Meaning, intentionality and communication

(to appear in Semantics: An International Handbook of Natural Language Meaning)

Pierre Jacob
Institut Jean Nicod,
EHESS/ENS-DEC/CNRS,
Ecole Normale Supérieure,
Pavillon Jardin,
29, rue d’Ulm,
75005 Paris,
France
Tel: 33144322696/7
Fax: 33144322699
email: Jacob@ehess.fr

Introduction

1. Intentionality: Brentano’s legacy
2. Early pragmatics: ordinary language philosophy and speech act theory
3. Grice on speaker’s meaning and implicatures
4. The emergence of truth-conditional pragmatics
5. Pragmatics and cognitive science

Introduction

This chapter lies at the interface between ground-level questions in the scientific investigation of human verbal communication and meta-level or metaphysical questions about the nature of meaning. Words and sentences of natural languages have meaning (or semantic properties) and they are used by humans in tasks of verbal communication. Much of
twentieth-century philosophy of mind has been concerned with metaphysical questions raised by the puzzling nature of meaning. For example, what is it about the meaning of the English word “dog” that enables a particular token used in the USA in the twenty-first century to refer (or to latch onto) particular hairy barking creatures that lived in ancient Egypt four thousand years earlier? As I explain in the first section, most if not all philosophy of mind in the twentieth century devoted to the metaphysics of meaning springs from Brentano’s puzzling definition of the medieval word “intentionality” in the late nineteenth century.

Much of the ground-level study of human communication in the twentieth century can be seen as a competition between two models: the “code model” and the “inferential model.” As Sperber and Wilson (1986: 12-13) have argued, decoding and inferential processes differ in important respects. A decoding process starts from a signal and results in recovering a message associated to the signal by an underlying code (i.e., a system of rules or conventions). An inferential process starts from a set of premises and results in a set of conclusions which are warranted by the premises. Neither is a conclusion associated to its premises, nor does a signal warrant the message it conveys. When an addressee understands a speaker’s utterance, how much of the content conveyed by the utterance has been coded into, and can be decoded from, the linguistic meaning of the utterance? How much of the content of an utterance retrieved by an addressee derives from his ability to infer the speaker’s communicative intention? These are the basic ground-level questions that arise in the investigation of human verbal communication.

As I explain in the second section, Austin, one of the leading ordinary language philosophers, emphasized the fact that by uttering sentences of some natural language, a speaker may perform an action (i.e., a speech act). But he espoused a social conventionalist view of speech acts, which later pragmatics rejected in favor of an inferential approach. In the

---

1 See Horwich (2005).
third section, I turn to Grice’s influential inferential model of verbal communication based on his novel concept of speaker’s meaning and his view that communication is a cooperative and rational activity. In the fourth section, I show how many of Grice’s insights have been further developed into a non-Gricean truth-conditional pragmatics. Finally, in section five, I show how the so-called “relevance-theoretic” approach (developed by Sperber and Wilson) has been able to create a bridge between pragmatics and the cognitive sciences. Human communication has been shown to strongly depend on the addressee’s ability to represent a special sort of intention, i.e., a speaker’s communicative intention. Thus, the development of pragmatics takes us from the metaphysical issues about meaning and intentionality inherited from Brentano to the cognitive scientific investigation of the human mindreading capacity to metarepresent others’ mental representations.

1. Intentionality: Brentano’s legacy

Brentano (1874) made a twofold contribution to the philosophy of mind and language: he provided a puzzling definition of intentionality and he put forward the thesis that intentionality is “the mark of the mental.” Intentionality is the power of minds to represent, stand for, or be about, things, properties, events and states of affairs. As the meaning of its Latin root (the verb tendere) indicates, the word “intentionality” denotes the mental tension whereby the human mind aims at what Brentano called “intentional objects.”

The philosophical concept thus expressed by the word “intentionality” should not be confused with the concepts expressed by two of its cognates: “intention” and “intension.” “Intention” refers to a specific type of mental state that, unlike beliefs, judgments, expectations, perceptions, fears, and so on, plays a distinctive role in the etiology, the preparation and the execution of actions. Since Brentano also held the thesis that
intentionality is a property of all and only psychological (or mental) phenomena, intentionality should not be seen as the sole (or distinctive) property of intentions.

“Intensional” and “intensionality,” which mean respectively “non-extensional” and “non-extensionality,” refer to logical features of sentences and utterances, some of which may describe (or report) an individual’s psychological states. To take Quine’s (1948) famous example, “creature with a heart” and “creature with a kidney” have the same extension because they apply to the same things: all creatures with a heart have a kidney and conversely. But the two expressions have different intensions because the property of having a heart and the property of having a kidney are different properties.\(^2\) In general, a linguistic context is non-extensional (or intensional) if, unlike an extensional context, it fails to license both the substitution of coreferential terms salva veritate and the application of the rule of existential generalization.

As Brentano defined it, intentionality is what enables a psychological state or act to represent a state of affairs or be directed upon what he called an “intentional object.” Intentional objects exemplify the property which Brentano called “intentional inexistence” or “immanent objectivity,” by which he presumably meant that it is not necessary for any psychological state (or act) to exemplify intentionality that its target exist in space and time or that the state of affairs mentally represented actually obtain or even be possible. For example, although unicorns do not exist in space and time and although round squares are not possible geometrical objects, thinking about either a unicorn or a round square is thinking about something, not thinking about nothing. To admire Sherlock Holmes or to love Anna Karenina is to admire or to love something, i.e., some intentional object.

Brentano’s puzzling characterization of intentionality gave rise to a gap in twentieth-century philosophical logic between intentional-objects theorists, who claimed that there must

\(^2\) Frege (1892) introduced the related distinction between the sense (intension) and the reference (extension) of a singular term.
be things that do not exist, and their opponents, who denied it. Meinong (1904) went so far as to suppose that for any set of properties, there is some intentional object that instantiates them, but only some, not all, such objects exist. Thus, on his view, it is one thing to say that there are unicorns. It is another thing to say that unicorns exist. By contrast, Russell (1905) designed his theory of definite descriptions in order to account for the possibility of true negative existential beliefs, i.e., of coherently denying that some things exist, without embracing the distinction between being and existence. Following Russell (1905), Quine (1948) and the majority of twentieth century analytic philosophers have denied that there are things that do not exist, on the threefold grounds that (i) existence is not a property (or the English verb “to exist” should be treated as a quantifier, not as a predicate); (ii) the theory of intentional objects incurs intolerably heavy ontological commitments; and (iii) an ontology of non-existent objects can hardly be reconciled with the ontology of the contemporary natural sciences.³

Brentano (1874) did not merely offer a puzzling definition of intentionality, he also held the thesis that intentionality is constitutive of the mental, i.e., that all and only psychological states and processes exhibit intentionality. Brentano’s first thesis that all psychological states possess intentionality has been much discussed in contemporary philosophy of mind, but it will not concern us any further in this chapter. Brentano’s second thesis that only psychological (or mental) phenomena possess intentionality led him to embrace a version of the Cartesian ontological dualist distinction between mental and physical things. Chisholm (1957) offered a linguistic version of Brentano’s second thesis, according to which the intensionality of a linguistic report is a criterion of the intentionality of the reported psychological state.⁴ He further argued that the contents of sentences describing an agent’s putative intentional psychological states cannot be successfully paraphrased into

⁴ See Jacob (2003) for critical discussion of Chisholm’s proposal.
the behaviorist idiom, i.e., into the contents of sentences describing the agent’s bodily movements and behavior.

Quine (1960) accepted Chisholm’s (1957) linguistic version of Brentano’s second thesis which he used as a premise for an influential dilemma: if Chisholm’s linguistic version of Brentano’s thesis is correct, Quine argued, then the intentional vocabulary (whereby we ascribe putative psychological states to conspecifics) cannot be part of the vocabulary of the natural sciences and intentionality cannot be “naturalized.” Quine’s dilemma was that one must therefore choose between embracing a physicalist ontology and intentional realism, i.e., the view that intentionality is a real phenomenon to be investigated (like electricity or heredity) by scientific methods. Unlike Brentano, Quine endorsed physicalism and rejected intentional realism.

Some of the physicalists who accept Quine’s dilemma (e.g., Churchland, 1989) have embraced eliminative materialism and denied purely and simply the reality of beliefs and desires. The short answer to this proposal is, however, that it is difficult to make sense of the belief that there are no beliefs. Others (such as Dennett, 1987) have taken the “instrumentalist” view that although the intentional idiom is a useful stance for predicting a complex physical system’s behavior, it lacks an explanatory value. But the question arises how the intentional idiom could make useful predictions if it fails to describe and explain anything.

As a result of the difficulties inherent to both eliminative materialism and interpretive instrumentalism, several physicalists have chosen to deny Brentano’s second thesis that only non-physical things exhibit intentionality and to challenge Quine’s dilemma according to which intentional realism is not compatible with physicalism, by rejecting Quine’s reliance on

---

5 Quine’s dilemma seems to presuppose acceptance of behaviorism.
6 The short answer does not take into account possible rebuttals from eliminative materialists. For further discussion of eliminative materialism and instrumentalism in the philosophy of mind, see Jacob (1997, 2003) and Rey (1997).
behaviorism. Their project is to “naturalize” intentionality, i.e., to account for the puzzling features of intentionality using only concepts recognizable by natural scientists. Even if one does not subscribe to behaviorism, intentionality is puzzling because it is fully exemplified by the contents of false thoughts and by the meanings of words that fail to apply to anything.\footnote{See section 3 on Grice’s notion of non-natural meaning.}

In recent philosophy of mind, the most influential proposals for naturalizing intentionality have been versions of the so-called “teleosemantic” approach championed by Millikan (1984, 1993, 2004), Dretske (1988, 1995) and others, which is based on the notion of biological function (or purpose). Teleosemantic theories are so-called because they posit an underlying connection between teleology (design or function) and content (or intentionality). According to the teleosemantic framework, a representational device is endowed with a function (or purpose). Something with a function may fail to fulfill its function. If a device has the function to indicate the presence of some property, then it may fail to do so. If and when it does, then it may generate a false representation or represent something that fails to exist.

Brentano’s thesis that only mental phenomena exhibit intentionality seems also open to the obvious challenge that words, phrases and sentences of natural languages, which are not mental things, have meaning, in virtue of which they too can represent, stand for, or be about things, properties, events and states of affairs. Although sentences of natural languages are not mental (or psychological) things, they exhibit intentionality. Most philosophers of mind, such as Grice (1957, 1968), Fodor (1987), Haugeland (1980) and Searle (1983, 1992), have responded to this objection to Brentano’s second thesis by offering a distinction between the underived (or primitive) intentionality of a speaker’s psychological states and the derived intentionality (i.e., the conventional meaning) of the sentences by the utterance of which she expresses her mental states. On their view, sentences of natural languages lack intrinsic
meaning and would be deprived of meaning altogether unless humans used them for the purpose of making their thoughts known to others.⁸

Some philosophers go one step further and posit the existence of an internal “language of thought” (or “mentalese”). They assume that thinking, having a thought or a propositional attitude is to entertain a token of a mental formula that is realized in one’s brain. On this view, like sentences of natural languages, mental sentences possess both syntactic and semantic properties, but, unlike sentences of natural languages, they lack phonological properties. Thus, the semantic properties of a complex mental sentence systematically depend upon the semantic properties of its constituents and their syntactic combination. As Fodor (1975, 1987, 1994) has relentlessly argued, the strongest arguments for the existence of a language of thought are based on such properties of thoughts as their productivity and systematicity, i.e., the fact that there is no upper limit on the complexity of thoughts and the fact that a creature with the ability to form certain thoughts must be equipped with the ability to form some other related thoughts. On this view, the intentionality of an individual’s thoughts and propositional attitudes derives from the semantic properties of symbols in the language of thought.

2. Early pragmatics: ordinary language philosophy and speech act theory

A sentence of a natural language is not an utterance. A sentence — qua linguistic type — is a grammatical sequence of words and morphemes, whose phonological and semantic properties systematically depend upon the phonological and semantic properties of its syntactic constituents. Sentences belong to natural languages. Each utterance consists in producing a particular token of a given sentence whose type belongs to a particular language.⁹

---

⁸ Dennett (1987), however, disagrees because, according to him, the distinction between derived and underived intentionality is just an illusion.
⁹ The word “utterance” can both denote the act of uttering a sentence and the result of this process. Unless I specify it, I usually intend to refer to the latter by the word “utterance.”
Different tokens of one and the same sentence-type all exemplify the same set of phonological, syntactic and semantic properties. Utterances pertain to language use: they are composed of words and sentences of natural languages, but they arise from the particular use (or tokening) of words and sentences of natural languages. So, unlike sentences (qua types), utterances are created by speakers, at particular places and times. Arguably, an utterance inherits all of its grammatical properties from those of the uttered sentence, but in addition it exemplifies further properties not exemplified by the sentence itself. For example, it may be a property of an utterance, not of the uttered sentence itself, that it was addressed to members of the Security Council of the United Nations, in New York City, on February 5, 2003, by the US Secretary of State in a hoarse voice, tape-recorded, broadcast on the French-German channel Arte and perceived by the French president Jacques Chirac, in Paris, on February 6, 2003.

One of the important uses speakers make of words and sentences of natural languages is verbal communication for the benefit of some audience. Not all human communication, however, need be verbal: bodily movements and facial expressions can serve in non-verbal communication. Nor do people use words and sentences of natural languages for the sole purpose of verbal communication: they also use expressions of natural languages for clarifying their thoughts, for reasoning and for making calculations. Just as an utterance inherits the phonological properties of the uttered sentence, it also inherits its semantic properties. But as I just mentioned, an utterance also has acoustic properties not exemplified by the sentence uttered. Similarly, the full meaning of an utterance goes beyond the linguistic meaning of the uttered sentence. The former is underdetermined by the latter, in two distinct aspects: both its representational content and its so-called “illocutionary force” (i.e., whether an utterance is meant as a prediction, a threat or an assertion) are underdetermined by the linguistic meaning of the uttered sentence.
Prior to the cognitive revolution of the late 1950’s, the philosophy of language was deeply divided between two opposing approaches: so-called “ideal language” philosophy and so-called “ordinary language” philosophy. The word “pragmatics,” which derives from the Greek word *praxis* (which means *action* or *practice*), was first introduced by ideal language philosophers as part of a threefold distinction between syntax (or syntactics), semantics and pragmatics. Syntax was defined as the study of internal relations among symbols or signs of a language. Semantics was defined as the study of the relations between signs and their denotations (or designata). Pragmatics was defined as the study of the relations between signs and their users.  

Ideal language philosophers (in the tradition of Frege, Russell, Carnap and Tarski) were mainly interested in the semantic structures of sentences of formal languages especially designed for the purpose of capturing mathematical truths. The syntactic structure of any so-called “well-formed formula” (i.e., sentence) of a formal language is defined by arbitrary rules of formation and derivation. Semantic values are assigned to simple symbols of the language by stipulation and the truth-conditions of sentences can be mechanically determined from the semantic values of its constituents by the syntactic rules of composition. From the overall perspective of ideal language philosophers, such distinctive features of natural languages as their context-dependence appeared as a defect. For example, unlike formal languages, natural languages contain indexical expressions (e.g., “now”, “here” or “I”) whose references can change with the context of utterance.

By contrast, ordinary language philosophers (in the tradition of Wittgenstein, Austin, Strawson and later Searle) were mainly concerned with the distinctive features of the meanings of expressions of natural languages and the variety of their uses in verbal communication. In sharp opposition to ideal language philosophers, ordinary language

---

10 Cf. Morris (1938), Carnap (1942).
philosophers stressed two main points, which paved the way for later work in pragmatics. First, they emphasized the context-dependency of the descriptive content expressed by utterances of sentences of natural languages, to which I shall return in section 4. For example, Austin (1962a: 110-11) denied that a sentence as such could ever be ascribed truth-conditions and a truth-value: “the question of truth and falsehood does not turn only on what a sentence is, nor yet on what it means, but on, speaking very broadly, the circumstances in which it is uttered.” Secondly, in reaction to ideal language philosophers, they criticized what Austin (1962b) called the “descriptive fallacy,” according to which the sole point of using language is to state facts or describe the world.

As indicated by the title of Austin’s (1962b) famous book, *How to Do Things with Words*, they argued that by uttering sentences of some natural language, a speaker performs an action, i.e., a speech act: she performs what Austin called an “illocutionary act” with a particular illocutionary force. For example, by using a sentence of some natural language, a speaker may give an order, ask a question, make a threat, a promise, an entreaty, an apology, an assertion and so on. In fact, Austin (1962b) sketched an entirely new framework for the description and classification of speech acts. As Green (2007) notes, speech acts are not to be confused with acts of speech: “one can perform an act of speech, say by uttering words in order to test a microphone, without performing a speech act.” Conversely, one can issue a warning without saying anything, merely by producing a gesture or even a “minatory facial expression.”

Austin (1962b) identified three distinct levels of action in the performance of a speech act, which he called respectively the “locutionary act,” the “illocutionary act,” and the “perlocutionary act,” which stand to one another in the following hierarchical structure. By uttering some sentence, a speaker performs the locutionary act of saying something that may be reported either by direct or indirect quotation. By performing the locutionary act of saying
something, the speaker also performs some illocutionary act with a particular illocutionary force (e.g., giving an order as opposed to asking a question). Finally, by performing a particular illocutionary act endowed with a specific illocutionary force, the speaker performs a perlocutionary act, whereby she achieves some psychological or behavioral effect upon her audience, such as frightening him, convincing him, or getting from him a response to her question.

Before he considered this threefold distinction within the structure of speech acts, Austin had made an earlier distinction between so-called “constative” and “performative” utterances. A constative utterance is supposed to describe some state of affairs and is true or false according to whether the described state of affairs obtains or not. By contrast, an utterance is performative if, instead of being a (true or false) description of some independent state of affairs, it is supposed to constitute (or create) a state of affairs of its own. Clearly, the utterance of a sentence in either the imperative mood (“Leave this room immediately!”) or the interrogative mood (“What time is it right now?”) is performative in this sense: it does not purport to register any pre-existing state of affairs. Instead, by making such an utterance, the speaker either gives an order or asks a question. In fact, in drawing the distinction between constative and performative utterances, Austin’s main concern was to criticize the descriptive fallacy and to emphasize the fact that many utterances of declarative sentences are performative utterances, not constative utterances.

In particular, Austin was interested in explicit performative utterances (such as “I promise I’ll come,” “I order you to leave,” “I apologize,” or “I hereby open the meeting”), in which the uttered sentence includes a main verb that names (or denotes) the very speech act that the utterance serves to perform. Austin’s attention was especially drawn towards those explicit performatives, whose performance is governed not merely by linguistic rules, but also by social conventions and what Searle (1969: 51) called “institutional facts” (as in “I thereby
pronounce you husband and wife”), i.e., facts that (unlike “brute facts”) presuppose the existence of human institutions. Specific bodily movements count as a move in the game of chess (or tennis), or as an act of betting, or as part of a marriage ceremony only if they conform to some conventions that are themselves part of some social institutions. Similarly, for a performative speech act to count as an act of baptism, an oath or an act of marriage, the utterance must meet some social and institutional constraints, which Austin calls “felicity” conditions. Purported speech acts of baptism, oath or marriage can fail some of their felicity conditions and thereby “misfire” if either the speaker (in e.g., a baptism or a marriage ceremony) lacks the proper authority or the addressee fails to respond with an appropriate uptake in response to e.g., an attempted bet sincerely made by the speaker.\footnote{Austin also considers what he calls “abuses,” as when a speaker makes an insincere promise.}

On the one hand, Austin came to abandon his former distinction between constative and performative utterances when he came to realize that some explicit performatives can be used to make true or false assertions or predictions. Just as one can make an explicit promise or an explicit request by uttering a sentence prefixed by either “I promise” or “I request,” one can also make an assertion or a prediction by uttering a sentence prefixed by either “I assert” or “I predict.” On the other hand, two of his assumptions led Austin to embrace a social conventionalist view of illocutionary acts. First, Austin seems to have assumed that explicit performatives offer a general model for illocutionary acts. Secondly, he seems to have assumed that explicit performatives, whose felicity conditions include the satisfaction of social institutional conventions, are paradigmatic of all explicit performatives. These two assumptions led Austin (1962b: 103) to embrace an institutional or social conventionalist view of illocutionary acts according to which the illocutionary force of a speech act is “conventional in the sense that it could be made explicit by the performative formula.”
Austin’s social conventionalist view of illocutionary force was first challenged by Strawson (1964: 153-54) who pointed out that the assumption that no illocutionary act could be performed unless it conformed to some social convention would be “like supposing that there could not be love affairs which did not proceed on lines laid down in the *Roman de la Rose*.” Instead, Strawson argued, what confers onto a speech act its illocutionary force is that the speaker intends it to be so taken by her audience. On the one hand, the illocutionary force of an utterance may be severely underdetermined by the meaning of the uttered sentence. By uttering “You will leave,” the speaker may make a prediction, a bet or order the addressee to leave. Only the context, not some socially established convention, may help the audience determine the particular illocutionary force of the utterance.

On the other hand, as noted by Searle (1975) and by Bach and Harnish (1979), speech acts may be performed indirectly. For example, by uttering “I would like you to leave,” a speaker directly expresses her desire that her addressee leave. But in so doing, she may also indirectly ask or request her addressee to do so. By uttering the interrogative sentence “Can you pass the salt?” — which is a direct question about her addressee’s ability —, a speaker may indirectly request him to pass the salt. As Recanati (1987: 92-93) argues, when a speaker utters an explicit performativie such as “I order you to leave,” her utterance has the direct illocutionary force of a statement. But it may also have the indirect force of an order. If so, then by making a statement, the speaker indirectly orders her audience to leave. There need be no socially established convention whereby a speaker orders her audience to leave by means of an utterance containing the verb that denotes the act performed by the speaker. As the following sections illustrate, later pragmatics has stressed the contribution of a speaker’s communicative intentions to human communication at the expense of social conventions.

---

12 See next section.
3. Grice on speaker’s meaning and implicatures

Austin and much ordinary language philosophy emphasized the social and pragmatic dimensions of language use. While Austin offered new tools for describing the complex structure of speech acts and classifying them, he embraced a broadly social conventionalist view according to which illocutionary acts are constituted by social and institutional rules. In 1957, Grice published a seminal paper, in which he did three things: first, he drew a contrast between “natural” and “non-natural” meaning. Secondly, he offered a definition of the novel concept of speaker’s meaning. Thirdly, he sketched an entire framework within which human communication is seen as a cooperative and rational activity in which the task of the addressee is to infer the speaker’s meaning on the basis of her utterance, in accordance with a few principles of rational cooperation. In so doing, Grice took a major step towards what Sperber and Wilson (1986) call an “inferential model” of human communication, and away from what they call the “code model.”

As Grice (1957) emphasized, smoke is a natural sign of fire: the former naturally means the latter in the sense that not unless there was a fire would there be any smoke. By contrast, the English word “fire” (or the French word “feu”) non-naturally means fire: if a person erroneously believes that there is a fire when there is none or if she wants to intentionally misleads others into thinking that there is a fire, then she can produce a token of the word “fire.” Thus, the notion of non-natural meaning is Grice’s counterpart of Brentano’s (derived or underived) intentionality.

In this and other subsequent papers, Grice (1957, 1968, 1969) introduced the concept of speaker’s meaning, i.e., of someone meaning something by exhibiting some piece of behavior that can, but need not, be verbal or linguistic. According to Grice, for someone S to mean something by producing some utterance x is for S to intend the utterance of x to produce some effect (or response r) in an audience A by means of A’s recognition of this very
intention. Hence, Grice characterizes speaker’s meaning in terms of communicative intention, where a communicative intention has the peculiar feature of being reflexive in the sense that part of its content is that an audience recognize it.

As I mentioned earlier, Strawson (1964) turned to Grice’s concept of speaker’s meaning as an intentionalist alternative to Austin’s social conventional account of illocutionary acts. In that same paper, Strawson (1964) pointed out that Grice’s complex analysis of speaker’s meaning or communicative intention requires the distinction between three complementary levels of intention. For $S$ to mean something by an utterance $x$ is for $S$ to intend:  

(i) $S$’s utterance of $x$ to produce a response $r$ in audience $A$;
(ii) $A$ to recognize $S$’s intention (i);
(iii) $A$’s recognition of $S$’s intention (i) to function at least as part of $A$’s reason for $A$’s response $r$.

This analysis raises two opposite problems: it is both overly restrictive and insufficiently so. First, as Strawson’s reformulation shows, Grice’s condition (i) corresponds to what, following Austin (1962b), could be called $S$’s intention to perform a perlocutionary act. But as Searle (1969: 46-48) has observed, it is not strictly necessary that $S$’s intention to perform her perlocutionary act be fulfilled for $S$ to successfully communicate with $A$. For example, suppose that $S$ utters: “It is raining,” with the intention (i) to produce in $A$ the belief that it is raining. Suppose that $A$ recognizes $S$’s intention (i), but, for some reason, $A$ mistrusts $S$ and $A$ fails to acquire the belief that it is raining. In that case, $S$ would have failed to convince $A$ (that it is raining); but $S$ would nonetheless have successfully communicated what she meant to $A$. If so, then fulfillment of $S$’s intention (i) is not necessary for successful communication. Nor is the fulfillment of $S$’s intention (iii), which, of course, presupposes the

---

13 By “utterance” here is meant any verbal or non-verbal communicative behavior.
fulfillment of $S$’s intention (i). All that is required for $S$ to communicate what she meant to $A$ is $A$’s recognition of $S$’s intention (ii). What is required is that $S$ has the higher-order intention to inform $A$ of her first-order intention to inform $A$ of something.

Secondly, Strawson (1964) also pointed out that his own reformulation of Grice’s definition of speaker’s meaning seems insufficiently restrictive. As in Sperber and Wilson’s (1986: 30) version of Strawson’s example, suppose that $S$ intends $A$ to believe that she needs his help to fix her hair-drier, but she is reluctant to ask him openly to do so. So $S$ ostensibly offers $A$ evidence that she is both trying and failing to fix her hair-drier, thereby intending $A$ to believe that she needs his help. So as in Grice’s definition, $S$ intends $A$ to recognize her intention to inform him that she needs his help. However, $S$ does not want $A$ to know that she knows that he is watching her. Since $S$ is not openly asking $A$ to help her, she is not thereby communicating with $A$. Although $S$ does have the second-order intention that $A$ recognizes her first-order intention to inform him that she needs his help, she does not want $A$ to recognize her second-order intention. To deal with such a case, Strawson (1964) suggested that the analysis of Grice’s speaker’s meaning include $S$’s third-order intention to have her second-order intention recognized by her audience. But as Schiffer (1972) pointed out, this opens the way to an infinity of higher-order intentions. Instead, Schiffer (1972) argued that for $S$ to have a genuine communicative intention, $S$’s intention to inform $A$ must be mutually known to $S$ and $A$. But as pointed out by Sperber & Wilson (1986: 18-19), people who share mutual knowledge know that they do. So the question arises: how do speaker and hearer know that they do? We shall come back to this issue in the last section.

While Grice (1968) held the view that his concept of speaker’s meaning could be used as a basis for offering a reductive analysis of semantic notions such as sentence meaning or word meaning, most linguists and philosophers have expressed skepticism about this aspect of

---

Grice’s program.\textsuperscript{15} By contrast, many assume that some amended version of Grice’s concept of speaker’s meaning can serve as a basis for an inferential model of human communication. In his famous 1967 William James Lectures,\textsuperscript{16} Grice argued that what enables the hearer to infer the speaker’s meaning on the basis of her utterance is that he is guided by the expectation that all utterances meet what he called the “Cooperative Principle” and a set of nine maxims or norms organized into four main categories which, by reference to Kant, he labeled maxims of Quantity (informativeness), Quality (truthfulness), Relation (relevance) and Manner (clarity). Grice’s normative principles reflect his assumption that communication is basically a rational and cooperative enterprise.

As ordinary language philosophers emphasized, in addition to what is being said by means of an assertion — what makes the assertion true or false —, the very fact that a speaker performs an illocutionary act with the force of an assertion has pragmatic implications. For example, consider Moore’s famous paradox: by uttering the sentence “It is raining but I do not believe it,” the speaker is not expressing a logical contradiction, as there is no logical contradiction between the fact that it is raining and the fact that the speaker fails to believe it. Nonetheless, the utterance is pragmatically paradoxical because by asserting that it is raining, the speaker thereby expresses (or displays) her belief that it is raining, but her utterance explicitly denies that she believes it.

Grice’s (1967/1975) third main contribution to the development of an inferential model of communication was his concept of conversational implicature, which he introduced as “a term of art.”\textsuperscript{17} Suppose that Bill asks Jill whether she is going out and Jill replies: “It’s raining.” For Jill’s utterance about the weather to constitute a response to Bill’s question about whether Jill is going out, additional assumptions are required, such as, for example, that Jill does not like rain (i.e., that if it is raining, then Jill is not going out) which, together with

\textsuperscript{15} See in particular Chomsky (1975, 1980).
\textsuperscript{16} First published as Grice (1975).
\textsuperscript{17} Cf. Grice (1989: 24).
Jill’s response, entails that she is not going out. Alternatively, if Bill were to combine Jill’s utterance with the contrary assumption that she enjoys rain (i.e., that if it is raining, then she is going out), then Bill could infer that Jill is going out. Either way, one such additional assumption is required for Bill to derive a response to his question from Jill’s explicit answer.

Grice’s approach to communication, based on the Cooperative Principle and the maxims, offers a framework for explaining how, from Jill’s utterance, Bill can retrieve an implicit answer to his question by supplying some additional tacit assumption. Bill must be aware that Jill’s utterance is not a direct answer to his question. Assuming that Jill does not violate (or “flout”) the maxim of relevance, she must have intended Bill to supply the assumption that e.g., she does not enjoy rain, and to infer that she is not going out from her explicit utterance. Grice (1967/1975) called the additional assumption and the conclusion “conversational” implicatures. In other words, Grice’s conversational implicatures are required to enable a hearer to reconcile a speaker’s utterance with his assumption that the speaker conforms to the Principle of Cooperation. Grice (1989: 31) insisted that “the presence of a conversational implicature must be capable of being worked out; for even if it can in fact be intuitively grasped, unless the intuition is replaceable by an argument, the implicature (if present at all) will not count as a conversational implicature.”

He further distinguished so-called “generalized” conversational implicatures, which are generated so to speak “by default,” from so-called “particularized” conversational implicatures, whose generation depends on special features of the context of utterance.

Grice’s application of his cooperative framework to human communication and his elaboration of the concept of (generalized) conversational implicature seem to have been motivated by his concern to block certain typical moves made by ordinary language

---

18 Instead, it would count as a so-called “conventional” implicature, i.e., a conventional aspect of meaning that makes no contribution to the truth-conditions of the utterance. For example, utterances of the type “p and q” and “p but q” have the same truth-conditions, but the latter, unlike the former conveys the conventional implicature that although “p” may be evidence for “r”, “q” tells against “r.”
One such move was exemplified by Strawson’s (1952) claim that, unlike the truth-functional conjunction of propositional calculus (entirely defined by its truth-table), the English word “and” can make different contributions to the full meanings of the utterances of pairs of conjoined sentences. For example, an utterance of “John took off his boots and got into bed” is ordinarily interpreted to mean that the event described by the first conjunct took place before the event described by the second. An utterance of “Mary insulted John and he left the room” can be naturally interpreted to mean that the event described by the second conjunct was a result of the event described by the first.

In response, Grice (1967/1975, 1981) argued that, in accordance with the truth table of the logical conjunction of propositional calculus, the utterance of any pair of sentences conjoined by “and” is true if and only if both conjuncts are true and false otherwise. He took the view that the temporal ordering of the sequence of events described by such an utterance need not be part of the semantic content (or truth-conditions) of the utterance. Instead, it should be seen to arise, he thought, as a conversational implicature, which the hearer retrieves through an inferential process in which he is guided by his expectation that the speaker is following the Cooperative Principle and the maxims, e.g., the sub-maxim of orderliness (one of the sub-maxims of the maxim of Manner), according to which there is some reason why the speaker chose to utter the first conjunct before the second.

Grice (1967/1975, 1987) also used his pragmatic apparatus to block another move made by ordinary language philosophers. For example, Malcolm (1949) claimed that Moore had made an incorrect use of the verb “to know” when he famously asserted that he knew that this was a human hand while ostensively demonstrating one of his own hands. Malcolm’s point was that unless some doubt arises or some inquiry is on its way about whether some state of affairs holds, it is inappropriate (incorrect or illegitimate) to apply the verb “to know”

---

19 In particular, Grice (1981) defended Russell’s (1905) theory of definite descriptions against Strawson’s (1950) criticism based on his theory of truth-value gaps.
to describe the cognitive relation between a person and the state of affairs in question. Under the influence of Wittgenstein, several ordinary language philosophers seem to have expressed the similar view that unless it is outright false or at least there is some doubt about whether some thing is red, it is illegitimate or incorrect to say “It looks red to me” (as opposed to “It is red”). In both cases, ordinary language philosophers were making claims about the content of a concept (e.g., knowledge) on the basis of the illegitimacy of the use of the verb “to know.”

In response, Grice (1967/1975) drew attention to the contrast between truth and oddity: what one says may be both true and odd or even misleading. By uttering the sentence “It looks red to me,” the speaker may no doubt indirectly convey the implicature that the thing she is referring to is in fact not red. If so, then if the thing the speaker is referring to turns out to be red, then her utterance may be misleading. But it could still be true. Indeed, as Carston (2002: 103) points out, if the speaker were to utter “It looks red to me and it is in fact red,” the speaker would not contradict herself, as she should if what Grice took to be an implicature was part of the content (and truth-conditions) of her utterance of “It looks red to me.”

4. The emergence of truth-conditional pragmatics

Grice’s seminal work made it clear that verbal communication involves three layers of meaning: (i) the linguistic (or conventional) meaning of the sentence uttered, (ii) the explicit content expressed (i.e., “what is said”) by the utterance, and (iii) the implicit content of the utterance (its conversational implicatures). Work in speech act theory further suggests that each layer of meaning also exhibits two dimensions: a descriptive dimension (e.g., the truth conditions of an utterance) and a pragmatic dimension (e.g., the fact that a speech act is an

---

20 As Grice (1989: 4) put it, “the precept that one should be careful not to confuse meaning and use is perhaps on the way toward being as handy a philosophical vade-mecum as once was the precept that one should be careful to identify them.” For in-depth analysis, see Neale (1992).

21 But see Travis (1991) for an important criticism of Grice’s strategy and a defense of ordinary language philosophy.

22 This is what Recanati (2004a) calls “the basic triad.”
assertion). Restricting itself to the descriptive dimension of meaning, the rest of this section will discuss the emergence of a new truth-conditional pragmatic approach, whose core thesis is that what is said (not just the conversational implicatures of an utterance) depends on the speaker’s meaning. This pragmatic approach to what is said further extends the inferentialist model of communication. As this section will show, it is inconsistent with two deeply entrenched principles in the philosophy of language: literalism and minimalism.\(^{23}\)

Ideal language philosophers thought of indexicality and other context-sensitive phenomena as a defective feature of natural languages. Quine (1960: 193) introduced the concept of an eternal sentence as a sentence devoid of any context-sensitive (e.g., indexical) or ambiguous constituent so that its “truth-value stays fixed through time and from speaker to speaker.” An instance of Quine’s concept of an eternal sentence might be: “Three plus two equals five.” Following Quine, many philosophers (see e.g., Katz, 1981) subsequently accepted literalism, i.e., the view that for any statement made in some natural language using a context-sensitive sentence in a given context, there is some eternal sentence in the same language that can be used to make the same statement in any context.\(^{24}\) Few linguists and philosophers nowadays would subscribe to literalism because they recognize that indexicality is an ineliminable feature of natural languages. However, many subscribe to minimalism.

Grice urged an inferential model of the pragmatic process whereby a hearer infers the conversational implicatures of an utterance from what is said. But by contrast, he embraced the minimalist view that what is said by an utterance departs from the linguistic meaning of the uttered sentence only as is necessary for the utterance to be truth-evaluable.\(^{25}\) If a sentence contains an ambiguous phrase (as in Grice’s own example, “He is in the grip of a vice”), then it must be disambiguated. If it contains an indexical expression, then it cannot be assigned its

\(^{23}\) I borrow the terminology from Recanati (2004a).

\(^{24}\) Recanati (2004a: 84). An alternative version might be Katz’s (1981: 226) “effability” principle: “every thought (or proposition) is expressible by some sentence in every natural language.”

\(^{25}\) As Grice (1989: 25) put it, “what someone has said” is “closely related to the conventional meaning of the words (the sentence) he has uttered.”
proper semantic value except by relying on contextual information. But according to minimalism, appeal to contextual information is always mandated by the presence of some linguistic constituent (e.g., an indexical expression) within the sentence. For example, in order to determine what is said by the utterance of a sentence containing the indexical pronoun “I,” the hearer relies on the rule according to which any token of “I” refers to the speaker who used that token. As Stanley (2000) puts it, “all truth-conditional effects of extra-linguistic context can be traced to logical form” (where by “logical form,” he means a grammatical level at which some of the semantic information about a sentence is represented).

Unlike the reference of a pure indexical like “I,” however, it is widely recognized that knowledge of the reference of a demonstrative (e.g., “he”) depends on representing the speaker’s meaning and cannot be determined automatically by a rule. But so does understanding the semantic value of “here” or “now.” A person may use a token of “here” to refer to a room, a building, a street, a neighborhood, a city, a country, the Earth, the solar system, and so forth. Similarly, a person may use a token of “now” to refer to a millisecond, an hour, a day, a year, a generation, a century, and so forth. One cannot determine the semantic value of a token of either “here” or “now” without representing the speaker’s meaning (or communicative intention).

The basic claim made by advocates of truth-conditional pragmatics is that what is said by an utterance is determined by pragmatic processes, which are not necessarily triggered by some syntactic constituent of the uttered sentence (such as an indexical). By contrast, minimalists, who reject truth-conditional pragmatics, are led to postulate, in the logical form of the sentence uttered, the existence of hidden variables whose semantic values must be contextually determined for the utterance to be truth-evaluable.

Perry (1986) has famously argued that although an utterance of “It’s raining” contains no surface syntactic constituent for a location, nonetheless the hearer automatically supplies
an “unarticulated constituent” for the location so that the speaker has said that it is raining at a particular place (not somewhere or other). On behalf of an “indexicalist” version of minimalism, Stanley (2000) has offered an example of an utterance in which a variable for the location of rain must be bound by a quantifier, as in: “Wherever John goes, it rains.” From the fact that the location of rain can be so bound by a universal quantifier, Stanley concludes that the logical form of a simple utterance of “It’s raining” must contain a free variable that must be contextually assigned a semantic value in order for the utterance to receive a truth-value. If so, then indeed, all truth-conditional effects of extra-linguistic context would be traceable to logical form. On behalf of truth-conditional pragmatics, Recanati (2004a) has offered the following re-analysis of the logical form of the quantified sentence: [For every place \( l \) such that John goes to \( l \), (in \( l \) (it rains))]. On this re-analysis, the simple sentence “It rains” contains no free variable. Only the more complex structure “in \( l \) (it rains)” does. So from the fact that the complex utterance “Wherever John goes, it rains” involves binding of a variable for locations, it does not follow that an utterance of “It rains” contains a free variable, which must be contextually assigned a semantic value.

The rise of truth-conditional pragmatics has been interpreted by some philosophers (e.g., Travis, 1991) as a rehabilitation (if not a revenge) of the contextualist emphasis by ordinary language philosophers on the dependency of an utterance’s truth-conditions upon what Searle (1978, 1983) calls the Background, i.e., a network of practices and unarticulated assumptions and practices. Only against alternative background assumptions could one assign widely divergent truth-conditions to utterances involving one and the same unambiguous verb, as in “John cut the grass” and “John cut the cake.” Plainly, speakers and hearers take it for granted that what counts as cutting grass differs from what counts as cutting a cake. However, some advocates of minimalism (e.g., Cappelen & Lepore, 2005) have argued that if, instead of using a lawn mower, John took out of his pocket a knife and carefully cut each
blade lengthwise, then the speaker’s utterance of “John cut the grass” would count as true (not false).26

Advocates of truth-conditional pragmatics argue that several pragmatic processes based on the representation of the speaker’s meaning are involved in determining what is said by an utterance. I will discuss three of the processes by which the truth-conditional content of what is said results from the modification of the concepts linguistically encoded by the utterance: free enrichment, loosening and transfer.

Free enrichment

Grice (1967/1975, 1981) offered a pragmatic account according to which the temporal or causal ordering between the events described by the utterance of a pair of sentences conjoined by “and” is conveyed as a conversational implicature. But consider Carston’s (1988) example: “Bob gave Mary his key and she opened the door.” Carston (1988) argues that part of what is said by such an utterance is that “she” refers to “Mary” and that Mary opened the door with the key Bob gave her. If this is correct, then surely the fact that Bob gave Mary his key before Mary opened the door is also part of what is said, not conveyed indirectly as an implicature of the utterance. Following Sperber and Wilson (1986: 189), suppose a speaker utters “I have had breakfast,” as an indirect way of declining an offer of food. By minimalist standards, what the speaker said was that she has had breakfast at least once in her life prior to the time of utterance. As noted in the previous section, according to Grice, the hearer must have the ability to infer a conversational implicature from what the speaker said. However, the hearer could simply not conclude that the speaker does not wish any food from the truism that she has had breakfast at least once in her life prior to the time of utterance. Instead, for the hearer to infer that the speaker does not wish to have food in

26 Weaker versions of minimalism offered by Bach (1994), Richard (1990) and Salmon (1986) have argued for the existence of a literal level of truth-conditions such that an utterance of “I have had breakfast” literally expresses the proposition that the speaker has had breakfast once in her life and the mother’s utterance of “You are not going to die” literally expresses the false proposition that the child is never going to die.
response to his question, what the speaker must have said is that she has had breakfast just
prior to the time of utterance. Consider further Bach’s (1994) example of a mother’s utterance
of “You are not going to die” addressed to her child who is crying in response to a minor cut.
On minimalist grounds, the mother’s utterance expresses the false proposition that the child
will not die. But what the child may understand his mother to have said is that he is not going
to die from the minor cut.

Loosening

Cases of free enrichment are instances of strengthening of the concept linguistically
encoded by the meaning of the sentence uttered — as in Carston’s (1988) example of the
strengthening of the concept encoded by “the key” into the concept expressible by “the key
Bob gave to Mary”. However, not all pragmatic processes underlying the generation of what
is said from the linguistic meaning of the sentence uttered are processes of conceptual
strengthening or narrowing. Some are processes of conceptual loosening or broadening. For
example, imagine a speaker’s utterance in a restaurant of “My steak is raw” whereby what she
says is not that her steak is uncooked but rather that it is undercooked. Similarly, to
understand what a speaker says by uttering “The ATM swallowed my credit card,” the hearer
must loosen some of the conditions of application of the concept encoded by the verb “to
swallow” since swallowing is restricted to living things and ATMs are not living things.

Transfer

Strengthening and loosening are both cases of modification of the content of a concept
linguistically encoded by the meaning of a word. Transfer is a process whereby a concept
encoded by the meaning of a word is systematically mapped onto a related but different
concept. Transfer can be illustrated by a couple of famous examples from Nunberg (1979,
1995): “The ham sandwich left without paying” and “I am parked out back.” In the first
element, the property expressed by the predicate “left without paying” is being ascribed to
the person who ordered the ham sandwich, not to the sandwich itself. In the second example, the predicate expresses the property of being the owner of the car parked out back, not the property of being parked out back.

The gist of truth-conditional pragmatics is that speaker’s meaning is involved in determining both the conversational implicatures of an utterance and what is said. As the following example shows, however, it is not always easy to decide, from the standpoint of truth-conditional pragmatics, whether a particular assumption is part of what is said or is a conversational implicature of an utterance. Consider “The picnic was awful. The beer was warm.” For the second sentence to offer a justification (or explanation) of the truth expressed by the first sentence, the assumption must be made that the beer was part of the picnic. According to Carston’s (2002b) account, the assumption that the beer was part of the picnic is a conversational implicature of the utterance (more precisely, an implicated premise). According to Recanati (2004a), the concept linguistically encoded by “the beer” is strengthened and enriched into the concept expressible by “the beer that was part of the picnic” that is part what is said by the utterance.

5. Pragmatics and cognitive science

One of the important issues in the development of pragmatics has been the respective contribution of decoding and inference to verbal communication. This question is made pressing by the central role played by generative linguistics in the cognitive revolution of the 1950’s. According to Chomsky (1965), speakers of a natural language tacitly know a system of grammatical rules that enable them to produce and understand novel sentences which they have never perceived before. The grammar tacitly known by speakers of a natural language must contain recursive rules capable of generating an unbounded set of sentences from a finite lexicon. Such a grammar is a code that systematically maps the phonological and the semantic
properties of sentences belonging to an open-ended set. The task of generative linguistics is to offer a computational account of the grammatical competence that enables a hearer to decode the linguistic meaning of an utterance.

Sperber & Wilson’s (1986) relevance-theoretic approach anchors pragmatic theorizing into cognitive science in at least two respects. First, with the advent of relevance-theoretic pragmatics the investigation of human communication can be pursued within a computational-representational framework consistent with both Chomsky’s approach to the language faculty and Fodor’s language of thought hypothesis. Relevance theory assigns a distinctive role to the process of semantic decoding as part of a general inferential model of verbal communication. According to the relevance-theoretic approach, the output of decoding (made possible by an addressee’s tacit knowledge of the grammar of her language) is treated by the addressee of an utterance as a piece of evidence about the speaker’s meaning (or communicative intention). Furthermore, the speaker’s thought expressed by an utterance is identified to a mental sentence in the language of thought that, unlike the logical form of the utterance, is fully truth-evaluable.

Secondly, relevance theory departs from Grice’s normative approach to communication in that it conceives the pragmatic system as a human cognitive capacity designed for the processing of a special kind of stimuli, i.e., ostensive communicative stimuli. In particular, relevance theory builds a bridge between pragmatics and the so-called “mindreading” capacity, which enables a human being to describe and predict the actions of her conspecifics on the basis of her representations of their psychological states.

Relevance-theoretic pragmatics offers a novel account of speaker’s meaning. As I mentioned in section 3, Schiffer (1972) pointed out that not unless S’s intention to inform A is mutually known to S and A could S’s intention count as a genuine communicative intention. If so, then the question arises: how could S and A know that they mutually know S’s intention
to inform A of something? Sperber & Wilson (1986) have argued that they cannot and that therefore the mutual knowledge requirement should be dropped altogether. Instead, they introduce the concept of manifestness whereby an assumption is said to be manifest to a creature at time $t$ if and only if the creature is capable of mentally representing it and accepting it as true at $t$. They then define an individual’s cognitive environment at $t$ as the set of assumptions that are manifest to him at $t$. As Carston (2002a: 67-68) notes, it is instructive to draw a contrast between the concept of an individual’s cognitive environment and Searle’s (1978, 1983, 1991, 1992) concept of the Background. Whereas Searle (1991) thinks of the Background as a non-intentional condition of intentionality, an individual’s cognitive environment is a set of manifest assumptions which may not be mentally represented, but are nonetheless representable as true.

On the basis of the concept of a manifest assumption, Sperber & Wilson (1986) define the notion of a speaker’s informative intention as an intention to make manifest or more manifest to an audience a set of assumptions $\{I\}$ and the notion of a speaker’s communicative intention as the intention to make it mutually manifest to audience and speaker that the speaker has the above informative intention. Sperber & Wilson (1986: 49) call ostensive a stimulus whereby an agent makes manifest her intention to make something manifest. So a piece of ostensive-inferential communicative behavior is a stimulus produced by an agent whereby she makes mutually manifest to herself and to her audience that she intends, by means of this stimulus, to make manifest (or more manifest) to her audience a set of assumptions $\{I\}$.

Relevance theory is so-called because Sperber & Wilson (1986) accept a Cognitive Principle of relevance according to which human cognition is geared towards the maximization of relevance. Relevance is a property of an input for an individual at a time: it depends on both the set of contextual effects and the cost of processing. Other things being
equal, the greater the set of contextual effects achieved by processing an input, the more relevant the input. Other things being equal, the greater the effort required for processing it, the lower the relevance of the input. Sperber & Wilson (1986) further accept a Communicative Principle of relevance according to which every ostensive stimulus conveys a presumption of its own optimal relevance. Furthermore, an ostensive stimulus is optimally relevant to an audience if and only if it is relevant enough to be worth the audience’s processing effort and it is the most relevant stimulus compatible with the communicator’s abilities and preferences. Combined with the definition of optimal relevance, the Communicative Principle of relevance suggests the following relevance-theoretic comprehension procedure: the hearer should follow a path of least effort in computing cognitive effects, test interpretive hypotheses in order of their accessibility and stop when his expectations of relevance are satisfied.

Most philosophers who contributed to pragmatics have focused on verbal communication and assumed that what a speaker intends to communicate either explicitly or implicitly are propositions. Within the framework of relevance theory, a communicator’s informative intention is characterized instead as an intention to modify the cognitive environment of an audience. Such modifications can vary enormously in strength. Relevance theory draws the following contrast between the explicit and the implicit content of utterances. Decoding of an utterance produces a logical form which is rarely, if ever, truth-evaluable. A so-called explicature is an ostensively communicated assumption which is truth-evaluable and results from the inferential modification of the logical form encoded by an utterance (via processes of decoding, disambiguation, reference assignment and pragmatic enrichment). A conversational implicature is an ostensively communicated assumption which is not an explicature: it can either be an implicated premise or an implicated conclusion. Sperber & Wilson (1986) and Wilson & Sperber (2002) have argued that relevance-theoretic
pragmatics has the resources to account for metaphors, ironical utterances and other tropes, which are overt violations of Grice’s normative maxims.

Much work in cognitive science has been devoted to the mindreading ability that enables humans to describe, predict and explain the actions of others on the basis of their intentions, beliefs and desires. Assuming that an individual’s intentions, beliefs and desires are mental representations, then the mindreading ability to represent others’ intentions, beliefs and desires can be conceptualized, as Sperber (1994) has emphasized, as an ability to form higher-order mental representations of mental representations or meta-representations. Whereas many studies have been devoted to the development and alterations of the ability to represent another’s false belief, human communication involves the ability to represent another’s communicative intention. On relevance-theoretic grounds, a communicative intention is itself a second-order informative intention: it is the intention to make manifest one’s first-order intention. Arguably, for someone to entertain a communicative intention is to intend another to represent one’s own informative intention. If so, then entertaining a communicative intention requires the ability to form a third-order meta-representation. In which case, representing another’s communicative intention requires the ability to form a fourth-order meta-representation.

Experiments in developmental psychology have showed that human children under the age of four years of age fail the so-called “false belief” task in which they are requested to predict where an agent will look for an object which she falsely believes to be in one of two boxes when, in fact, unknown to the agent, the object has been moved to the other box.\(^{27}\) By contrast, by the age of two, human children’s language understanding is well on its way. On the relevance-theoretic joint assumptions that language acquisition requires the representation of a speaker’s communicative intention and that the latter requires in turn the ability to form a

fourth-order meta-representation, then it may seem puzzling that children under four should fail the false belief task. But in fact the situation is not as paradoxical as it may seem at first blush. First of all, as several developmental psychologists have argued, passing the false belief task requires, not just the ability to represent another’s false belief about the location of an object, but also the inhibition of one’s own true belief about the location of the object. Secondly, Onishi & Baillargeon (2005) have provided evidence, based on the so-called “dishabituation” paradigm, that 15-month-old babies are able to represent an agent’s false beliefs.

As Sperber (2000: 131) has argued, from the relevance-theoretic standpoint, there is an interesting asymmetry between representing the intentions of agents involved respectively in (communicative) ostensive and non-ostensive intentional behaviors. When one observes a piece of non-ostensive intentional behavior (e.g., hunting), one can ascribe to the agent some intention on the basis of the desirable outcome of his behavior, which can be independently identified (e.g., hit some target). One can ascribe an intention to the agent of a piece of non-ostensive intentional behavior whether or not the behavior is successful (whether the hunter hits or misses his target). However, the desirable outcome of a piece of ostensive communicative behavior is that the addressee recognizes the agent’s communicative intention. The desirable effect of a piece of ostensive behavior cannot be achieved until the addressee recognizes the agent’s intention to achieve that effect. In the case of ostensive intentional behavior, the very recognition of the intended effect secures the achievement of the effect itself. Given the complexity of the computational resources required for representing a communicative intention (i.e., the ability to form fourth-order meta-representations), Sperber (2000), Origgi & Sperber (2000) and Sperber & Wilson (2002) have

---

29 Meltzoff (1995) has showed that 18-month-old children are able to represent the intention of the agent of non-ostensive intentional behavior when the agent fails to achieve his intended outcome.
argued that it is likely that “relevance guided inferential comprehension of ostensive stimuli is a human adaptation, a sub-module of the human mindreading ability.”

References


---

I am grateful to Neftali Villanueva Fernández for comments on this chapter.


