

Semantics/Pragmatics:
(Purposes and Cross-Purposes)^{1,2,3}

' 1. About Purposes

¹ Much of the first half of this chapter has been revised and adapted, with the kind permission of Hegeler Institute, from "Purposes and Cross-purposes: On the Evolution of Language and Languages," The Monist 84.3 (July 2001) 392-416.

² I use "Purposes and Cross-purposes" as a subtitle here because in a footnote to chapter 10 in Varieties of Meaning (2004) I promised, I now think inadvisably, to present the material on understanding implicatures and non-literal usages that appears here (' 4) in a chapter named that.

³ I am grateful to Carston for a helpful reading of an earlier draft of this "second edition."

Products emerging from histories of natural selection under consistent selection pressures can be described as having "natural purposes," purposes to perform the functions or to produce the effects for which they were selected. A number of different kinds of selection mechanisms generate natural purposes, and selection operates on a number of different kinds of replicators. Besides genes and behaviors for which genes are directly responsible there are behaviors established by conditioning, behaviors learned by consciously intended trial and error, action alternatives selected through trial and error in thought ("Popperian selection"), and a variety of cultural replicators of the kind Dawkins called "memes"⁴ --items that are replicated by people copying one another's behaviors and products. Purposes derived from natural selection are often contrasted with human purposes. I have argued that this is a mistake, that these purposes are not of a fundamentally different kind.⁵ Even those human purposes that are determined by explicit goals and conscious intentions are merely sophisticated forms of natural purposes. Representations of goals, explicit intentions, and so forth, have fulfillment of their represented ends as "derived proper functions,"⁶ functions that are derived from the complex biological functions of the cognitive and conative mechanisms that produced them, which mechanisms were designed by natural selection to vary their

⁴ Dawkins 1976/89, Ch. 11..

⁵ Millikan 1984, Millikan 2000 ch. 1, Chapter 5 this volume

⁶ Millikan 1984 Ch.2, Millikan 2002

productions to accord with relevant variations in inner and outer conditions and circumstances.

I will be concerned here with interactions between two broad kinds of purposes. First are the purposes of human behaviors. These form a variety, ranging from directly genetically determined purposes such as the purpose of the eye blink reflex, through purposes that result, for example, from subtle or less subtle conditioning or that govern the fine-tuned intermediate portions of automatized skills, to purposes defined by explicit intentions and represented goals. Second are the purposes of conventional public language forms such as words, idioms, syntactic forms and so forth. Public language forms are "conventional" in the sense described in the first chapter of this book. They are copied or "reproduced" in a language community owing to precedent determined by historical accident, rather than owing to properties that make them more intrinsically serviceable than other forms might have been. That is, they are relatively arbitrary. Conventional linguistic forms are memes of a particular sort, memes that are selected for because they are serving coordinating functions.⁷ Their proliferation or "survival" depends on a sufficiency of occasions on which they promote at once both a speaker's and a hearer's interests. Their coordinating functions, their memetic purposes, determine their conventional meanings of both the first and the second kinds discussed in chapter 3 --"stabilizing functions" and "semantic mappings." Given this, my topic, broadly conceived, is the relation of semantics to pragmatics and, with post-Gricean pragmatics

⁷ Chapters 1-5, this volume.

in mind, the question what kinds of speaker purposes and hearer understandings are involved during normal semantic, and during normal pragmatic, linguistic processing.

I'll start (' 2) with a brief review of the way in which conventional linguistic forms acquire memetic functions and a preliminary discussion of the distinction and relation between public language semantics and speaker meanings. Then (' 3) I will question the sharpness of the distinction between these two kinds of meaning and introduce a role that context may play in thoroughly conventional forms of communication. Last (' 4) I will explore the ways in which mature hearers and also children arrive at understandings both of public language meaning and of speaker meanings keeping the contemporary question in the foreground whether hearers have to possess some sort of "theory of mind" in order to do this.

' 2. Linguistic Purposes

For a mechanism of selection to contribute to the emergence of items with natural purposes, the method of replication it depends on has to be quite accurate. Copies of copies must continue to be like the originals in relevant respects with very few exceptions. Dawkins (1976) calls this "fidelity." In the case of human language, at least two kinds of dedicated filters, constraining variety and inhibiting drift in replication, seem to exist. First, there is evidence that the human auditory systems are specifically designed to accomplish efficient mastery of the prosodic and phonological structures of a language. Phonological structures are particulate and compositional and they determine what will be heard, in the given language, as the same linguistic pattern repeated and what as a different pattern. They define the basic same-different scheme for a spoken

language, what counts as correct reproduction of an element such as a word or a sentence, thus helping to ensure fidelity in linguistic replication. And they enable a learner to know in advance what aspects of the speech signals produced by herself or others will be the instrumental aspects, aspects that matter to sameness and difference of meaning. Second, universal grammar may be a mechanism helping to effect faithful reproduction of linguistic forms. To posit universal grammar is to posit a filter controlling the aspects of the language a child hears that it will reproduce, or, in practice the same thing, determining what aspects will be perceived as functionally significant aspects.

Prior agreement on the kind of materials that are to be used in communication and the aspects of these materials that are to be significant produces a genuinely new kind of faithful replicator, ready for selection, but it does not, of course, determine what the replicated linguistic forms shall mean, what their functions shall be. Their functions are determined by whatever accounts for their continued proliferation and this, with few exceptions, is some service that they are performing for cooperative speaker-hearer pairs. Correlatively, the replicators that form the basis of a language in use are not utterances only. They are two-part patterns, an utterance followed by a conventional cooperative hearer response. The whole pattern must be reproduced for the purpose of the language form itself, as opposed to the purposes merely of the current speaker and/or hearer, to be accomplished. This theme is spelled out at some length in my (1984) chapters 3-4 and 10-14, so I will just give a couple of easy examples here.

Consider, for any language, the syntactic form that gets labeled "indicative." This form may have a number of alternative functions, just as one's tongue has alternative

functions, being designed, for example, to help both with mastication and with speech production. But the form will not be called "indicative" unless one of its central functions is this. It effects production of a true belief having whatever propositional content the various other aspects of the sentence are designed to impart.⁸ This effect is often of interest both to speakers and to hearers. Production of false hearer beliefs may occasionally interest speakers, but rarely serves the interests of hearers. A hearer unable to interpret the indicative sentences she heard so as sometimes to extract genuine information from them would soon cease to form beliefs on their basis. She might first try out other interpretations of the form, and different interpretations of other linguistic elements used with it, but eventually she would have to give up on it altogether.

And if hearers ceased using indicative sentences as guides in forming beliefs, speakers would stop trying to use them to impart beliefs. Production of true beliefs, then, is a linguistic function, a purpose, of the indicative form itself, whether or not a particular speaker and/or hearer have as their own purpose to use it that way on a given occasion. Similarly, to instigate actions that accord with their propositional contents is a linguistic function of imperative mood sentences. If it were not sometimes in the interest of hearers to comply with directives Cadvice, instructions, directions, friendly requests, sanctioned directives, and so forthC imperative syntactic forms would either become obsolete or change their functions. Giving a different sort of example, to help to produce

⁸ This function is, of course, highly relational. See my (1984), chapter 2 and chapter 4, appendix A.

representations of elephants in hearers' minds is a purpose of the word "elephant," further aspects of its memetic function depending on linguistic context.

Besides their conventional or "stabilizing" functions, tokens of language forms in use also have functions/purposes that are derived from the purposes of the individual speakers who use them. A tool designed by a person for a purpose has that purpose as its own purpose. It has a "derived" function, derived from the purpose of the person who made it.⁹ Similarly, a token of a linguistic form that has been produced by a speaker for a purpose has a derived function, derived from the purpose of the speaker in speaking. Thus it has two kinds of purposes, a memetic purpose and the speaker's purpose, and these may coincide, or they may differ or even conflict.¹⁰ Indeed, following tradition, it seems clear that failure of these two kinds of purpose to coincide is a common occurrence. In the case of metaphorical use, sarcastic use, or the use of other figures of speech, there is actually a conflict between stabilizing function and speaker purpose. The "literal meaning" is not what the speaker means. And in uses of the kind that Grice labeled "conversational implicatures," what the speaker means either conflicts with the stabilizing function of the form or has some additional purpose beyond.

Once a specific language is in place, using its conventional forms in the

⁹ See my (1984) chapter 2.

¹⁰ It is very common for a behavior or product of behavior to have more than one source of function or purpose, and not uncommon for these purposes to conflict. Chapter 1 of my (2004) contains an extensive discussion of this point.

conventional way is what evolutionary biologists call an "evolutionarily stable solution" (ESS) for many of the coordination problems of communication. No single participant can gain by unilaterally changing the basic strategies or rules by which she communicates. A settled language community is not easily invaded by incompatible policies of language use. Still, because the forms of a specific language are arbitrary within very broad limits, they remain subject to slow copying drift. Slow changes do occur both in meanings and in phonological and syntactic structures. More germane to my intended topic, there are also changes that come about through purposive innovation. Speakers and hearers may cooperate to improvise new uses, and these uses may be copied and may later become conventional.

' 3. The Blurry line between Semantics and Pragmatics

If stabilizing function and speaker purpose for a linguistic form conflict often enough, the result will be either a change in function, or the addition of a new function. If the speaker's purpose is cooperative, the hearer understands the purpose, and this speaker-hearer pattern is reproduced often enough, the new use will become conventional. It has become conventional when it starts being copied from one speaker-hearer pair to another rather than reinvented each time it appears, so that it wouldn't occur, or occur nearly so often, were it not for this precedent. Conventionality is not a matter of linguistic rules having suddenly materialized in some obscure quarter or medium. Nowhere are there any rules written or any "codes" laid down. Conventionality is merely a matter of the likelihood of persons saying and understanding linguistic forms in certain ways had they not experienced their being used in those ways before (Chapter

one, this volume). Conventionality, then, is a matter of degree, and if the distinction between semantics and pragmatics is interpreted as the distinction between what is conventional in language use and understanding and what is not,¹¹ it follows that the semantics/pragmatics distinction is also a matter of degree. I propose now to expand this theme into several dimensions. Implications for the theory of language understanding will be discussed in ' 4 below.

Consider, first, the process by which a fresh metaphor may become trite and eventually quite dead. For example, the metaphor "dead metaphor" is fairly dead; were it not, "fairly dead" would come through as an oxymoron. In the beginning, some people may copy the use of an easy metaphor from friends, while others arrive at it spontaneously. Then there may be a long period, perhaps even encompassing the whole rest of its history, when the form is often used both because its new meaning is understood and copied and because its old meaning continues to be understood. Examples of this usage will belong to two lineages --will exemplify two different least types-- at once (Chapter 3, this volume). The form is copied with both the old and the new uses in mind. Both are resonating, as it were (or perhaps as it is) in the connectionist net. That is what gives many usages their strong color long after they have ceased to be merely metaphorical. Establishment of a new meaning as an independent literal meaning might be defined as occurring when the new meaning would continue to

¹¹ I recommended this position in my (2004) chapter 11.

proliferate even if the old meaning should die out. But although this may be a theoretically clarifying idea, it is hard to see how one might, in practice, gain evidence for a usage having this status, unless, of course, the old usage actually does die out (as often happens).

Besides metaphors and other figures of speech that may die and become literal quite gradually, there are other factors that have a broadly blurring effect on the semantics/pragmatics distinction. I have mentioned that certain parameters of the reproduction of linguistic memetic patterns are constrained by the conventions of phonological structure and syntax, perhaps even aided by innate tendencies. But certain other important aspects of linguistic reproduction seem to be left unfettered, allowing different speakers and hearers to interpret and to copy what they hear somewhat differently. At least these three aspects of speech reproduction seem to allow some slippage:

- C The length of the speech stream segment that is copied
- C The exact effects owing to which a segment is copied, thus what its function or meaning is understood to be
- C The degree of embedding in extra-linguistic context that may be copied along with the segment

I will argue that none of these parameters has a determinate setting for the proliferation of language forms. Thus there may be variation in how different persons process the same forms, indeed, in how the same person processes them on different occasions. This adds to the reason why a clear distinction often cannot be drawn

between an expression's having one sense or several senses, between its being used in more than one literal sense or only one plus various extended or figurative sense, between what has been said and what merely conveyed. The distinction between semantic and pragmatic phenomena rests merely on statistical regularities in psychological processing, and these regularities often are not well defined. Moreover, when different patterns of processing achieve very much the same practical results, there is no pressure for uniformity in processing.

Consider first the length of the reproduced segment of the phonological stream. Just as there is no set length of chromosome that gets copied in sexual reproduction, there is no set length of linguistic expression to be copied. Sometimes a whole sentence form is copied. When answering the phone we say, "This is N" or "N speaking," not "I am N" or "Here is N," but when introducing ourselves in person we say "I am N" or "My name is N," not "This is N" or "N speaking." Other languages do it differently. There are countless conventions of this sort, people copying one another's phrases rather than conveying the same ideas in other equally possible ways. We speak of a flock of sheep or of geese, a herd of reindeer or of cows, a pack of wolves, a pride of lions, a crowd of people, and once it was a bevy of girls and a blush of boys. We say "next year" for the year after this one and "next week" for the week after this one but not "next day" for the day after this one. Instead, we say "tomorrow," indeed "next day" would not even be understood. We speak of the mouth of a river, the mouth of a bottle, the mouth of a balloon, the mouth of a cave, but not the mouth of a house or the mouth of a room, indeed, probably these last would not be understood at all in normal conversation. Using

set phrases rather than composing one's own from smaller linguistic parts, rather, that is, than saying things in "unconventional" ways, is what constitutes speaking "idiomatically." Beginning at 18 months, children learn perhaps ten words a day, steadily, until well into their teens (Bloom 2000). There is no reason to suppose their capacity for learning conventional phrasings is any less dramatic.

People not only speak in chunks, they understand in chunks. Small children often learn phrases first, taking them apart only later. A close look at the average essay written by a freshman who hasn't understood the material well reveals phrase after phrase that has been swallowed whole and returned still semantically, and sometimes even syntactically, unparsed. These phrases have been understood in a fuzzy, holistic way; some semblance of their intended meanings has penetrated, but lacking articulation, hence precision. I have myself only recently penetrated "going haywire," "casting aspersions," and "weighing anchor." A result is that the meaning of a whole easily separates from the compositional meaning that would be derived from its parts, and may then evolve independently, as in the slippage from "weighing anchor" to "Anchors Away," and from "God be with you" to "Goodby."

Conventions of phrasing are sometimes called "conventions of language use" as opposed to "conventions of meaning,"¹² but I think this invites confusion. Consider, first, cases in which there has been no slippage so that although the phrase is copied whole, the function performed, the meaning, is the same as the meaning that would be

¹² Compare (Searle 1975) and (Morgan 1978).

compositionally constructed from its parts. Thus, given the prior conventions governing the individual components of the phrase "This in N," there is nothing arbitrary about what this phrase means when used in answering the phone. Meanings of such phrases are readily derivable from living principles governing their parts. Still, they have something in common with half dead metaphors, for their meanings are derived alternatively or, often, simultaneously in two different ways, (1) compositionally and (2) from holistic reproduction. It can happen then that these two sources of meaning eventually come apart, the holistic meaning floating free of the compositional structure and acquiring an independent phonology, as in "Goodbye" and "bye."

Where a chunk's meaning has not pulled completely free from its moorings, some people may hear it as built out of its parts while others hear it holistically. Indeed, whether a given person hears through to its parts may be a matter of degree, or may vary, even for the same person, from occasion to occasion. Where a phrase is sometimes understood holistically and sometimes taken apart, so long as these different ways of processing don't produce misunderstandings between speakers and hearers, its lineage will remain unbroken through these differences. It will have only one sense in the public language.

Similarly, when words or longer expressions are frequently used in extended ways, ways with a new and different meaning, as in metaphors, euphemisms or other figures of speech, different people, whether as speakers or hearers, may also hear the expression differently. Indeed, the same person may hear it differently on different occasions. Sometimes it carries strong echos from its original literal usage, sometimes

less strong, and sometimes none. Tokens that are produced with echos in mind are hybrids, produced by the crossing of two gradually separating lineages. They are often reproduced by speakers on two models, copied due to familiarity with the new use but also, in part, due to familiarity with the old. Likewise, tokens that echo for hearers are understood on two models at once. Expressions of this intermediate kind have two conventional or semantic meanings at once. They continue to owe some of their proliferation to, hence to derive some of their natural purpose from, the function their original lineage still serves, but they also belong to a new lineage with a new function.

The same analysis applies to half dead conversational implicatures. Tokens of the much celebrated expression "Can you pass the salt?" no longer function merely as conversational implicatures. They also have a literal imperative meaning. They belong to two lineages at once, a wide-ranging syntactic family of English expressions that proliferates owing to its interrogative function, and a restricted family that is an idiomatic form used for requests. These tokens are literal requests. But they are also literal questions, which is why they have a tentative rather than a demanding feel.

The lesson seems to be that we need not always choose between this being the literal meaning and that being the literal meaning. Occam's razor employed to prohibit proliferation of semantic meanings can be as useless as it is for prohibiting the proliferation of living species. Using a closer analogy, it can be as useless as trying to prohibit the memorization of those addition and multiplication "facts" that we use from memory rather than calculating again each time they are needed. On the other hand, complete separations may eventually occur between lineages that had earlier reproduced

tokens jointly. A classic case is the "The dog went to the bathroom on the living room rug" in which "went to the bathroom" floats entirely free of its original moorings (Morgan 1978).

Now consider the question, for the sake of what effect is a linguistic segment replicated? That is the same as asking what its linguistic function, its meaning, is. What is to prevent different people from noticing somewhat different effects, or from generalizing in somewhat different ways to new cases? How broad or narrow, abstract or concrete is a word's meaning? How many distinct senses does it have? None of these questions appears to have answers that are bounded in as sharp a way as are standard pronunciation and grammar.

Consider the colors of things, for example, consider red hair. If a dress were that color, it would never be called red. Our cat that everyone calls orange exactly matches our wooden kitchen floor that everyone calls brown. Do "red" and "orange" each have several senses which are disambiguated according to context? Or are "red hair" and "orange cat" understood as chunks (like "mouth of a river")? Or does "red x" mean red for an x, as "large mouse" means large for a mouse (Wheeler 1972)? (Is it true that what is orange for a cat is not orange for a wooden floor?) Does "long" mean the same thing or different things when applied to space and to time, that is, to the first, second or third versus the fourth dimension? How many different semantic senses does "clear" have when I clear the table, clear the ground, when the water is clear, the coast is clear, the bell is clear and the message is clear? Is the term "lineage" used literally when applied to chains of copied words as in Chapter 3 above, or is this an extended use? Suppose that

you have "gone around" a squirrel in the sense that you have circled from north to east to south to west of a tree it is climbing, but suppose it kept ducking behind so that you never saw its backside.¹³ Is this an extension (an abbreviating) of the meaning of "going around," or does "going around" have two different meanings one of which covers this case, the other not?

Nothing determinate settles questions of this sort, not even within single idiolects. Neither exactly what has been copied from what, nor why it was copied, hence where the copying chains have begun to meander in "new" directions, is anywhere written. Similar remarks may go for the seasoned question whether such words as "some" and "two" have more than one semantic meaning (at least some/only some, at least two/exactly two) or whether the shift from one of these meanings to the other is pragmatic. Similarly, for whether "She had a baby and got married" carries information about temporal order semantically or only pragmatically. There is no reason to suppose that the various lineages concerned here are cleanly formed into separate least types, indeed every reason to suppose they are not. Occam's razor can't prevent the multiplication of actual complexities.

Last, consider whether what gets copied is expressions or expressions-in-contexts. Where the relevant contexts are linguistic, this question may merge with questions about chunking. Compare "Have you had lunch?" with "Have you

¹³ The example is from William James

had chickenpox?" There is a shift here from meaning "Have you had ... today?" to meaning "Have you had ... ever?" Does this shift take place in the semantics, or is it merely a pragmatic phenomenon? Are there two different semantic meanings of "Have you ... ?" which are disambiguated in context? Or does "Have you ...?" mean the same in both cases, referring only to some past stretch of time, which stretch being calculated from pragmatic context? And there is also a chunking possibility. Perhaps what is copied is the use of "Have you...?" coupled with reference to an event of a sort that happens only once or a few times in a person's life, and another use of "Have you...?" coupled with reference to an event of a sort that recurs periodically. In that case it is not the meaning of "Have you...?" that changes. Rather, the meaning is holistic, not strictly built up out of parts.

Similarly, if John and Bill went to Boston then John went to Boston, but if Whitehead and Russell wrote Principia Mathematica it doesn't follow that Whitehead wrote Principia Mathematica (Harnish 1976). Does "A and B" have two meanings, A and B combined versus A and B separately, or does it have one vague meaning that must be specified further pragmatically? The chunking possibility is that subject-plus-verb-ascribing-responsibility is copied whole so as always to mean the subject has sole responsibility. Then the difference between "John and Bill went to Boston" and "Whitehead and Russell wrote Principia" has nothing to do with the meaning of "and." Blind faith in compositional semantics may be rather like belief in "beanbag genetics." Just as the context of other genes in which a particular gene finds itself may radically change its phenotypic expression, similarly, the context of other words in which

a particular word finds itself may change the contribution it makes to the semantics of the whole.

There seems no reason to suppose that one rather than another of the suggested three ways of copying or understanding how "Have you..." works, or how "X and Y did Z" works, is somehow imposed on all speakers of English, nor even that it is determinate for individual speakers exactly which aspects of usage they themselves copy. The same phrase might be copied on different models different times, or copied from two models at once. Nor will it matter to the hearer exactly which aspects of language use are the ones that were copied. If "Have you...?" has two semantic meanings, then the hearer must use context to determine which of the two is being employed. If "Have you...?" is univocal but refers to an indefinite time span, the hearer must use context to determine, not the semantic meaning, but the speaker's meaning. If "Have you...?" is copied in a chunk along with reference to a periodically recurring event, this chunk having a holistic semantic meaning, understanding still won't be merely decoding, for the hearer will have to determine whether the event is to be interpreted as of a recurring kind or not. Does "Have you been in Paris?" ask whether you have just been in Paris or whether you have ever been in Paris? How about "Have you eaten snails?" (Suppose the last is asked by a physician examining your hives.) Imagine English-speaking Martians whose nutrition at home comes only in liquid form but some of whom occasionally visit earth. One asks another, "Have you had lunch?", curious about what that strange experience is like. Speakers sometimes purposefully mislead or make jokes by producing tokens that would have one meaning if derived one way but another if derived another.

Not just position in linguistic context, but position in extra-linguistic context may also be copied by speakers. Saying "hit me!" when playing blackjack to ask the dealer for another card is copied whole. That is, what is copied, hence what has a conventional meaning, is saying-"hit me!"-while-playing-blackjack. Similarly, saying-"break-a-leg!"-to-an-actor-before-a-performance is copied whole by speakers in order, conventionally, to wish the actor luck. These phrases-in-context are reproduced items, replicators, each with a special memetic purpose or meaning. Contrast the case of "This is N" used to introduce oneself on the phone. Its use in a particular context is also copied, but saying "This is N" does not have a separate meaning in that context. It does not mean what it means because it is said on the phone.

Perhaps the simplest example of a conventional use of context to fill out meaning is putting signs or labels on things: "First Prize" or "Weight limit 20 tons." Demonstratives and indexicals are obvious examples of language forms with meanings that depend on context. They also illustrate well how the conventional shades into the pragmatic. The relations that demonstratives bear to their referents in the external environment or in the discourse environment seem often to be conventional but may also be improvised. Use of a pointing gesture along with a demonstrative may typically be copied, different cultures exhibiting different forms of pointing and perhaps different forms for different occasions. There is use of the whole hand, the first finger, the middle finger, the head, and even the lips. Besides pointing are many other ways to establish joint attention as one refers with demonstratives, ways of making salient or using the natural salience of objects in the external or discourse environment. Some of these ways are improvised, perhaps, and

others copied, one person improvising where another one copies. What is at first merely natural may acquire a memetic function, slowly entering the semantic dimension, perhaps also becoming stereotyped or "conventionalized." To what degree a certain method of demonstrative reference is part of the language and to what degree it is improvised from natural materials hence merely pragmatic will be largely a matter of statistics on individual psychological processing, hence not easily open to a priori inspection or argument. Nor should we expect to find this distinction clearly within individual idiolects, for there is no reason to suppose that the same person always processes the same form the same way nor, given contemporary connectionist views, that these different ways of processing always remain distinct within individual minds.

Like the referents of demonstratives, the domains of quantifiers are also understood relative to context, and again it is an intrinsically blurry question how much the use of context is improvised and how much it follows conventional forms. The domain of a quantify is often whatever domain is currently the object of joint attention, either naturally so, or because the speaker has done or said something that conventionally makes it so. With both demonstratives and quantifiers, the speaker herself or certain facts about her, along with the natural symptoms she shows of attention, are a major part of the context. What parts of a speaker's repertoire of methods for establishing joint attention are natural, what parts improvised from natural materials, what parts imitated, and what parts stereotyped and conventionalized, cannot be a very determinate matter. Nor can it be determinate exactly how large the copied language-plus-environment chunks are, hence exactly which chunks have semantic meaning. Once

again, it cannot be determinate where the semantics ends and the pragmatics begins. When misunderstandings occur, often there is nothing to determine whether the speaker has said or conventionally indicated something wrong, or merely meant something her hearer failed to interpret.

Nor does this fuzziness between semantics and pragmatics raise any special problems for the theory of language interpretation. It has been common to presume a wide divide between the way semantically conveyed information is interpreted or "processed" by a hearer and the way pragmatically conveyed information is processed. Semantically conveyed information is said to be merely "decoded" while pragmatically conveyed information must be "recovered" by inference about the speaker's intentions. In '4 below I will explore the role that consideration of speaker intentions plays or does not play in language understanding. But it is important to observe first that semantically conveyed information is never simply "decoded."

The deep reason for this is that, on the present analysis, there can be no strictly dedicated forms in a language. Language is just a raggedy collection of reproduced speaker-hearer patterns having various origins and independent histories. There is nothing to prevent the same physical sign pattern from emerging from the employment of separate linguistic conventions, indeed, it seems inevitable that linguistic conventions should sometimes accidentally cross. It is true that political circles can be drawn around groups of people who then become subject to certain laws, rules, or regulations, including, perhaps, that they are required to use certain designated forms only for certain purposes. For example, a state law might require that Roberts Rules of Order be

followed in certain political meetings, hence that in these meetings one must not raise one's hand during a hand vote in order to request to speak. In that circumstance, hand-raising is to be a dedicated gesture constituting a vote and it will always count as such. It counts as such de jure. But there are no such circumscribed groups or prescriptions associated with natural languages; there are no linguistic rules. No idiom is dedicated de jure. Many may be unique de facto, but only within a certain tradition. Other traditions may always develop that cross over and intercede.

Glance at the section on equivocation in any informal logic text to find dozens of amusing examples of crossing conventions. Simple cases involve homonyms and syntactic ambiguities. For a different kind of example, notice that although the whole configuration, "Hit me"-said-while-playing-blackjack has its own special memetic meaning, not every instance of someone saying "Hit me!" while playing blackjack need be a replication of this pattern. Someone might say "Hit me" playfully or perhaps defiantly while playing blackjack, having composed this phrase compositionally on the model of the ordinary words and syntax involved. Then although "Hit me!" would be uttered while playing blackjack it would not be a member of the least type that means deal me another card. That a certain lineage composed of tokens of a certain sound in a certain context has a certain semantic function does not entail that all sounds of that kind in that context are from that lineage. It does not force that configuration of sound and context to be everywhere dedicated to that function. Wider context may always suddenly be needed for interpretation --needed in order to make it clear from which lineage of tokens this token has been copied, hence which is its true memetic function. Understanding the

semantics of a linguistic form never reduces merely to decoding.

' 4. Understanding what speakers intend

I have suggested that there may be much more that is conventional in language use than has usually been supposed. On the other hand, the fact that a use is conventional does not imply that its interpretation is automatic; the fact that the speaker uses a linguistic form in simple accord with its semantic meaning does not imply that it requires only to be "decoded." This is particularly obvious for words and phrases that sound alike but that are composed of semantically separate least types. Consider, for example, the "two" that means at least two and the "two" that means exactly two. It may be that these two "two"s compose separate least types. But the fact that tokens having these different semantic values were derived from separate lineages with separate histories isn't written on their faces. The hearer cannot tell from which lineage one has descended just by its sound. Indeed, figuring out their semantic values seems, ironically, to be a pragmatic problem of exactly the same kind as would be if the difference between them was merely pragmatic, for example, if one was always derived independently by Gricean implicature from the other. The results we have turned up so far then are merely negative. We have seen reason to avoid arguments about whether this or that understood meaning of a linguistic form is determined directly by its semantic value or instead is an extended meaning, the product of a figure of speech, or of an implicature. These distinctions are intrinsically vague, difficult to define and to measure. And it also seems that not much turns on them. The interesting question for pragmatics, how a hearer manages to interpret the linguistic forms she hears, is not answered merely

by the discovery that more forms than we had supposed may be conventional.

According to the Gricean tradition, during normal linguistic communication the hearer, H, recognizes that the speaker, S, has a certain intention in speaking, say, the intention to get H to believe something or, perhaps, to do something. H also recognizes that S intends H to recognize S's intention and, if sufficiently trusting, H proceeds, for this very reason, to fulfill S's intention. Given this analysis, the basic problem posed for H is to figure out what it is that S intends H to do. In my (1984 ch.3) I argued that this Gricean analysis is very implausible if taken at face value as requiring that speakers and hearers harbor multiply embedded mental representations of one another's mental representations during normal conversation.¹⁴ One reason I gave was that it seemed to have been shown by students of child language that children younger than about four, although fairly proficient in the use of language, don't yet have concepts of such things as beliefs, desires and intentions. Since then there has been much more discussion and

¹⁴ It is hard to see how Gricean intentions are to be understood other than as explicit mental representations of this sort since explicitness would seem to be necessary for embedding. It is true that Grice himself apparently had in mind purely behavioral criteria in the spirit of Gilbert Ryle (and Daniel Dennett (1987)), but that alternative is surely quite unattractive in a context that uses Gricean intentions in an explanatory way. Mere behaviors don't explain anything. Only their underlying causes are explanatory.

considerable experimental work done on the question when normal children, and also autistic children, downs syndrome children, deaf children and blind children, acquire a "theory of mind" and there seems to be very solid evidence that this depends both on the prior development of skill in handling dependent clauses ("x says, sees, is afraid, ...that p") and on enough social interaction involving discussion of mental attitudes, rather than acquisition of language depending on already having a "theory of mind." Particularly interesting is that deaf children who are otherwise normal often seem not to acquire "theory of mind" until 15 or 16 years old unless they were lucky enough to learn sign language from native signers in infancy.¹⁵ But on the other hand, strong arguments have been made that the fast and accurate way small children learn words and other linguistic forms requires active employment of a theory of mind with which to interpret the intentions behind the speech that they hear (Blume 2000). The ability to follow the focus of another person's mind, for example, by following their eyes or pointing gestures, seems to be essential to normal language acquisition. These two sets of findings and arguments seem to conflict sharply with one another. I will try to show that the evidence behind these opposing positions can be reconciled if we adopt a certain view of language perception and a flexible enough view of the different ways in which one mind can

¹⁵ For a review of the literature and an excellent discussion see Garfield et al 2001 (which strikes me as a definitive essay on this subject).

recognize another.

The studies that show that the development of a "theory of mind" depends on the prior development of rather sophisticated prior linguistic skills have relied almost exclusively on what are termed "false belief tests" to determine when their subjects have acquired a theory of mind. These tests require the subject to recognize a situation in which misleading circumstances would naturally give rise to a false belief, or they require the subject to recognize that one of their own prior beliefs has been false. What these tests seem to probe is the point at which subjects begin to recognize both their own thoughts and the thoughts of others as taking a form analogous to that of intentional representations such as sentences; the mark of an intentional representation is, of course, that unlike natural signs, it can be false. As emphasized particularly by Wilfred Sellars,¹⁶ mental states are a bit like theoretical entities. They are not readily publicly observable, nor does the ability to have thoughts imply the ability to think about one's thoughts, any more than the ability of a baby to experience pains and itches implies that the baby has concepts of pains and itches. The idea that there are such things as thoughts which are somewhat like sentences in being sometimes true and sometimes false is, indeed, a sort of "theory" of mind, call it a "representational theory of mind." To acquire a representational theory of mind is a considerable achievement, quite likely an historical cultural achievement, and certainly a considerable achievement for the individual. There is nothing surprising, then, in the fact that there is some delay even in

¹⁶ Best known is Sellars (1956).

normal children in acquiring this sort of theory of mind. The puzzle is to understand how very young children can be aware of the intentions and of the focus of attention of those from whom they learn language without yet having this sort of sophisticated theory of mind.

I will make three suggestions which when taken together may explain how this is possible. The first is a fairly simple suggestion about how the purposiveness of another's actions may be understood and taken into account without employing a representational theory of mind. The second is a more radical suggestion about the psychological processing involved in language interpretation. I argue that interpreting the meaning of what you hear through the medium of speech sounds that impinge on your ears is much like interpreting the meaning of what you see through the medium of light patterns that impinge on your eyes. When communication proceeds "Normally,"¹⁷ that is, in the way that accounts for the ultimate survival of linguistic forms and of stable speaker and hearer dispositions in their use, it is the world, not meanings, and not speaker intentions, that is immediately perceived when language is understood. Natural language operates on the senses in very much the same way that natural signs such as non-speech sounds or light operate on them. Both require to be interpreted in context, both require a great deal of filling in or "enrichment," and both sometimes produce illusion. Third, I will suggest how indications of where the speaker's attention is focused can be recognized during the

¹⁷ The capitalization is to distinguish this sense of "normal" from one that implies average or most usual. See my (1984) chs. 1-2.

process of language interpretation without the hearer's having to employ a representation of the speaker's mind or of its contents. Together these suggestions are designed to explain how a hearer can come to understand, believe or do what a speaker intends without having to think about that intention as such. The hearer needs to understand what the speaker intends, but not under that description.

First, then, about recognizing purposes. I began this essay by mentioning a number of forms that purposes can take other than the form of explicit intentions. Compatibly, recognizing purpose in an activity need not be the same thing as projecting an explicit intention as its cause. Many behaviors of lower animals that we do not take to result from explicit intentions are none the less quite obviously purposive. Moreover, it is clear that many animals, perhaps most mammals, are capable of recognizing one another's purposes. Dogs and cats, for example, could neither play together nor fight with each other without having some grasp of the directedness of one another's behaviors so as to anticipate one another's moves.. A simple form that this can take is understanding purposive behavior to be "goal directed," that is, as part of a flexible pattern with a strong tendency to produce a given effect regardless of interfering circumstances. The animal that is goal directed behaves so as to reach a certain result one way or another despite potentially deflecting forces. This has to be the understanding of both the dog and the squirrel it is chasing, and of the cat as it plays or fights with another cat. Each has to understand the other as being attracted toward certain goals in this manner, as being strongly disposed to continue a particular spurt of behavior until a certain kind of result is reached. Goals appear in this light as though they

were future states that are causing present events to happen. Aristotle called purposes "final causes," a term that expresses this idea rather neatly.

Plausibly, very small children understand the purposiveness of other people in exactly this sort of way. Studies that show, for example, that a small child can identify the candy on which another child's eyes are focused and will volunteer that that is the one that the other child wants need not indicate possession of a representational theory of mind.. The child understands toward which candy bar the other child's behavior will be directed. Similarly, children lacking a representational theory of mind need have no problem understanding that someone else is trying to help them. People who cooperate are directed toward the same goals. The young child often finds that it's mother is directed towards the child's own goals and easily learns ways to bring this felicitous circumstance about. Similarly, no representational theory of mind is needed to understand that the speech of another is purposeful to attend to its purpose, or to attempt to divine its purpose.

But this doesn't answer the deeper question what children understand the purpose of another person's speech to be, or how they divine what this purpose is in specific instances. To answer that question I must first expand on the second theme mentioned above, the proposal on language understanding. This proposal was defended at some length in my (2000, 2004) and is reviewed in this volume at the end of Chapter 6. I will briefly review my conclusions and then add some clarifications important for the present issue.

The basic idea is that during Normal conversation, it is not language that is most

directly perceived by the hearer but rather the world that is most directly perceived through language. Distal states of affairs are perceived through speech sounds just as they may be perceived, for example, through the medium of structured light during normal vision. But a number of prejudicial assumptions have to be set aside to grasp this as a genuine possibility. One assumption is that believing what one hears or sees directly must be crucially different from believing what one hears said in that the former is (1) so much more reliable and (2) alone yields information about the spatial and temporal relation of the perceiver to what is perceived. I have argued that these are differences of degree only and not relevant to the basic similarities germane to an understanding of how language is normally processed (2000, Ch. 6; 2004, Ch 9 above). A second assumption is that acquiring beliefs through the testimony of language must involve inference from premises about the speech sounds heard, hence must be indirect in a way that ordinary perception is not. In (2004, ch. 9) I discussed the distinction between direct and indirect perception, in particular, the neurology of these, and concluded that there is no reason to suppose that one of these ways of forming beliefs is more direct than the other. A related assumption is that every end organ of sense has a proprietary level, a single depth, at which it perceives directly, so that if it is possible to perceive sentences directly, it cannot also be possible to perceive their distal causes in the world directly. I will say some further words about this third assumption below. Next is the assumption that the objects of perception are always perfectly concrete, never abstract, whereas the information given in language is always highly abstract and sketchy. Last is the claim that because the information provided by language is so abstract and sketchy, the only

way for a hearer to fill out this sketchy information to yield the richer understandings we routinely communicate in conversation is through inference about the probable intentions of speakers, a kind of inference obviously not required for ordinary perception. These last two assumptions will be discussed here for the first time. I will argue that ordinary perception is also abstract. On the other hand, I will argue that much that is presented in ordinary perception is not given in the actual data that informs perception but is "filled in," as it were, during perceptual processing, nor, obviously, does this filling in involve considering anyone's intentions. The same, I will argue, applies to perception of the world through language.

I have used the different levels at which a television screen can mediate perception as an example of shifting levels of direct perception. Usually, one sees at the level of the objects and events depicted; one sees the newscaster frown or smile or sees him point to the map. Other times one might study the dots on the screen, noticing patterns in the static. The TV repairman may be able to see through the dots to perceive exactly what part of the mechanism inside is ailing.¹⁸ The housewife sees the grime on the screen and reaches for a rag to clean it. Looking out a train window when the inside of the carriage is well lit, sometimes one can either watch the people reflected inside or watch the scenery outside with equal ease, neither vision interfering with the other.

¹⁸ This perception may be just as direct as the auditory perception of the auto mechanic that the trouble is your main bearing is shot, or the perception of the physician who is able to read an echogram.

Similarly, although one normally hears through the language one hears to the events being described, educated adults can attend to phonemes instead if they try, or to phonetics (say, in order to place a foreign accent) or to words, to morphs, to sentences or to grammatical forms. That words and sentences are not what we usually hear (any more than light is what we usually see) is suggested by the ease with which people can remember the content of what was said ten seconds ago in contrast to the difficulty they have in remembering the words themselves. Little children usually cannot segment words into their component phonemes or recognize phonemes as recurring entities at all until they are five or six.¹⁹ They also have difficulty with the very concept of a word.²⁰

Indeed, I think there is evidence that small children don't perceive language at the level of words and sentences for cognitive purposes at all. Consider, for example, the following dialogue (usually used to illustrate that small children are immune to attempts to correct their speech, hence that they don't learn to talk correctly through the imposition of social sanctions).

Child: Nobody don't like me

Mother: No, say "nobody likes me"

¹⁹ Liberman et al 1974.

²⁰ Susan Cary, private correspondence.

Child: Nobody don't like me

(eight repetitions of this dialogue)

Mother: No, now listen carefully; say "nobody likes me"

Child: Oh! Nobody don't likes me²¹

What seems pretty obvious here is that the child simply is not hearing the words, any more than you usually see the dots on the television screen. Both of you are bypassing the vehicle, focusing your powers of perception directly on the content. Compatibly, three to four year olds are typically unable to say whether they have just observed something for themselves or whether they have just been told about it, mostly saying that they have observed what they really have only been told.²² Several possible interpretations for this

²¹ MacNeill (1966)

²² Whitcombe and Robinson 2000; Gopnik and Graf 1988; O'Neill and Gopnik

have been offered, none of which seem wholly adequate. No one has suggested, however, that for the child, perception through language might just be perception, observing for oneself.²³ Compare observing for oneself through a telescope.

1991; Wimmer, Hogrefe and Perner 1988.

²³ Robyn Carston writes that there is some evidence that young children treat what they have been told as being less reliable than what they observe for themselves (private correspondence). As Carston notes, this is not inconsistent with their failure to perceive linguistic forms as such. It may be that children, or that we all, also favor what we have seen over what we have felt or over perceived through non-linguistic sounds.

That perception is often²⁴ abstract is illustrated by occasions on which one can easily perceive the presence of a determinable property without perceiving its determinate value. For example, you can see that an object is some shade or other of red but would need a better light to tell which shade. Or you may clearly perceive that something is looming or moving over there in the dark or that something just shot by you but not be able to perceive what it is. Or consider More striking testimony to the abstract nature of perception are certain illusions of motion, for example, the waterfall illusion²⁵ which presents to the perceiver a landscape that is at once clearly moving and clearly staying in exactly the same place. It moves, but not from one place to another. Contradictions, presumably, are not found in completely concrete representations but only in abstract ones. The world is presented through language very abstractly, but this does not argue against the essentially perceptual nature of the presenting.

²⁴ Indeed, in another context, I would argue that it is intrinsically abstract.

²⁵ Crane 1989

Perception may be abstract, but on the other hand, it is generally more concrete than dictated by the actual data available to the perceptual process. What is perceived is filled in by context. One perceives a chair or a cat, but what is responsible for one's perception is only one side of a part of the chair or the cat. An animal nearly completely occluded behind bushes or any surface with some small openings is easily seen for what it is if its color contrasts with the foreground or if it is moving. If a surrogate dime that is distinctly larger than a real dime is placed at a distance in front of you, in the absence of clear depth clues you will perceive the dime as being dime-sized but closer than it actually is. That is, size is perceptually filled in according to context rather than actually seen, and so is depth. If you listen to the syllable the "ga" while watching the apparent speaker saying "ba" you will compellingly hear the syllable "da."²⁶ Foregrounded objects and their backgrounds are reported more accurately when they are consistent with common experience than objects in inconsistent settings.²⁷ If tiny lights are attached to a person's main joints and a video made showing only these points of light as the person walks about in the dark, watching the video one distinctly sees a person walking about.²⁸ A circle with two marks above inside and a line below is seen as a face, the angles of the marks and curvature of the line determining whether the face is happy, angry or sad.

²⁶ McGurk and MacDonald 1976.

²⁷ Davenport and Potter, 2004.

²⁸ Johansson 1975.

Stick figures are readily seen as people or as animals, standing, walking, running, lifting things, boxing, waving goodbye. If you see one of Wittgenstein's duck-rabbits surrounded by drawings of ducks you will see it as a duck, but if surrounded by rabbits then as a rabbit. If you in draw water surrounding it, you will again see it as a duck. Consider how you immediately see a certain squiggly line in a cartoon strip to be shoes in the character's hand rather than, say, a bottle because the fellow is obviously tiptoeing up the stairs in stocking feet, the clock on the wall saying 2:30. The examples go on and on.

Considering this, if language is a medium of direct perception, it certainly is not surprising that half a sentence, or even a single word uttered in the right context, often can convey the same as a full sentence would have. It is not surprising, for example, that when some one calls that they are ready, one generally knows for what they are ready. Nor is it surprising that what is conveyed by language can be much more concrete than the abstract or indeterminate semantic meaning carried by linguistic convention. For example, when one hears that Jane is sad about losing her watch and Jenny sad about losing her husband, one understands these to be rather different kinds of unhappiness, and when one hears that France is hexagonal, that a building is hexagonal and that a machine nut is hexagonal one interprets these hexagonals to have different degrees of perfection. The semantics of a possessive allows it to refer to any pairing relation coupling possessors uniquely with things possessed, such as ownership, physical possession, current responsibility for and so forth, so that "John's book" may the one he owns or carries, the one he wrote or bought or brought, and so forth. Still it is not surprising that, in context, it usually is easy to perceive immediately what specific kind of

relation is being represented. As for metaphors, gross distortions of the right kind are sometimes more readily recognized in perception than the same items shown more accurately, as illustrated by good caricatures of people's faces. Metaphor is similar. It is often understood effortlessly in context, even sometimes striking a somewhat humorous note just as pictorial caricatures do. I'm certainly not suggesting that it is obvious how perceptual identifications of these various kinds are achieved! But it does seem clear that no thoughts of speaker's minds need be involved in the process.

Most linguistic forms that have multiple senses are effortlessly disambiguated in linguistic context. Other forms are disambiguated just as easily given the immediate external context. One only has to know that we are playing blackjack ("Hit me!"), or that we are eating together ("Can you pass the butter?"). Often, perhaps typically, no consideration of speaker purposes is needed for these interpretations any more than, say, when interpreting what stick figures are doing or perceiving that it's a squirrel rather than a rabbit (obviously) that just ran up the tree. But, of course, there are also many occasions when immediate context is not enough. Instead, an understanding of what the speaker is focusing on is needed both for interpreting what is being said conventionally and for interpreting what may be being meant beyond the conventional.

If someone tells you, "John sprained his ankle" when no John is present, despite the fact that this token of "John" is being used in an entirely conventional way and has a proprietary conventional referent, clearly you will be lost, it seems, unless you know something about the speaker. The *phonologically*-typed word "John" is a hugely multiple homonym that divides into thousands of semantic least types, each conventionally

referring to a different person, and you need somehow to discern which of these various lineages this particular token of "John" has been copied from. Similarly, the content of quantifiers and definite descriptions is nearly always determined by the domain the speaker is focusing on, not just by the immediate context of speaking. And of course there are numerous other examples. These are the cases that have been thought most clearly to require that the hearer think about the speaker's intentions.

But if thinking about speaker intentions is required, does it follow that small children don't ever understand such references? On the other hand, if it is not required, should we suppose that adults employ a representational theory of mind where children are able to do without? The trick is to show how it is possible to make use of natural information about where another person's mind is focused in interpreting their words without employing a representational theory of mind. If we can understand how that is possible, we will have a hold on three subjects of interest. We will have a hold on how children manage to learn language so fast and so accurately (Bloom 2000). We will have a hold on how hearer's manage to interpret many non-conventional uses of language, for presumably they do so in much the same way that children manage in first instances to interpret conventional uses. And we will have a hold on how references that can't be interpreted from linguistic form and immediate context alone but only by following the focus of the speaker's mind might be understood both by children and adults without relying on a representational theory of mind. Employing a full-fledged representational theory of mind in order to interpret what another person is saying or meaning is always a possibility for an adult, but it is plausible, I believe, that this mechanism is seldom

required.

We have seen how small children can recognize what another person is trying to do by understanding their activity as goal directed. Now take seriously that understanding language is originally just a way of perceiving the world through another medium. As the infant matures she learns what things feel like, what those same things look like, what they sound like, often what they smell and taste like, and she learns what they sound like as filtered through the perception and speech of the people around her. She understands speakers as purposefully showing her things when they speak, and she tries to interpret what she is being shown. But how can she tell which aspects of the world around her are being manifested through which speech sounds? How can she make out what it is she is hearing about?

In the case of non-speech sounds, she can tell, in large part, because hearing is directional. She has learned or innately knows to look or feel in the place or direction from which the sound comes. In this way she soon learns to perceive many ordinary sounds as bearing information about ordinary happenings, such as doors closing, drawers opening, spoons scraping, the eggbeater beating, the toaster popping up, the toy that she drops hitting the floor, her own snuffles and cries, her rustling clothes and her footsteps. We could put it the other way around just as well. She learns to understand what it is she is seeing by its direction relative to what she hears and feels. The capacity to identify the same as the same from any of numerous perspectives including through various sensory modalities is of the essence of conceptualization (Millikan 2000).

Now consider what the child hears through language. When things go Normally,

that is, in the way that accounts for the survival of languages, another person who tells you things is like a pair of binoculars, or a camera, or a television set. They pick up patterns in the ambient energy surrounding them, interpret these patterns as manifesting certain configurations of distal objects, translate this interpreted structure into new patterns of structured energy, and pipe them over to you. Just as the child's brain (like that of a chimpanzee) is equipped at birth with a neuronal organization that is easily tuned to interpret the visual arrays around it, the child's brain (unlike that of a chimpanzee) is equipped with a neuronal organization that is easily tuned to interpret the kinds of informational patterns that language presents. You know what you are seeing through binoculars in part because you know in which direction the binoculars are pointing. Similarly, whether this is learned or innate, children grasp that what people talking to them are talking about is often where these people's eyes or hands are pointing. Nor do they have to think about the construction of the insides of these people's minds in order to understand this, any more than one has to think about how the binoculars are constructed in order to tell what one is seeing through them. Thus the phenomenon of joint looking and joint attention, so much discussed in the literature.

But, of course, people also talk about many things they are not observing. Consider then how you recognize what's in a photograph. Some things in the photograph you may recognize straight off, perhaps certain people, or a very familiar configuration of buildings. Similarly, some of the words talking about things not currently present will already be familiar to the child. At least two other techniques for recognition may be available as well.

The photograph shows a real configuration of various objects and properties on which a camera was once focused. Because the scene was once real, recognizing what is in one part of the picture may give context for recognizing what's in another part. If that's a tiger, then that's bars in the zoo and, now it's clear, that's another cage with a black panther in it, not a toy or a house cat. That's Aunt Sally in the foreground so it must be her new house in the background and the baby pulling the cap down over its eyes must be Billy. Similarly, when part of what is being talked about is understood, what past event or what general kind of event the speaker is focusing on may become clear, bringing that event into focus for the listening child as well. In this way, the content of various new words may be recognized and remembered. Somewhat similarly, if you are telling me about an event or a kind of event that I recognize, my mind will be focused where your mind is focused, and I will understand the proper names you use, your descriptions and the domains of your quantifiers accordingly, without any concern for what's inside your mind. We are focusing on the same scene or the same type of scene.

Also, one may know where the camera that took a certain photograph has been shooting. If these pictures came from your camera just home from Italy, then I know that would not Notre Dame but maybe the Milan Cathedral. Or since the picture of the elderly couple falls on the roll between the picture of the Aunt Sally and the one of baby Billy in the bath, the couple must be her neighbors who came over for tea. Similarly, people talk about what they have observed in the past, most often about what they have recently observed. If my husband announces as he returns from school that "John" has sprained his ankle I will interpret him one way, other ways if he announces it as he returns from

playing tennis or from playing music with friends. What people talk about depends on where they have been, what they have heard about or experienced. Thus the expression, when having difficulty understanding what someone is talking about, "I had no idea where he was coming from!" Most frequently, people talk about what they have just experienced. As a child can understand, say, that what a digital camera shows on its monitor will be something to which it has been exposed, children can understand that people are often talking about what they have experienced. Mother is talking to Daddy about what they, the child and Mother, just did in the park. And children can do this without wielding a representational theory of mind, just as they don't need a theory of how the digital camera work inside.

Earlier I mentioned that the child learns to understand what she is seeing by finding out how it sounds, how it feels and perhaps how it tastes and smells exactly as she learns to understand what she hears, feel or tastes by finding out how it looks. No sensory modality exhibits the true nature of a sensed object more than does any other, nor are the natural signs that any one sensory modality relies on more arbitrary, from the standpoint of the interpreter, than for another. The look of a thing no more exhibits its one true essence than does its feel or its smell. The same holds true for perception through language. Objects that a child has only perceived through language, objects that she has heard talked about but has never encountered more directly, may be thought of and judged about as easily and directly as objects she has seen but not heard named. Knowing the meaning of a word in linguistic context is exactly like knowing the meaning of a sight or of a sound in context nor is the meaning of one more arbitrary than that of the

other. The word allows you to recognize the object you are receiving information about exactly as would its sight or its sound. The meaning of a sight or a sound is learned or amplified for one when one learns the name for its object/referent exactly as the meaning of a word is learned or amplified when one learns how its object/referent looks or feels. Once initiated into the realm of language, language begins to speak to the child directly, just as sight, sound and touch do. "The child begins to perceive the world not only through his eyes, but also through his speech.....speech becomes an essential part of his cognitive development" (Vygotsky 1978, p. 32).

I have said that following the mental focus of another person can be like following the focus of binoculars or of a camera. On the other hand, so as to be understood reliably, speakers also learn to leave trail markers as they shift their focus from one domain to another. They accommodate their discourse to the hearer, taking account of what the hearer might be expected already to know or to have experienced. Whether leaving trails tailored to accommodate the individual hearer requires the speaker to employ a representational theory of mind might be a good question for empirical study. For example, how good are small children at tailoring their conversation to the needs of the particular hearer, and if they are good at it, how do they do it? But I'm not going to speculate about that.

Instead I will finish with some comments about Gricean implicatures, for it is producing and understanding Gricean implicatures that has traditionally been taken most obviously to require thoughts about thoughts. I have suggested that many interpretations of language forms that Grice described as generalized implicatures may actually be

entirely conventional. For example, if there are forms that any cooperative speaker will use only in certain circumstances, and if language conventions proliferate because they are serving cooperative functions, then these forms will inevitably come to mean that those certain circumstances are realized. Thus if the form, "Some As are B" will be used by a cooperative speaker only when she does not possess information that all of the A's are B, then in circumstances in which the speaker may be presumed to know whether all A's are B, "Some As are B" will actually entail that not all As are B. Nor does a hearer need to employ a representational theory of mind to be sensitive, in many cases, to whether the speaker has the relevant information. Knowing that another has certain information can involve no more than knowing the other has been exposed to this information or anticipating that the other could exhibit this information through language.

Occam's razor cannot discipline the ways that people actually process forms, preventing them from purposefully reproducing and routinely reading context as part of the sign to be interpreted. But if there are forms such as "Some As are B" that have different readings in different contexts, then out of context they are polysemic or ambiguous. And because it is perfectly obvious that words are part of the communication system but not so obvious, in many cases, that the context in which words are uttered (or the tone of voice in which they are uttered) is part of that system, it is easy for speakers who are not entirely ingenuous to claim that they have not "said" what they have conventionally signified because their words (taken alone) could have meant something different. Similarly, children will sometimes claim in fun to have "said" only what a certain form would mean if read compositionally, whereas the form taken as a

whole conventionally has a different or added meaning: "I didn't ask her for more cookies; I just asked whether she could reach them!"

It is true, however, that observations of this kind don't apply to many of the uses of language that Grice would have labeled "particularized implicatures." If A says to B "Let's go play tennis" and B replies "It's raining," the obviously intended interpretation that B is unwilling to play tennis surely is not conveyed conventionally. On the other hand, without entertaining any thoughts of what's in B's mind, the mere information that it is raining would undoubtedly be enough to squelch A's hope that B would be willing to play tennis, hence would serve as a reply to A's invitation. And that B is purposefully replying to A's invitation --that B is goal directed toward revealing to A his lack of inclination (read that behaviorally) toward playing tennis --can also be understood without employing a representational theory of mind. In general, when a speaker's purpose in what he says involves bringing something closely associated or implied to a hearer's mind, a hearer can understand that this effect on her is purposive without using a theory of mind. She can do this just as she can understand that the speaker is purposefully insulting, hurting, or instead trying to help or comfort with his words. I certainly don't want to claim that all particularized Gricean implicatures can be handled in this sort of way. But how far it is possible to get in understanding such implicatures prior to acquiring a representational theory of mind is not, I believe, obvious a priori. It remains an interesting question for empirical study.

It is, of course, obvious that for a hearer to understand what a speaker means is for her to grasp what the speaker intends (or at least what he purposes, taking into

account that not all purposes take the form of explicit intentions). The question that has been at issue here is how frequently it is necessary that the hearer grasp "what the speaker intends" where that phrase is to be read opaquely, and how frequently as read merely transparently. If the hearer grasps that which the speaker is trying to impart, that is, if she grasps that such and such is the case or that she is to do such and such, that will usually be enough. In deriving this, the hearer need not represent to herself that the speaker has imparting this as his intention. Indeed, in the usual case, for the hearer to diagnose the hearer's intentions she must first interpret the speaker's words (in context), not the other way around.

References Cited

- Bloom, P. 2000 How Children Learn the Meanings of Words (Cambridge MA: The MIT Press)
- Crane, T. 1989. The waterfall illusion. Analysis 483, 142-147.
- Dawkins, R. 1976, revised 1989 The Selfish Gene (Oxford: Oxford University Press).
- Davenport, J.L. and M.C Potter 2004 "Scene consistency in object and background perception" Psychological Science 15.8, 559-564.
- Davis, S. (ed.) 1991 Pragmatics: A Reader (Oxford: Oxford University Press)
- Dennett, D.C. 1987 "Intentional Systems in Cognitive Ethology: The 'Panglossian Paradigm' Defended," The Intentional Stance 237-269 (Cambridge MA: The MIT Press).
- Garfield, J., C. Peterson and T. Perry, " Social cognition, language acquisition and the development of the theory of mind" Mind and Language 16.5, 494-541
- Gopnik, A. and P. Graf 1988 "Knowing how you know; Young children's ability to remember and identify the sources of their beliefs" Child Development 56, 1366-1371.
- Harnish, R. M. 1976 "Logical Form and Implicature" In T. D. Bever et al eds., An Integrated Theory of Linguistic Ability (Harper & Row). Reprinted in (Davis 1991) 316-54.
- Johansson, G. 1975 "Visual Motion Perception" Scientific American 232, 76-89.
- Lieberman I.Y., D. Shankweiler, F.W. Fisher and B. Carter 1974 "Explicit syllable and phoneme segmentation in the young child" Journal of Experimental Child

Psychology 18:201-12

Liberman, A.M. 1999 "The Reading Researcher and the Reading Teacher Need the Right

Theory of Speech" Scientific Studies of Reading 3, 95-111.

McGurk, H. and T. MacDonald 1976 "Hearing lips and seeing voices" Nature 264, 746-748.

McNeill, D. 1966 "Developmental Psycholinguistics" In F. Smith and G.A. Miller, eds., The Genesis of Language (Cambridge MA: The MIT Press).

Millikan, R.G. 1984 Language, Thought, and Other Biological Categories. Cambridge MA: MIT Press.

Millikan, R.G. 1998 "Language Conventions Made Simple" The Journal of Philosophy 95.4, 161-180.

Millikan, R.G. 2002. "Biofunctions: Two Paradigms" In A. Ariew, M. Perlman and R. Cummins, eds. Functions in Philosophy of Biology and Philosophy of Psychology (Oxford: Oxford University Press).

Millikan, R.G. 2004. Varieties of Meaning (Cambridge MA: MIT Press)

Morgan, J. L. 1978 "Two Types of Convention in Indirect Speech Acts." In P. Cole, ed., Syntax and Semantics, vol. 9: Pragmatics (Academic Press) 261-280. Reprinted in (Davis 1991) 242-53.

O'Neill, D.K and A. Gopnik 1991, "Young children's ability to identify the sources of their beliefs" Developmental Psychology 27, 390-397

Searle, J. 1975 "Indirect Speech Acts." In P. Cole, ed., Syntax and Semantics, vol. 9: Pragmatics (Academic Press) 59-82. Reprinted in (Davis 1991) 265-77.

- Sellars, W. 1956. Empiricism and the philosophy of mind. In K. Gunderson (ed.), Minnesota Studies in the Philosophy of Science, vol. I, 253-329. Minneapolis: University of Minnesota Press. Reprinted in Sellars 1963.
- Sellars, W. 1963. Science, Perception and Reality. London: Routledge and Kegan Paul. Reprinted, Atascadero CA: Ridgeview 1991.
- Vygotsky, L. 1878 Mind in Society: The Development of Hogher Psychological Processes, M. Cole, V. Johns-steiner, A. Cribner and E. Souberman eds., Cambridge: Harvard University Press.
- Wheeler, S. 1972 "Attributives and Their Modifiers" Nous 6.4, 310-34.
- Whitcombe, E.L. and E.J. Robinson 2000, "Children's decisions about what to believe and their ability to report the source of their belief" Cognitive Development 15 no.3, 329-346.
- Wimmer, H., G.J. Hogrefe and J. Perner 1988, "Children's understanding of informational access as a source of knowledge" Child Development 59, 386-396