhog Day,* a movie with the actor Bill Murray, the protagonist gets trapped into going through the same day over and over. He wakes up and his day starts out in the same way (with the alarm clock ringing at 7 a.m. on groundhog day); as he walks outside, he meets the same waitress who greets him in the same way ("weather so-so, today"). Yet this sentence, though being the same day after day, and being uttered in circumstances as identical as they can conceivably be, clearly conveys a different sense or information unit on each occasion of its use (the hearer going from noticing that something is fishy about this verbatim repetition, to the painful discovery of the condemnation to live through groundhog day for eternity).

Ultimately, we want to understand how communication takes place. But we cannot nail down every aspect of a speech act, just as we cannot know (not even in principle, I believe) every aspect of the physical or mental life of a particular human being. At the same time, while speech acts are unique events, there is much that is regular and invariant about them; that is what can be fruitfully investigated. One family of such invariants concerns form: similar sound patterns may be used in different speech acts. Another family of invariants concerns content: similar states of affairs may be described through a variety of expressions. The notion of truth is useful in describing the latter phenomenon. A pair of sentences may be judged as being necessarily true in the same circumstances. This is so, for example, for (5a) vs. (5b). Yet, such sentences clearly differ in many other respects. One is much more long-winded than the other; it uses rarer words, which are typical of high, formal registers. So in spite of having the same truth conditions, such sentences may well be used in different ways. Having the same truth condition is generally regarded as a semantic fact; being able to be used in different ways is often regarded as a pragmatic fact. While this gives us a clue as to the role of these two disciplines (both of which deal with meaning broadly construed), the exact division of labor between semantics and pragmatics remains the object of controversy.

Truth conditions are a tool for describing semantic invariants, structural regularities across communicative situations. Whenever I utter a declarative sentence, I typically do so with the intention to communicate that its truth conditions are satisfied (which of course raises the question of nondeclaratives, emotive expressions, and the like; see for example textbooks such as Chierchia and McConnell-Ginet (2000) or Heim and Kratzer (1998); cf. also Kratzer (1999) for a discussion of relevant issues). Truth conditions depend on the reference (or denotation) of words and the way they are put together, i.e., they are compositionally projected via the reference of words (or morphemes). If I say to you, as we are watching a movie, "Batman has found all of the five clues but one," you understand me because you sort of know (or guess) who Batman is, what sort of things clues are, what finding something is, what number the word *five* refers to; you also understand the "all ... but ..." construction. The reference/denotation of words is set (and is modified, as words may change their denotation in time) through use, in complex ways we cannot get into within the limits of the present work. The denotation of complex expressions (e.g., of a verb phrase such as [$_{VP}$*found five clues*] and truth conditions of sentences) are set by rule (the semantic component of grammar). Semantic rules presumably work like syntactic rules: they display variation as well as a common core, constitutive of universal grammar. Insofar as semantics is concerned, what is important for our purposes is that truth conditions can be compositionally specified. This paves the way for an algorithmic approach to meaning. We already remarked that sentences are formed by composing morphemes together via a limited number of syntactic operations. So to arrive at the truth condition of an arbitrary sentence, we can start by the contribution of the words (their reference). Then, for each way of putting words together, there will be a way of forming the reference of complex expressions, and so on until we arrive at the truth condition of the target sentence.

So far, we have discussed sentences that have the same truth conditions (such as those in (2a), (2b), and (2c)); but this is not the only semantic relation that can be characterized in terms of the notion of truth. Consider the following examples.

(7a) Every Italian voted for B	a'. Most Italians voted for B.
(7b) Leo voted for B	b'. Leo voted for B.

Sentence (7a) is related to (7b) in a way that differs from the relation between (7a') vs. (7b'). Here is the difference. If (7a) is true, and Leo is Italian, then (7b) has to be true, too; this is clearly not so for (7a') vs. (7b'): (7a') may be true without (7b') being true. If whenever A is true, B also must be, we say that A *entails* B (B's meaning is part of of A's meaning). Two sentences with the same truth conditions entail each other (i.e., they hold a symmetric relation); when entailment goes only one way (as from (6a) to (6b)), we have an asymmetric relation.

Entailment is pervasive. Virtually all semantic intuitions are related to it. As an illustration, consider the pair of sentences in (8a) and (8b).

(8a) John promised Bill to take him to the station.
(8b) John ordered Bill to take him to the station.

Pronouns, like *him* in (8a) and (8b), take their denotation from the context; they can take it from the extra linguistic context (a person salient in the visual environment, a person the speaker points at, etc.) or from the linguistic context (e.g., from NPs that occur in the same discourse; *John* or *Bill* in (8a) and (8b)); one widespread terminology is to speak of *indexical* uses in the first case and of *anaphoric* uses in the second. We can conceptualize this state of affairs by viewing pronouns as context-dependent items, incomplete without pointers of some sort. Now we shall focus on the anaphoric interpretation of (8a) vs. (8b).

(9a) John promised Bill that John would take Bill to the station.
(9b) John ordered Bill that Bill should take John to the station.

These appear to be the only options. That is to say, sentence (8a) cannot convey something like 'John promised Bill that Bill should take John to the station.' The point of this example is that we have intuitions that govern how the denotation of a pronoun is to be reconstructed out of contextual clues; such intuitions tell us that (8a) and (8b), though structurally so similar, allow for a distinct range of interpretive options. At the basis of intuitions of this sort, we again see entailment at work: on its anaphoric construal, (8a) entails (9a).

Another important set of truth-based semantic relations are presuppositions. Consider the contrast between the sentences in (10a) and (10b).

(10a) Fred stole the cookies.
(10b) It was Fred who stole the cookies.

There is a noticeable semantic contrast between (10a) and (10b). How can we characterize it? Clearly the two sentences are true in the same circumstances (they entail each other). Yet they differ semantically. Such a difference can be perhaps caught by looking at what happens by embedding (10a) and (10b) in a negative context.

(11) So, what happened this morning?
(11a) Everything went well. Fred didn't steal the cookies; he played with his toys.
(11b) ?? Everything went well. It wasn't Fred who stole the cookies.

The answer in (11a) is natural. The one in sentence (11b) would sound more natural as an answer to

(12a) Who stole the cookies?
(12b) It wasn't Fred.

The difference between the question in (11) and the one in (12a) is that the latter (but not the former) tends to presuppose that cookies where stolen. In other terms, the situation seems to be the following. Both sentences in (10a) and (10b) entail:

(13) Someone stole the cookies.

If either (11a) or (11b) are true, then (13) must also be true. Furthermore, sentence (13) must be true for (10b) to be *denied* felicitously. The illocutionary agents must take for granted the truth of (13) to assert, deny, or otherwise use sentence (10b), as the naturalness of the following continuations for (13) illustrate:

(14) Someone stole the cookies . . .
(14a) It was Fred.
(14b) It wasn't Fred.
(14c) Was it Fred?
(14d) If it was Fred, he is going to get it . . .

This brings us to the identification of presupposing as a distinctive semantic relation: a sentence A presupposes B if the truth of B must be taken for granted

Encyclopedia of Language & Linguistics (Second Edition), 2006, Pages 564-579

in order to felicitously assert, deny, etc., A. Presuppositions are quite important in language. So much so that there are distinctive syntactic constructions (such as those in (10b), known as cleft sentences) specifically keyed to them.

Let me illustrate the wealth of semantic relations and their systematic character by means of another example, which will bring us to the interface between semantics and pragmatics. Consider:

(15a) Who stole the cookies?
(15b) Fred looks mischievous.
(15c) Fred stole the cookies.

If to a question such as (15a), I reply with (15b), I do suggest/convey something like (15c). Sentence (15c) clearly is not part of the literal meaning of (15b) (however hard defining such a notion might be). Yet, in the context of the dialogue in (15a), (15b), and (15c), speakers will converge in seeing that (15c) is strongly suggested by (15b). Here, too, we have, thus, a systematic semantic intuition. The suggestion in (15c) can be retracted; that is, one can continue (15b) with '... but I know he didn't do it'. However, in the absence of such an explicit correction, illocutionary agents upon hearing (15b) will tend to infer (15c). This phenomenon has been studied by H. P. Grice (1989), who dubbed it implicature. His proposal is that it arises through interaction of the core meaning assigned to sentences by rule with principles that govern conversational exchanges. The basic idea is that for conversational exchanges to be successful they have to be basically cooperative acts; cooperating means that one sticks to relevant topics, one only gives information believed to be truthful, one gives no more and no less than what is relevant, etc. Applying this to the case at hand, in a situation in which question (15a) is topical, answering (15b) would seem to be blatantly irrelevant; the hearer, however, tends to interpret it as relevant and sets in motion an inferential process that tends to link it to some piece of information that does address the topical question; such a link is to be found with the help of the information available in the context to the illocutionary agents (e.g., in the common knowledge that if people commit a mischief, such as stealing cookies, they may well look mischievous, etc.). Thus, this type of semantic judgment (the implicature) appears to be best accounted for in terms of the interaction between grammar and general conditions on reasonable language use (that fall under the scope of pragmatics).

Sometimes it is not immediately clear whether something is a matter of conventionalized meaning or pragmatics. To illustrate, consider the oscillation in meaning of a word like *or*. It can be illustrated with the following examples. Consider first (16a):

(16a) If I got it right, either John or Mary will be hired.
(16b) If I got it right, either John or Mary but not both will be hired.

Normally, one tends to interpret (16a) as truth conditionally equivalent to (16b); i.e., the disjunction in (16a) is interpreted exclusively (as incompatible with the simultaneous truth of each disjunct). However, this is not always so. Contrast (16a) with (17a).

(17a) If either John or Mary are hired, we'll celebrate.
(17b) (?) If John or Mary (but not both) are hired, we'll celebrate.
(17c) If John or Mary or possibly both are hired, we'll celebrate.

The most natural interpretation of (17a) is not the exclusive one (namely (17b), which is somewhat odd pragmatically); rather it is the inclusive one, made explicit in (17c). (Notice that the emphatic word *either* is present both in (16a) and (17a); in spite of this, the interpretation of *or* shifts.) We might see in these phenomena a lexical ambiguity of disjunction. Words expressing disjunction, we may feel inclined to conclude, have a varying interpretation, as it happens with words such as *bank* or *lap* ('sit on my lap' vs. 'he swam three laps'). We may assume that such interpretations are always in principle available, but then we select the most suitable to the context of the speech act. While this seems *prima facie* possible, there are reasons to doubt it. In particular, true lexical ambiguities are resolved across languages (in Italian, there are two different words for the two senses of *lap*). Ambiguities are never universal. The meaning shift of *or*, *per contra*, seems to be universal: in every language disjunction appears to have a similar oscillation in meaning. A convincing case for two lexically distinct disjunctions, one exclusive, the other exclusive, has not been made (sometimes it has been proposed that Latin *vel* vs. *autem* is just that; for arguments against this, cf., e.g., Jennings (1994)). Moreover, other areas of the lexicon have been found that display a similar behavior (e.g., the number words). This strongly suggests that a different explanation for such behavior should be found. Grice himself has proposed that the phenomenon under discussion is to be accounted for in terms of the interaction between semantics and pragmatics. The idea is that the basic meaning of *or* is the inclusive one, as it is the most liberal interpretation; the exclusive construal arises as an implicature, i.e., a

Encyclopedia of Language & Linguistics (Second Edition), 2006, Pages 564-579

pragmatic enrichment, albeit a generalized one. The advantage of this move is that it would explain the oscillation in meaning of disjunction without positing a covert ambiguity. We will come back to how the generalized implicature associated with *or* might come about in the later section "The Semantics/ Pragmatics Interface".

Wrapping up, the picture that emerges is roughly the following. In using language, speakers display complex forms of spontaneous knowledge. They put together words in certain ways and not others. This is how knowledge of syntax manifests itself. They also accept certain paraphrases and not others, draw certain inferences and not others, etc. It turns out to be possible/useful to categorize the latter in three major families of semantic relations.

(18a) Entailment-based (entailment, mutual entailment, contradictoriness, analyticity, etc.)
(18b) Presupposition-based (presupposition, question/answer pairs, etc.)
(18c) Implicature-based (generalized implicature, particularized implicature, etc.)

All of them can be readily defined in terms of the notion of truth:

(19a) A entails B = for any conceivable situation s, if A is true in s, B is also true in s.
(19b) A presupposes B = to use A appropriately in a situation s, the truth of B must be taken for granted by the illocutionary agents in s.
(19c) A implicates B = use of A in a situation s suggests, everything else being equal, that B is true in s.

The definitions in (19a), (19b), and (19c) can be readily associated with "operational" tests that enable speakers to assess whether a given relation obtains or not. For example, to check whether (20a) entails (20b), you might check whether you could sincerely assert (20a) while denying (20b), viz. whether you could sincerely and felicitously utter something like (20c):

(20a) It is indeed odd that Mary is home.
(20b) Mary is home.
(20c) It is indeed odd that Mary is home, even if she in fact isn't.

To the extent that you can't really say something like (20c), you are entitled to conclude that (20a) entails (20b). It is useful, in these cases, to use contrast sets such as (21a) and (21b).

(21a) It is indeed conceivable that Mary is at home.
(21b) It is indeed conceivable that Mary is home, even if she in fact isn't.

The semantic relations in (18a), (18b), and (18c) can be viewed as intuitions of semantic relatedness speakers have about sentences of their own language, as judgments that may be elicited, and the like. By analogy with well-formedness judgments, there are some cases in which things are not so clear and we may not be sure whether, say, a certain entailment holds or not. In such a case, more complex arguments, indirect evidence of various sorts, or psycholinguistic experimentation may be called for (see, e.g., Crain and Thornton (1998) on experimental methodologies for truth-based semantic judgments). But in indefinitely many cases, simple introspection yields relatively straightforward judgments. The capacity for making such judgments is constitutive of our semantic competence. Such a competence cannot be simply a thesaurus, a store of pairs of sentences, with the relative judgment tacked on, for the number of judgments speakers can make on the fly is potentially infinite. Semantic competence must be a computational device of some sort. Such a device given an arbitrary pair of sentences <A, B> must be able to determine in principle whether A entails B, presupposes it, etc. The task of semantics is to characterize the general architecture of such a computational device. While there are many foundational controversies that permeate the field, there is a broad convergence that this is roughly the form that the problem of meaning takes within modern formal semantics.

Semantic Modeling

In the present section I will sketch how a (necessarily, much simplified) calculus of semantic relations may look. Suppose you have a lexicon of the following form:

(22a) N: John, Bill, dog, cat, table, ...
(22b) V: runs, smokes, drinks, . . .
(22c) DET: the, a, some, every, no . . . .

Think of syntax as a device that combines lexical entries by merging them in complex phrases and assigning them a syntactic analysis that can be represented by tree diagrams or labeled bracketings of the following form:

(23a) $[_{VP}$ John smokes]
(23b) $[_{DP}$ every boy]
(23c) $[_{VP}$ $[_{DP}$ every boy] smokes]

I assume, without being able to justify it, that lexical items have phrasal projections. In particular, VP is the phrasal projection of V and constitutes a clausal nucleus composed of the verb and its arguments linked in a predicative structure. Such a nucleus forms the

Encyclopedia of Language & Linguistics (Second Edition), 2006, Pages 564-579

innermost skeleton of the sentence (I will have to ignore matters pertaining to inflection, agreement, tense, and the like). The lexical features of verbs are crucial in determining the characteristics of clausal nuclei. DP is the phrasal projection of D, and it is constituted by a determiner and a (common) noun. Clausal nuclei can be formed by merging a verb with a (proper) name or a DP, as indicated. In the spirit of the discussion in the section on **Truth and Semantic Competence**, semantics assigns recursive truth conditions to sentences in terms of the reference assigned to lexical entries. There are several ways to do this. Ultimately, the choice one makes on the exact format of interpretive rules has far-reaching consequences for our understanding of grammar. However, our choices here are only in small part dictated by our current understanding of semantics in universal grammar; for the major part, they result from considerations such as ease of exposition, keeping prerequisites at a minimum, and the like. To get started, we should assign a reference (or denotation, terms we will use interchangeably) to lexical entries. To do so, we assume we have a certain domain $D_s = \{a, b, c, \ldots\}$ at each given discourse situation s that constitutes our universe of discourse. A discourse situation can be thought of as the time at which the utterance takes place. A domain is just a set of individuals, pragmatically selected (e.g., those salient to the illocutionary agents). Interpretations are relative to an utterance situation s and the corresponding domain of discourse D_s. Reference of proper nouns, for example, is suitably chosen from the domain of discourse. Suppose, for example, that a and b are salient humans in our universe of discourse, then we might have:

(24) For any conceivably relevant utterance situation s, the name *John* denotes a in s; the name *Bill* denotes b in s . . .

It doesn't matter how a or b are characterized (via a description, an act of indication, etc.) to the extent that one successfully succeeds in linking the noun to its bearer. Also, it is useful to have a uniform category-neutral notation for semantic values; we will use for this the double bar notation || ||; accordingly, for any expression α, $\|\alpha\|^s$ will be the semantic value of α in situation s. Thus, (24) can be abbreviated as:

(25) $\|\text{John}\|^s = a$ (where $a \in D_s$, the domain of discourse at s)

(Technically, || || can be viewed as a function from expressions and situations into denotations; so sometimes we will speak of the interpretation function.) The denotation of a simple (intransitive) verb such as those in (22b) can be thought of as a function that for each (appropriate) individual in the domain discourse tells us whether that individual performs a certain action or not. Here is an example:

(26) *smokes* in a situation s denotes a function $\textbf{smoke}_s$ that applies to animate individuals and returns truth values. If a is such an individual, then $\textbf{smoke}_s(a)$ returns 'true' (which we represent as the number 1) if that individual performs the action of smoking in s (where smoking involves . . . .); otherwise $\textbf{smoke}_s$ (a) returns 0 (i.e., 'false').

If a is not animate (e.g., if a is a stone and s is a 'normal' situation), then $\textbf{smoke}_s$ (a) is not defined (lacks a value). The final part in definition (26) reflects the fact that sentences like (27a) and (27b), out of the blue, are (equally) strange: smoking normally requires its subject argument to be animate.

(27a) That stone smokes.
(27b) That stone doesn't smoke.

The deviance of sentences like (27a) and (27b) has been variously characterized as a violation of selectional restrictions or as sortal deviance. Here we are couching the relevant phenomenon in presuppositional terms (to illustrate a further application of such a concept). The fact that sentences of this sort remain deviant across negation may be taken as evidence that the verb *smoke* imposes an animacy presupposition on its arguments (see e.g., Chierchia and McConnell-Ginet (2000) for more discussion). A definition like (26) can be stated more compactly:

(28) $\|\text{smokes}\|^s = \textbf{smoke}_s$,
where for each a in D_s, $\textbf{smoke}_s(a)$ is defined iff a is animate in s; if defined, $\text{smoke}_s(a) = 1$ if a smokes in s (where smoking involves . . .); $\textbf{smoke}_s$ (a) $= 0$, otherwise.

The definition of (or constraints on) smoking (i.e., the dots in (28)) can be elaborated further in several ways by refining our lexical analysis of the verb *smoke*. Although much progress has been made on this score, many important issues remain open (including, e.g., whether a presuppositional treatment of selectional restrictions is ultimately viable). What is important, from the point of view of compositional semantics, is the *logical type* or *semantic category* of the denotation of a verb like *smoke*. Such verbs are treated here as functions from individuals into truth values. These are called *characteristic functions*; they divide the (relevant portion of) the domain of discourse of the utterance situation in two: the things that satisfy the verb from those that don't. Characteristic functions correspond to sets (which might be called the *extension* of the function), as the following example illustrates:

Encyclopedia of Language & Linguistics (Second Edition), 2006, Pages 564-579

(29) Let universe of discourse be {a, b, c, d}; let a, b, and c be people. Of these, let a and b smoke in s while b but not a also smokes in a different situation s'. We can represent all this as follows:

$$\mathbf{smoke}_s = \begin{matrix} a \rightarrow 1 \\ b \rightarrow 1 \\ c \rightarrow 0 \end{matrix} \quad \text{corresponding extension: } \{a,b\}$$

$$\mathbf{smoke}_{s'} = \begin{matrix} a \rightarrow 1 \\ b \rightarrow 0 \\ c \rightarrow 0 \end{matrix} \quad \text{corresponding extension: } \{a\}$$

As is evident from the example, sets and characteristic functions are structurally isomorphic (encode the same information). In what follows it will be useful on occasion to switch back and forth between these two concepts. Use of characteristic functions as a formal rendering of verb meanings is useful in giving truth conditions for simple subject predicate sentences:

(30a) A sentence of the form $[_{VP}$ N V $]$ is true in s iff $||V||^s\,(||N||^s) = 1$
Example:
(30b) $[_{VP}$ Bill drinks $]$ is true in s iff $||drinks||^s\,(||Bill||^s) = 1$

The truth conditions of any sentence with the syntactic structure specified in (30a) boil down to applying a characteristic function to an individual (and thereby ascertaining whether that individual belongs to the set that constitutes the extension). To find out whether Bill in fact smokes in s, we need factual information about the situation obtaining in s. To understand the sentence, we don't. We merely need to know its truth conditions, which in the case of simple subject–predicate sentences are an instruction to check the value of a characteristic function for the argument specified by the subject. The rules in (30a) and (30b) can be reformulated more compactly as in (31):

(31) $||\,[_{VP}\ N\ V\,]\,||^s = ||V||^s\,(||N||^s)$

This can be viewed as the kernel of a predication rule (that tells us how subject and predicates combine semantically).

Everything so far looks like a formally explicit (and perhaps somewhat pedantic) way of sketching a denotational, information-oriented semantics, and the reader may get the feeling of not yet finding striking insights on what meaning is. In order to grasp the potential of this method, one needs to look at a little more of its computational apparatus. So let us turn now to DPs. Things here are definitely more challenging. DPs are constituents formed by a determiner plus a common noun. Common nouns can be given, at least in first approximation, the same analysis as (intransitive) verbs, i.e., the meaning of, say, *cat* can be thought of as a characteristic function that selects those entities that are cats out of the universe of discourse (or, equivalently, we can say that *cat* identifies a class/set across situations). But what about things like *no cat* or *every cat*, which are the typical constituents one finds in, e.g., subject position and the like? What does *no cat* denote? And, even worse, what do *no* or *every* or *some* denote? Our program is to assign a denotation to lexical entries and then to define in terms of it truth conditions for sentences. So we must find suitable denotations of Ds and DPs.

To address questions of this sort, we apply a heuristic that goes naturally with our general setup: whenever the denotation of an expression is not directly accessible to your intuition, look at what that expression contributes to the truth conditions of the sentences it occurs in (the epistemological primacy of sentences, again). So, consider for example:

(32) No boy smokes.

We know/assume/conjecture that *boy* and *smoke* denote characteristic functions and that sentences contribute truth values (i.e., they are true or false, as the case may be, in different situations). We may think of *no* as a function, too. As is evident from (32), such a function combines first with a characteristic function/set (corresponding to the noun); then the result combines with a second characteristic function (corresponding to the verb) to yield a truth value. Schematically, here is what we have:

(33) $\mathbf{no}(\mathbf{boy}_s)\ (\mathbf{smoke}_s) = 1$ or 0

Now we can look at our intuitions. When is (32) true? The answer is pretty clear. When among the boys, nobody smokes. Or, equivalently, when the class of boys (i.e., the extension of $\mathbf{boy}_s$) has no member in common with the smokers (i.e., the extension of $\mathbf{smoke}_s$), (32) is true. In set talk, the intersection between the boys and the smokers must be empty:

(34) $\text{no}(\mathbf{boy}_s)\ (\mathbf{smoke}_s) = 1 \text{ iff } BOY_s \cap SMOKE_s = \varnothing$

(where BOY_s, $SMOKE_s$ are the extensions corresponding to $\mathbf{boy}_s$, $\mathbf{smoke}_s$, respectively) This is perfectly general. Replace boy/smokes with any other noun/verb. The contribution of *no* stays constant: no(N) (V) is true just in case no member of the extension of N is in V. We thus discover that *no* has a perfectly sensible (if abstract) denotation: a function that encodes a relation between sets. Our contention here is that speakers behave as if they had such a function in mind (or something similar to it) in using *no*.

Encyclopedia of Language & Linguistics (Second Edition), 2006, Pages 564-579

The next step is to see that *all* determiners express relations among sets (characteristic functions), just like *no* does. Here are a few examples, along with some comments.

(35a) *Some*
(35a.i) Example: some boy smokes
(35a.ii) Truth conditions: **some**(**boy**$_s$) (**smoke**$_s$) = 1 iff $BOY_s \cap SMOKE_s \neq \emptyset$
(35a.iii) Comment: *some* is the contrary of *no*; *some boy smokes* is true just in case you can find someone among the boys who is also among the smokers; i.e., the intersection between the class of boys and the class of smokers must be non empty. The indefinite article *a* can be analyzed along similar lines.
(35b) *Every*
(35b.i) Example: every boy smokes
(35b.ii) Truth conditions: **every**(**boy**$_s$) (**smoke**$_s$) = 1 iff $BOY_s \subseteq SMOKE_s$
(35b.iii) Comment: *every* expresses the subset relation: *every boy smokes* is true just in case all the members of the class of boys also belongs to the class of smokers
(35c) *Most*
(35c.i) Example: Most boys smoke
(35c.ii) **most**(**boy**$_s$) (**smoke**$_s$) = 1 iff the number of member of $BOY_s \cap SMOKE_s$ is bigger than half the number of members of BOY_s.
(35c.iii) Comment: *most* involves actual counting. *Most boys smoke* is true just in case the number of boys who smoke (i.e., the intersection of the boys with the smokers) is greater than half the number of boys (i.e., more than half of the boys are smokers).
(35d) *The*
(35d.i) Example: The blond boy smokes.
(35d.ii) Truth conditions: **the** (**blond boy**$_s$) (**smoke**$_s$) is defined only if there is exactly one blond boy in s. Whenever defined, **the** (**boy**$_s$) (**smoke**$_s$) = **every** (**boy**$_s$) (**smoke**$_s$).
(35d.iii) Comment: this reflects the fact that *the blond boy smokes* is only interpretable in situations in which the universe of discourse contains just one blond boy. If there is more than one blond boy or if there is no blond boy, we wouldn't really know what to make of the sentence. So *the* is a presuppositional determiner; it presupposes the existence and uniqueness of the common noun extension. (This analysis of *the* goes back to Frege.)

In spite of the sketchiness of these remarks (that neglect important details of particular determiners), it should be evident that the present line of analysis is potentially quite effective. A class of words and phrases important and tendentially stable across many languages falls into place: determiners ultimately express natural relations between sets (the set associated with the common noun and the set associated with the verb phrase). Our denotational perspective seems to meet rather well the challenge that seemingly denotationless items pose. It is useful to see what becomes of our rule of predication (viz. (31) above). Evidently such a rule needs to be split into two (main) subcases, depending on whether the subject is a simple N (a proper name) or a complex DP. Here is an exemplification of the two cases:

(36a) Mary smokes.
(36b) No boy smokes.

In case (36a), we have semantically two pieces: an individual (whomever *Mary* denotes) and a characteristic function (**smoke**$_s$); so the latter applies to the former. In case (36b) the two pieces are: a complex function (namely **no** (**boy**$_s$)) that looks for a characteristic function to yield a truth value, and, as before, the characteristic function **smoke**$_s$; in this case the former applies to the latter. In either case, the end result is a truth value. So our predication rule becomes:

(37a) $\| [_{VP}\ N\ V\ ] \|^s = \|V\|^s (\|N\|^s)$
(37b) $\| [_{VP}\ DP\ V\ ] \|^s = \|DP\|^s (\|V\|^s)$

This suggests that the core rule of semantic composition is functional application. Consider for example an ungrammatical sentence of the form:

(38) * $[_{VP}$ boy smokes]

Such a sentence, as things stand, would be generated by our (rudimentary) syntax. However, when we try to interpret it, we find two characteristic functions of individuals, neither of which can apply to the other. Hence, the sentence is uninterpretable, which explains its ungrammaticality. There are languages like, for example, Russian or Hindi where singular common nouns without a determiner can occur in subject position:

(39a) Russian: mal'cik kurit boy smokes 'the boy smokes'_
(39b) Hindi: kamre meN cuuha ghuum rahaa hai (from Dayal 2004) room in mouse moving is 'a mouse is moving in the room'

Notice that (39a) is the verbatim translation of (38) and is grammatical in Russian. The line we are taking suggests that in such languages it must be possible to turn some covert forms of common nouns into

Encyclopedia of Language & Linguistics (Second Edition), 2006, Pages 564-579

argumental DPs, i.e., things that can semantically combine with predicates; for example it is conceivable that in a language without articles, like Russian, the semantic functions associated with the articles can be applied covertly (as part of the interpretive procedure), so as to rescue the semantic mismatch that would otherwise ensue. This may, in turn, involve the presence of a phonologically null determiner (for alternative developments of this line of analysis, as well as details concerning the available interpretations, see, e.g., Chierchia (1998), Longobardi (2001), and Dayal (2004)).

The picture that emerges is the following. The basic mode of syntactic composition is *merge*, or some analogously simple operation that puts together two constituents (subject to parametrization pertaining to, e.g., word order, case, etc.). The basic mode of semantic composition is *apply*: constituents are compositionally analyzed as functions (of more or less complex semantic type) and arguments (individuals or other functions); so whenever we find a function and an argument of the appropriate sort, we simply apply the former to the latter. If things go wrong at any level, the derivation crashes and the result is ungrammatical. The semantic side of this process has come to be known as 'type driven interpretation,' the main idea being that the semantic categories of functions and arguments drive the interpretation process.

The present approach directly yields a computationally tractable theory of entailment and presupposition. We have defined entailment roughly as follows: a sentence S entails a sentence S' iff whenever S is true, S' is also true. The apparatus we have developed allows us to prove whether a certain entailment holds or not. Let me show, as an illustration, that (40a) entails (40b) but not vice versa.

(40a) Every scientist smokes.
(40b) Every mathematician smokes.

To show this we need to assume that if one is a mathematician, one is a scientist; i.e.,

(41) For every individual a,
(41a) if $\mathbf{mathematician}_s$ (a) = 1, then $\mathbf{scientist}_s$ (a) = 1 or, equivalently:
(41b) $\text{MATHEMATICIAN}_s \subseteq \text{SCIENTIST}_s$

Consider now the semantics of (40a), according to our analysis. It is the following:

(42) $\mathbf{every}(\mathbf{scientist}_s)\ (\ \mathbf{smoke}_s)$

In virtue of (35b), this is tantamount to

(43) $\text{SCIENTIST}_s \subseteq \text{SMOKE}_s$

This being so, every subset of the set of scientists must also be included among the smokers (by elementary set theoretic considerations). Since, in particular, mathematicians are scientists, it follows that

(44) $\text{MATHEMATICIAN}_s \subseteq \text{SMOKE}_s$

But this is just the semantics of (40b). So, if (40a) is true in s, then (40b) must also be true in s. Evidently, this reasoning goes through no matter which situation we are in. Hence, (40a) does entail (40b). On the other hand, it is easy to conceive of a situation in which (44), and hence (40b), hold, but say some economist doesn't smoke; in such a situation, (43) would fail to obtain. Hence, (40b) does not entail (40a).

A fully parallel way of reasoning can be put forth for presuppositions. We said that S presupposes S' iff S' must be taken for granted in every situation in which S is asserted, denied, etc. This can be cashed in as follows. We can say that for S to be true or false (i.e., to have a semantic value that makes it suitable for assertion or denial), S' must be known to be true in the utterance situations by the illocutionary agents, i.e., S can be true or false in s iff S' is true in s. Using this definition (known as the 'semantic' definition of presupposition), we can formally prove (though we will not do so here) that, for example, (45a) presupposes (45b):

(45a) The blond boy smokes.
(45b) There is exactly one blond boy around.

The general point of these examples is the following. Intuitions about entailment and the like are *a priori*; speakers have them just by inspecting the meaning of the relevant sentences. In the present setup, this central fact is captured as follows. Semantics can be viewed as a set of axioms that (a) determines the interpretation of lexical entries and (b) assigns truth conditions to sentences. Such apparatus yields a calculus of entailment (and other semantic relations) that reemerge as theorems of semantics. We have not formalized each single step of the derivation (relying on the readers' patience and understanding of elementary set theory); but such a formalization is, evidently, feasible. We not only thereby gain in clarity. We also obtain a device that constitutes a reasonable (and falsifiable) model of speakers' linguistic abilities. The claim is not that the specific rules we have given are actually implemented in the speakers' mind. The claim is that speakers, to the extent that they can be said to compute entailments must be endowed with computational facilities that bear a structural resemblance to the ones sketched here. This, in turn, paves the way for inspecting the architecture of our linguistic abilities

Encyclopedia of Language & Linguistics (Second Edition), 2006, Pages 564-579

ever more closely. Without excessive optimism and in full awareness of the controversies that permeate the field, this seems to constitute a step in the right direction.

One further remark on the general picture that emerges from the sketch above cannot be avoided. Our approach to meaning is denotational: we assign a denotation to words and morphemes and (in terms of such denotations) truth conditions to sentences. This can be understood in several ways, of which I will present two much simplified extremes. We can take truth condition assignment as a way of exposing the link between language and the world, which is, arguably, the ultimate goal of semantics. Words/morphemes are actually mapped into aspects of the world (e.g., names are mapped into actual individuals); sentences are symbolic structures that code through their fine structure how things may be arranged in the world. However, it is also possible to view things somewhat differently. What really matters, it can be argued, is not the actual mapping between words and aspects of reality and between sentences and the conditions under which they, in fact, are true. What we do is give a form or recipe or potential for actual truth conditions; we merely constrain the form that truth conditions may take. What we get out of this is what really matters: a calculus of semantic relations (entailment, presupposition, etc.). Unlike what happens in, say, pure logic, such a calculus is not a normative characterization of sound reasoning; it is an empirically falsifiable characterization of semantic competence (i.e., of what speakers take to follow from what, when). Under the latter view, truth conditions (or truth condition potentials, or whatever it is that we map sentences on) are a ladder we climb on to understand the working of semantic relations, i.e., relations that concern the information content of linguistic expressions.

It is evident that we are not going to settle these issues here. As a small consolation (but also, if you wish, as evidence of the maturity of the field), I hope to have given the reader reasons to believe that progress is possible even if such foundational issues remain open.

We haven't discussed implicatures and other pragmatically driven intuitions about meaning. To understand the full scope of the present proposal, it is important to do so. This requires extending a bit what we have done so far.

The Semantics/Pragmatics Interface

In the section **Truth and Semantic Competence**, we mentioned implicatures, a broad and varied type of meaning relations. We will elaborate by looking more closely at the oscillation in the meaning of *or*. The purpose is to illustrate how generalized implicatures come about and how this bears on the view of semantics sketched in the preceeding section on **Semantic Modeling**.

The first step is to attempt a semantic analysis of *or*. To this we now turn. Imagine we extend our grammar by introducing coordination and negation along the following lines:

(46a.i) [$_{VP}$ John doesn't smoke]
(46a.ii) [$_{VP}$ NEG VP]
(46b.i) [[$_{VP}$ John smokes] and/or [$_{VP}$ Bill smokes]]
(46b.ii) [VP and/or VP]

The syntax of negation and coordination poses many thorny questions we simply cannot address here. Although for our purposes any number of assumptions concerning syntax might do, let us maintain, again without much justification, that a negative sentence like (46a.i) has the structure in (46a.ii) out of which the observed word order is derived by moving the subject left from the inner VP. Furthermore, we will assume that coordinated sentences, whether disjunctive or conjunctive, such as (46b.i), are obtained through schemas such as (46b.ii). Insofar as semantics is concerned, the introduction of negation, conjunction, disjunction, etc., poses problems similar to that of determiners. The relevant expressions are function words, and it is not obvious how to analyze them in denotational terms. This question, however, can be addressed in much the same way as we have done with the determiners: by looking at what the relevant elements contribute to the truth conditions of the sentences they occur in. For sentential operators, we can draw on a rich logical tradition. In the attempt to characterize the notion of valid inference, logicians have discussed extensively propositional connectives (like *not*, *and*, *or*), and the outcome is an analysis of such elements as truth functions or, equivalently, in terms of 'truth tables.' For example, the contribution of negation to meaning can be spelled out in terms of conditions of the following sort:

(47a) *John doesn't smoke* is true in s iff *John smoke* is false in s
(47b) $\| \text{NEG VP} \|^s = 1$ iff $\| \text{VP} \|^s = 0$
(47c)

VP	NEG VP
1	0
0	1

In (47c) we display in the form of a truth table the semantics given in (47b). Essentially, this says that in uttering a negation like (47a), the speaker intends to convey the falsity of the corresponding positive sentence. By the same token, conjunctions can be

Encyclopedia of Language & Linguistics (Second Edition), 2006, Pages 564-579

analyzed as in (48a), (48b), and (48c), and disjunction as in (49a), (49b), and (49c):

(48a) *John smokes and Bill smokes* is true if both *John smokes* and *Bill smokes* are.

(48b) $\| [VP_1 \text{ and } VP_2] \|^s = 1$ iff $\| VP_1 \|^s = \| VP_2 \|^s = 1$

(48c)	VP_1	VP_2	[VP_1 and VP_2]
(48c.i)	1	1	1
(48c.ii)	1	0	0
(48c.iii)	0	1	0
(48c.iv)	0	0	0

(49a) *John smokes or Bill smokes* is true if either *John smokes* or *Bill smokes* or both are true.

(49b) $\| [VP_1 \text{ and } VP_2] \|^s = 1$ iff either $\| VP_1 \|^s = 1$ or $\| VP_2 \|^s = 1$ or both

(49c)	VP_1	VP_2	[VP_1 or VP_2]
(49c.i)	1	1	1
(49c.ii)	1	0	1
(49c.iii)	0	1	1
(49c.iv)	0	0	0

This is the way in which such connectives are analyzed in classical (Boolean) logic. Such an analysis has proven extremely fruitful for many purposes. Moreover, there is little doubt that the analysis in question is ultimately rooted in the way in which negation, etc., works in natural language; such an analysis indeed captures at least certain natural uses of the relevant words. What is unclear and much debated is whether such an analysis stands a chance as a full-fledged (or nearly so) analysis of the semantics of the corresponding English words. There are plenty of cases where this seems *prima facie* unlikely. This is so much so that many people have concluded that while Boolean operators may be distilled out of language via a process of abstraction, they actually reflect normative principles of good reasoning more than the actual semantics of the corresponding natural language constructions. Of the many ways in which this problem might illustrated, I will choose the debate on the interpretation of *or*.

The interpretation of *or* provided in (49a), (49b), and (49c) is the inclusive one: in case both disjuncts turn out to be true, the disjunction as a whole is considered true. As we saw, this seems adequate for certain uses but not for others. The exclusive *or* can be analyzed along the following lines:

(50) Exclusive *or*

	VP_1	VP_2	[VP_1 or VP_2]
(50.i)	1	1	0
(50.ii)	1	0	1
(50.iii)	0	1	1
(50.iv)	0	0	0

As the readers can verify by comparing (49a), (49b), and (49c) with (50), the two interpretations of *or* differ only in case (i); if both disjuncts are true, the whole disjunction is true on the inclusive interpretation and false on the exclusive one.

So, the thesis that *or* is ambiguous can be given a precise form. There are two homophonous *or*s in English. One is interpreted as in (48a), (48b), and (48c), the other as in (50). Illocutionary agents choose among these options on pragmatic grounds. They go for the interpretation that is best suited to the context. Determining which one that is will involve knowing things like the topic of the conversation (e.g., are we talking about a single job or more than one), the purpose of the conversational exchange, the intentions of the speaker, etc.

We mentioned that Grice proposed an alternative view, however. We are now in position to spell it out more clearly. If you look closely at the two truth tables in (49a), (49b), and (49c) vs. (50), you'll notice that in all the cases in which the exclusive *or* comes out true (namely case (ii) and case (iii)), the inclusive one does, too, i.e., in our terms, [p $or_{exclusive}$ q] entails [p $or_{inclusive}$ q]. The former is, thus, stronger, more informative than the latter in the following precise sense: it rules out more cases. If you get the information that [p $or_{exclusive}$ q] holds, you know that case (ii) or case (iii) may obtain, but case (i) and case (iv) are ruled out. If you know instead that [p $or_{inclusive}$ q] obtains, you know that you might be in case (i), (ii), or (iii); only case (iv) is ruled out. Your degree of uncertainty is higher. So $or_{exclusive}$ is more restrictive; $or_{inclusive}$ is more general (more liberal we said). Things being so, suppose for a moment that *or* in English is unambiguously inclusive (i.e., its interpretation is the most general, less restrictive of the two); this does not rule out at all the possibility that we are in case (ii) or case (iii). The exclusive construal, in other words, might arise as a special case of pragmatic strengthening. It is as if we silently add to, say, (51a) something like (51b).

(51a) John or Mary will be hired.
(51b) (. . . but not both)

The silent addition of (51b) to (51a) might be justified through a reasoning of the following sort:

(52) The speaker said (51a); let us assume she is being cooperative and not hiding on purpose any relevant information. This entails that she has no evidence that both John and Mary have been hired, for otherwise she would have said so. Assuming, moreover, that she is well-informed about the facts, this furthermore entails that she thinks that in fact (51b) holds.

Encyclopedia of Language & Linguistics (Second Edition), 2006, Pages 564-579

So in this view, the base interpretation (viz. (51a)) is enriched through an inferential process that draws on principles of rational conversational exchanges and on factual knowledge about the context. The relation between (51a) and (51b) can thus be analyzed as a case of implicature (cf. on this, e.g., Horn (1989), Levinson (2000), and references therein).

The debate on how the two interpretations of *or* come about is important and shows different ways in which semantics is taken to interact with broader considerations pertaining to communication. Whether the two interpretations of *or* are a matter of ambiguity or arise as an implicature, I want to point out a generalization concerning their distribution, which I think shows something important concerning how language works. I will argue that the cases in which *or* is construed preferentially inclusively are (1) predictable, and (2) determined by structure. Then, I will put forth a hypothesis as to why this is so.

We have seen that a sentence like (16a), repeated here as (57a), is interpreted as in (57b), namely exclusively:

(53a) If I got it right, either John or Mary will be hired.
(53b) If I got it right, either John or Mary but not both will be hired.

Now take the consequent (i.e., the main clause) in the conditional in (53a) and move it to the antecedent, and the interpretation tends to shift:

(54a) If either John or Mary are hired, we'll celebrate.
(54b) If John or Mary or both are hired, we'll celebrate.

So, moving a disjunction from the consequent to the antecedent seems to have a systematic effect on the interpretation of *or*. The same holds for the pair in (55a) and (55b):

(55a) Every student will either take an exam or write a paper.
(55b) Every student who either takes an exam or writes a paper will satisfy the requirements.

In (55a), *or* is within the VP, which corresponds to the second argument of *every*, according to the analysis sketched in the section **Semantic Modeling**. Its preferred interpretation is exclusive. In (55b), *every* is in a relative clause which is part of the subject NP (namely, the first argument of *every* according to the analysis in **Semantic Modeling**). Its preferred interpretation is clearly inclusive.

A further class of contexts that displays a similar effect are negation and negative verbs. Compare (56a) and (56b):

(56a) I believe that either John or Mary will be hired.
(56b) I really doubt that either John or Mary will be hired.

Sentence (56a) is likely to get the interpretation 'I believe that either John or Mary but not both will be hired.' Sentence (56b), on the other hand does not have a parallel reading. It rather means 'I really disbelieve that John and Mary stand a chance.'

The list could go on. But these examples should suffice to instill in the reader the idea that there is a systematic effect of structure on the interpretation of *or*. A doubt might linger, though, as to whether it is really in the nature of structure to have this impact. Take, for example, the pair in (55a) and (55b). Is it the position of disjunction that makes a difference? Or is it rather our knowledge of how classes normally work?

This is a legitimate question. Noveck *et al.* (2002) address it experimentally. They designed a reasoning task, in which logically naïve subjects are asked to judge whether a certain inference is sound or not. For example, subjects were asked to judge whether one can infer (57c) from (57a) and (57b):

(57a) If there is an A, then there is a B or a C.
(57b) There is an A.
therefore:
(57c) There aren't both a B and a C.

Subjects were told that this was about inferences that could be drawn (on the basis of the given premises) concerning letters written on the back of a certain blackboard. What would your answer be? The experimental subjects overwhelmingly accepted the inference in (57a) and (57b). What is interesting is that in terms of classical Boolean logic (which takes *or* to be inclusive) this inference is invalid. It is only valid if *or* in (57a) is interpreted exclusively. At the same time, subjects rejected inferences of the following form:

(58a) If there is an A, then there is a B and a C.
(58b) There is an A.
therefore:
(58c) There is a B or a C.

Again, this seems to make sense only if *or* in (58c) is interpreted exclusively. Things change dramatically if *or* is embedded in the antecedent of a conditional:

(59a) If there is an A or a B, then there is a C.
(59b) There is an A; there is also a B.
therefore:
(59c) There is a C.

Subjects overwhelmingly accepted this inference as valid. But this is only possible if *or* in (59a) is

Encyclopedia of Language & Linguistics (Second Edition), 2006, Pages 564-579

construed inclusively. Our raw intuition thus finds experimental confirmation, one that passes all due controls (the inferences were mixed with others containing other connectives and quantifiers, so that subjects were not conditioned to devise an answering strategy, and the order of presentation was duly varied, etc.). What is interesting is that these experiments only involved meaningless letters A, B, C . . . so scripts, contextual clues, knowledge of the world can hardly be imputed any role in the outcome. If there is a systematic effect on the interpretation of *or*, this must be due to the meaning of conditionals, of disjunction, and to the positioning of the latter. Nothing else is at play.

The reader may wonder how one manages to find out which structures affect the interpretation of *or*. The answer is that such structures were familiar from another phenomenon: the licensing of Negative Polarity Items (NPIs). NPIs are lexical items like *any* or *ever* that seem to require the presence of a negative element:

(60a) * There is any cake left
(60b) There isn't any cake left.

NPIs are acceptable in the contexts that favor the inclusive interpretation of *or* over the exclusive one:

(61a) * If we are in luck, there are any cookies left
(61b) If there are any cookies left, we are in luck.

(62a) * Everyone had any cookies left
(62b) Everyone who had any cookies left shared them.

This correlation is striking, for the two phenomena (the distribution of *any* and of inclusive vs. exclusive *or*) seem to have little in common.

The next question is whether the relevant contexts have some common property. The answer seems to be positive and, surprisingly, points in the direction of a rather abstract, entailment-based property. Positive contexts typically license inferences that go from sets to supersets. For example, (63a) entails (63b) and not vice versa.

(63a) There are Marlboros.
(63b) There are cigarettes.

The set of cigarettes is a superset of the set of Marlboros; so the entailment goes from a set to its supersets. Negation reverses this pattern: (64b) entails (64a) and not vice versa:

(64a) There aren't any Marlboros.
(64b) There aren't any cigarettes.

Now the VP portion of a sentence with *every* (i.e., its second argument) patterns with (63a) and (63b):

(65a) Everyone had Marlboros.
(65b) Everyone had cigarettes.

Sentence (65a) entails sentence (65b) and not vice versa. So does the consequent of a conditional

(66a) If you open the drawer, you'll find Marlboros.
(66b) If you open the drawer, you'll find cigarettes.

But the NP argument of *every* (its first argument) inverts this pattern just like negation, as we saw in the **Semantic Modeling** section:

(67a) Everyone who had Marlboros shared them.
(67b) Everyone who had cigarettes shared them.

Here it is (67b) that entails (67a) and not vice versa. The same applies to the antecedent of conditionals:

(68a) If you smoke Marlboros, you'll be fined.
(68b) If you smoke cigarettes, you'll be fined.

Sentence (68b) entails (68a); on the other hand (68a) could be true without (68b) necessarily being true (in a town in which Marlboros but no other brand is banned).

In conclusion, the contexts that favor the inclusive interpretation of *or* share a semantic property that has to do with entailment patterns: they all license entailments from sets to their subsets. Such a property has come to be seen as the property of being *downward entailing* (where *down* refers to the directionality of the entailment from sets to smaller ones). If this characterization is correct, this means that speakers to the extent that they interpret *or* as shown, must differentiate such contexts, and hence must be able to compute the entailments associated with the relevant structure.

The next question is why *or* tends to be interpreted inclusively in downward entailing structures. I will only hint at what strikes me as a highly plausible answer. As we saw above, in plain unembedded contexts, exclusive *or* is stronger (i.e., asymmetrically entails) than inclusive *or*. The set of cases in which exclusive *or* is true is a subset of the set of cases in which the inclusive one is true. We evidently prefer, everything else being equal, to go for the strongest of two available interpretations. Now, negation and, in fact, all downward entailing structures, as we just saw, reverse this pattern. Under negation, first becomes last; i.e., strongest becomes weakest. In the case of disjunction, the negation of inclusive *or* is stronger (i.e., entails) the negation of exclusive *or*. I'll leave it to the readers to persuade themselves that this is so. Now why is this observation relevant? Suppose we go for the strongest of two alternatives (i.e., we maximize informativeness, everything else being equal); for disjunction, in downward-entailing

Encyclopedia of Language & Linguistics (Second Edition), 2006, Pages 564-579

contexts inclusive *or* is the strongest interpretation; in nondownward-entailing contexts exclusive *or* is the strongest. This explains the observed behavior in terms of a rather simple principle that optimizes information content on the basis of the available expressive resources.

So pragmatic strengthening (via a generalized implicature) correlates harmoniously with the entailment properties of various elements.

Conclusions

We have sketched a view of semantic competence as the implicit knowledge a speaker has of how the information content of various expressions is related. We have proposed to classify the hosts of semantic relations in three major families: entailment-based, presupposition-based, and implicature-based. Given two sentences, speakers can judge whether they entail each other or not, whether they presuppose each other or not, and so on; and they can do so with finite cognitive resources. We have sketched a denotational semantics that accounts for such a competence (i.e., provides a model for it). Our semantics takes the form of a calculus in which entailments, presuppositions, and even (certain) implicatures re-emerge as theorems. Such a model is formal in the sense of being explicit (building on the tradition of logic and model theory). It is, however, also substantive, in that it models a human cognitive capacity (i.e., the ability to semantically relate sentences to each other). We have seen two simple applications of this approach, to the analysis of determiners and connectives. We have also discussed a case of pragmatic enrichment. What we found is that the interpretation of *or* as exclusive or inclusive follows a pattern sensitive to downward entailingness (much like what happens with negative polarity items). If this is so, then entailment patterns are not simply an invention of logicians or linguists. They must be constitutive, in an unconscious form, of the spontaneous knowledge that endows speakers with their linguistic abilities.

See also: Boole and Algebraic Semantics; Compositionality: Semantic Aspects; Implicature; Monotonicity and Generalized Quantifiers; Presupposition; Propositional and Predicate Logic: Linguistic Aspects; Quantifiers: Semantics.

Bibliography

Chierchia G (1998). 'Reference to kinds across languages.' *Natural Language Semantics 6,* 339–445.

Chierchia G & McConnell-Ginet S (2000). *Meaning and grammar* (2nd edn.). Cambridge, Mass: MIT Press.

Crain S & Thornton R (1998). *Investigations in Universal Grammar.* Cambridge, Mass: MIT Press.

Dayal V (2004). Number marking and (in)definiteness in kind-terms. *Linguistics and Philosophy 27,* 393–450.

Grice H P (1989). *Studies in the ways of words.* Cambridge, Mass: Harvard University Press.

Heim I & Kratzer A (1998). *Semantics in generative grammar.* Oxford: Blackwell.

Horn L (1989). *A natural history of negation.* Chicago: University of Chicago Press.

Jennings R E (1994). *The genealogy of disjunction.* Oxford: Oxford University Press.

Kratzer A (1999). Beyond ouch *and* oops. How expressive and descriptive meaning interact. Paper presented at the Cornell Conference on Context Dependency, unpublished manuscript. Amherst Mass.

Levinson S (2000). *Presumptive meanings.* Cambridge, Mass: MIT Press.

Longobardi G (2001). 'How comparative is semantics? A unified parametric theory of bare nouns and proper names.' *Natural Language Semantics 9,* 335–369.

Mates B (1950). 'Synonymity.' In *Meaning and interpretation.* University of California Publications in Philosophy, 25.

Noveck I, Chierchia G, Chevaux F, Guelminger R & Sylvestre E (2002). 'Linguistic–pragmatic factors in interpreting disjunctions.' *Thinking and Reasoning 8,* 297–326.