Aporia Vol. 10 number 1, 2000

Coherent Systems and the Location of Meaning

KIMBERLY A. PATTERSON

INTERNALISM, in the philosophy of language, refers to the belief that the meaning of an utterance (however defined) is entirely identifiable with the mental state or states of the speaker. Meaning, on this view, is "in the head." Extreme externalism, as I will use the term, refers to the belief that meaning is entirely determinable from the publicly available facts about the speaker's physical and social environment. Internalism and extreme externalism, particularly as formulated by Searle and Davidson, respectively, are as fundamentally divergent as any two philosophical approaches can be, yet both ground coherent theories of meaning. An analysis of the arguments for and against each position reveals no clear victor but an impasse, with no decisive philosophical reason to adopt one scheme over the other.

Moderate externalism holds that meaning is not entirely identifiable with the speaker's intentional states but that other factors (like the speaker's physical and social environment) also determine meaning. I propose a look at empirical evidence to suggest that this position, specifically C. Terry Warner's form of it, does a better job of explaining meaning and language than either internalism or extreme externalism.

Extreme Externalism and Indeterminacy

Our first task is to show that extreme externalism is resistant to internalists' attempts to refute it and also to point out why they are motivated to try to refute it.

Kimberly Patterson is a senior at Brigham Young University. This essay was awarded first place in the 2000 David H. Yarn Philosophical Essay Competition.

The most salient element of an extreme externalist view is indeterminacy. For Quine, this is an indeterminacy of translation; for Davidson, one of interpretation. These points are distinguishable but are so similar that for the purpose of a general critique of extreme externalism it is not necessary to draw the distinction. Indeterminacy, as I will use it, refers to the idea that there is no "fact of the matter" about meaning for the hearer (interpreter/translator) to discover in understanding (interpreting/translating) the speaker. Let's begin by seeing why extreme externalism entails indeterminacy.

Davidson and Quine propose a theory of meaning. They take the sentence to be the basic unit of linguistic meaning—a reasonable move, since so many necessary words lack meaning entirely in the absence of context in a sentence. A theory of meaning, then, would be based on sentences. Here Davidson and Quine adopt Tarski's theory of truth as the model for a theory of meaning: they assert that understanding the meaning of a sentence is based on knowing under what conditions that sentence is true. Interpretation in practice is a process of matching assertions to their truth-conditions, of matching parts of assertions to other assertions and their truth-conditions, of learning to identify conditions under which some assertions are not held true, and so on. This also seems a reasonable move: we would say a person who knows that the sentence 'there's a rabbit' is true if and only if there is a rabbit present—who knows that any sentence 'there's a(n) x' is true if and only if x is present (and, generally, also only if the speaker is indicating x)—knows the meaning of 'there's a rabbit'. However complex or sophisticated, the process is fundamentally based on matching sentences, sentence forms, or sentence parts to their truth-conditions. And therefore, knowing what a person or group of persons meant by sentences in their language could be entirely derived from knowing the conditions under which that person or people held those sentences to be true.

It follows from the Tarskian model of meaning, therefore, that *understanding* (interpreting/translating) meaning is an empirical matter, a matching up of an assenting attitude toward certain sounds to the presence of certain conditions. A theory of meaning for a language is the set of all sentences in that language matched with the conditions under which each sentence is true.

Davidson insists, and I do not feel that he could be wrong about this, that the ultimate point of language is communication, "[being] taken to mean what one wants to be taken to mean" ("Social" 6). Language is a necessarily social activity, engaged in with other human beings not for the

blind and purposeless end of stating "what I mean" as if having mental states was an end in itself. People engage in language for the purpose of *being taken to mean* something. It is unavoidably and necessarily public. And so meaning must be public (publicly available) as well, otherwise communication would be impossible.[1]

Indeterminacy results from all these considerations. There will be alternative, though empirically equivalent, ways for the hearer to match sentences with empirically observable circumstances. So long as the hearer's understanding (translation/interpretation) is internally consistent and supported by the empirical evidence, it is "correct." However, since all meaning is public and empirically based, there is no empirically motivated reason to choose one scheme of understanding over another internally consistent and empirically supported scheme.

This is not to say that there is nothing at all to choose among possible schemes of understanding (interpretation/translation). As the hearer develops her scheme (or theory, as Davidson calls it) of the language she is trying to understand, she does (and must) assume that the speaker is a rational, logical being enough like herself that there will be logical connectives in the speaker's language that correspond to her own. She must also assume that the speaker generally assents to the truth of the assertions he utters. Of course there will be exceptions to these conditions, but if the hearer does not assume a general standard of truth and a shared standard of rationality, she will be unable to even recognize the exceptions, much less interpret them as such (Davidson, "Belief" 461–63). There are other moves the hearer is likely to make, though she can develop a productive and empirically consistent scheme without them. For example, there is no empirical difference between attributing to a person the belief that "there's a rabbit" and that "there's an undifferentiated rabbit part." However, the

[1]Davidson, as far as I can tell, does not argue the point precisely this way, nor would every extreme externalist do so. For Quine, who is a behaviorist, meaning is public because the relationship between external events and personal response fully constitutes meaning, and is always empirically observable. But both assume the publicity of meaning as an aspect of their externalism (cf. Searle, "Indeterminacy") and I think the compelling claim of the publicity of language is a good reason for them to do so. It is the assumption, not the argument for it, that concerns me here.

hearer will probably, and naturally, attribute the first because it is simpler (Davidson, "Indeterminism" 6). That is, extreme externalism does not deny that the hearer may have reason and motivation to adopt one scheme or theory over another. it claims that, because meaning is public and empirical, the reasons and motivations for choosing among empirically indistinguishable schemes is only a matter of preference, not of fact. There is no way to determine conclusively what the speaker "really means," just more or less appealing interpretive schemes to adopt.

Now, since meaning is entirely public, if the meaning cannot be isolated publicly, that can only be because there is no one meaning to isolate. Meaning must be indeterminate.

The position of extreme externalism and its attendant indeterminacy has been very influential and often adopted. It has also been heavily criticized, primarily by those unwilling to accept indeterminacy. Burge, Dummett, and Putnam are externalists who have criticized Davidson, and Searle is his most vocal internalist foe. I am not interested in following the dialogue among all of them, but will outline what seem to me the strongest arguments against extreme externalism and why they fail.

The conclusion of extreme externalism is that there is no fact of the matter about what a speaker means. Searle is loudest in insisting that this conclusion contradicts our experience. He insists the speaker is in an unassailable position to know, not only what she means, but *that she means it* (cf. "Indeterminacy"). So for Quine et al. to conclude that there is no fact of the matter of the speaker's meaning can only establish that their arguments are flawed. (Contra Searle, I am not convinced that this reductio tells us which step in the argument is the problem, but the point remains that if this criticism holds, *something* is wrong with the extreme position.) This is the most intuitively appealing criticism offered, I think: no matter how many empirically equivalent interpretations of me are possible, there is one right one and I as the speaker know what it is.

But the "I know what I mean" objection does not faze the extreme externalist. The internalist can insist "I mean rabbit, not undifferentiated rabbit part" all he likes, but there is still no way for the hearer to distinguish the meaning of that claim from the meaning of 'I mean undifferentiated rabbit part, not rabbit'. The claims are empirically indistinguishable, and the hearer, no matter what the objector says, still doesn't know for sure what the speaker has in mind for 'there's a rabbit' except that it is true only when a rabbit is present.

Behind this objection is the idea that the speaker has a special relationship to her meaning that her listeners do not and cannot have—it is, essentially, the claim that meaning is not public. This purported relationship is generally referred to as "first-person authority," the exclusive authority that a speaker has to determine the "correct" interpretation of her words. Extreme externalism, denying that there is any "real" meaning or correct interpretation at all, denies that the speaker has this kind of authoritative relationship to her own meanings. But, for the indeterminist, it is no great harm to the theory to admit that the speaker has no greater insight into what the "real" meaning of a sentence, since there is no such "real" meaning to be accounted for.

There is another side to this question: the speaker should certainly be expected to have a special relationship to the meaning of her own utterances that she does not have to others' utterances, whether or not this relationship is authoritative in the manner described above. There has been a good deal of debate over whether or not extreme externalism can preserve any kind of special relationship between the speaker and the meaning of her own utterances, but I think it clearly can. Since the extremist presupposes that meaning is not identifiable with the mental states of the speaker, the speaker is an interpreter of himself as well as of others. But the speaker's interpretation of himself is immediate and does not normally require an appeal to empirical evidence. In interpreting others, the speaker requires multiple orders of beliefs: "I believe that she believes that…." However, in interpreting herself, these orders collapse together as "I believe that…" (Davidson, "Knowing").

On Davidson's account, one's self-interpretation is and ought to be more immediate and more clear than one's interpretations of others' meanings, and ought to seem impervious to correction. So extreme externalism allows for a special relationship between the speaker and the meaning of her utterances, even when meaning itself is indeterminate.

One could then try to argue that Davidson's approach fails because it does not discriminate between sentences that are empirically indistinguishable but which, nevertheless, have different meanings (cf. Searle, "Indeterminacy"). For example, 'there's a rabbit' and 'there's an undifferentiated rabbit part' always have the same truth-conditions in the same circumstances, and yet they obviously refer to two different and mentally distinguishable things. The extreme externalist would agree that those sentences are not the same, but that has no impact on the

indeterminacy argument. The whole point is that many schemes of understanding are different, yet empirically indistinguishable. What she would deny is that the difference amounts to a difference in meaning. Since there is no empirical way to ever determine whether a speaker means "undifferentiated rabbit part," "time-slice in the life of a rabbit," or "rabbit" when he says "rabbit," then they do mean the same thing. To assume otherwise is to presuppose the internalist view, not to argue against the externalist.

The extreme externalist position, with indeterminacy, is, as far as I can tell, a perfectly internally consistent theory. Once one grants the Tarskian model of meaning, the sentence as the basic structure of meaning, and the publicity of meaning—which are reasonable, defensible moves—indeterminacy follows perfectly. Objections to it, as we have seen, point out intuitive problems with the overall position but fail to show any internal inconsistency or to demonstrate why the premises should be rejected But the position has one glaring flaw: no one in their right mind would believe it. We assume, in actual interactions, that the people speaking to us have particular meanings that they intend to convey, and that our interpretations of them will be literally right or wrong depending on whether or not they agree with the speaker's intent. To discard speaker's intent as a literal determiner of meaning is to distort beyond all recognition what most of us have always thought meaning was.

Of course, people are generally wrong about all sorts of things. Unintuitiveness is not a *prima facie* reason to give up a theory. But it is a reason to look for better answers elsewhere, answers that explain language as it is actually used and accord with our common sense intuitions about what meaning is.

Internalism, Incommunicability, and Background

For internalism, all the meaning of an utterance is identifiable with the intentional mental states of the speaker. If meaning depends on the speaker's intent or purpose, then the truth-value of an utterance is not as important as the purpose for which it was uttered. And if the purpose of the utterance is its primary meaning-component, then language does not consist so much in a set of sentences as in a set of intentional actions—speech-acts.

Internalism seems to accord much more closely with intuitive and common sense ways of thinking and talking about meaning. The Tarskian model, for example, has difficulty assigning meaning to nonpropositional

sentences like questions or commands because they have no intrinsic truth conditions but have to be linked to assertions in order to be discussed in terms of truth-value. This approach can work (that is, can offer a coherent explanation of how one could develop a theory of meaning of such utterances) but is glaringly unintuitive and seems unnecessarily complicated. Internalism offers the obvious explanation that the meaning of a question or command (etc.) consists in the questioning or commanding contents of the speaker's mind. And, of course, for extreme externalism there is the problem of indeterminacy—consistent or not, the indeterminacy of meaning conclusion is wholly unintuitive. Internalism grants a fact of the matter about meaning, and plants it squarely in the speaker's mind, just where the speaker herself intuitively locates it.

To this point, however, it has not evaded the perennial communication question: assuming there is a fact of the matter about meaning, how exactly is it conveyed from the speaker to the hearer? If meaning is associated with the internal mental states of the speaker, which by definition cannot be shared, then there is no way to be sure one conveys those meanings/mental states to others. A theory of meaning that cannot account for communication, or at least the appearance of communication, is essentially useless. It is no more advantageous for a model of meaning to have a "fact of the matter" than not to have one if that fact cannot be communicated.

Searle anticipates that objection by referring to a background, a move that ought to be consistent with any internalist position even if not adopted by all (cf. *Intentionality*, chapter 5). The background, for Searle, is the set of biological capacities and non-intentional mental capacities that make language use possible. For example, a speaker's ability to meaningfully utter (rather than just parrot) a sentence like 'snow is white' requires that she be biologically capable of sight, of distinguishing colors, of intentionally producing noises, etc. These would be among the biological elements of the background against which the agent is able to perform speech-acts. But there are non-biological capacities exercised here as well. Even before the speaker has formed a belief about snow, or an intention to use the sentence to effect a particular purpose, and so on, she has a sort of "stance" before the world that determines what beliefs and intentions will be possible for her to form. For one thing, she encounters a world full of objects, as opposed to just encountering a barrage of colors, textures, and lines. She hears words instead of noises, etc. Frankly, Searle is not particularly explicit on this point (cf.

Intentionality, chapter 5), but one can get an idea of what he is talking about. I will say a few more things about the background that I think are helpful to the internalist view and to which Searle might not object.

The notion of a background undergirding language use is not unique to this dialogue. Heidegger and Wittgenstein are two who posit an inarticulable social background only against which can an agent have intelligible experience with the world. For Levinas, the possibility of intelligible experience is provided for this way: objects are seen in the first instance not just as blocks of certain colors, sizes, or shapes, but as things to use and appropriate; people are seen not just as language-using, complex objects but as fundamentally inappropriable. The general idea here is that we do not have raw perceptual contact with the world as it is. We have perceptual contact with the world against a very specific background that renders some things meaningful and some not. The background is partially socially determined, partially physically determined, and partly determined by the speaker's own intentional states, but no language is possible without it.

For example, when I am wondering what to say to entertain my five year old at the doctor's office, it never crosses my mind to tell her about the paper I am working on. It is not that I think of it and then reject the idea, but that it will not even occur to me to discuss such a thing. If I end up next to a professor at a lecture and need to make small talk before it begins, it will not even cross my mind to ask him if he got any "owies" today. The topics that will occur to me in the setting of an academic lecture are vastly different from those that will occur to me at the pediatrician's office. And again, it is not because I have to sort through all possible utterances to decide which of them will be appropriate in each case. The sorting, so to speak, is already done by the time I begin searching for words to say. In fact, in most situations I do not search consciously at all, but speak what comes naturally, given how the situation is organized or sorted. My physical capacities, my social conditioning, and my own intentions determine, within a given context, what words will even come to mind to say. Experience is not raw data, but is *organized* in terms of all kinds of aspects of the human experience. This sorting and organization is effected by the background.

This, then, is an answer to the incommunicability problem. All human beings have a shared biological background, and most have remarkably similar cultural backgrounds. I see the world in terms of whole

objects, not in terms of undifferentiated parts or time-slices. I have no reason to believe that anyone else is different in that, especially since I can read others talking about rabbits in contrast to their undifferentiated parts. I can be confident that the 'Gavagai' refers to a whole rabbit and not an undifferentiated part because the speaker and I share a background in which objects and animals are seen in the first instance in their entirety, and not as a set of parts. Likewise, when I say "rabbit" I can be sure that I am understood as meaning rabbit and not an undifferentiated part. The background gives me ground for choosing one interpretation over another, and for feeling that there is a right or wrong answer about what I choose.

But this answer is not conclusive. It simply puts the problem of incommunicability one step back. There is certainly some background for each speaker, but there is no way for the speaker to be sure that anyone else *shares* elements of her own background. There is always the possibility that some other people think in terms of time-slices or 100-miles-south-ofs and that the hearer would never be able to determine the difference. But, again, the fact that one cannot be sure whether a background is shared does not mean it is not. The notion of background is a tremendously useful one in describing how language works, but it does not decisively establish internalism over extreme externalism. It just offers a different model of meaning that is underdetermined rather than indeterminate.

Putnam's Twin Earth argument, augmented and clarified by Burge, offers what I consider the strongest cast against internalism (cf. Putnam). The argument is a thought experiment that asks you to imagine a planet that is identical in every respect to our Earth, except with regards to, say, aluminum. Twin Earth has a substance that looks, feels, and behaves just like aluminum, and is called 'aluminum' by its inhabitants, but which has a different elemental structure than aluminum has. Suppose also that the scientific communities of each planet have discovered elemental structures of both aluminum and not-aluminum, and that they and other educated citizens would be able to tell the two apart. Twin Earth also has a doppelganger, or twin, of Putnam, with exactly the same individualistically-described physical and non-intentional mental history as Putnam. (That is, if you described the history of the internal physical and mental states of each, totally removed from the social or physical contexts in which they lived, you would not find any differences between them.) Neither one knows the elemental structure of what he calls aluminum. Both think, "there's aluminum." But they each mean something different by 'aluminum'—Putnam means

aluminum, and his doppelganger means something that is *not* aluminum, something with a different elemental structure entirely. So even though, individualistically-described, Putnam and his doppelganger are thinking the same words, they are not thinking the same thing. If they had spoken about aluminum, they would not be talking about the same thing. But we have assured that this is not because one could find any difference in their individualistically-described physical or mental history, but because their physical and social environments are different. The thought experiment shows that "the identity of one's mental contents, states, and events is not independent of the nature of one's physical and social environment" (Burge 149). Or, more briefly, it shows that "'meanings' just ain't in the *head*!" (Putnam 13). Meaning cannot be entirely in the head, they conclude, if two people indistinguishable except for their external environments "mean" something different by the same proposition and even by the same thought.

The problem with this objection is that it is not clear that two people in the same mental state (with the same thought) must be referring to the same thing. For example, Searle suggests that references of (some) terms are defined indexically, by an appeal to the internal structure of the item being referred to. When Putnam thinks "water is good" he means that all the stuff identical in structure to the stuff he calls 'water' is good. Searle's doppelganger in the same situation means that all the stuff identical in structure with what *he* calls 'water' is good (cf. "Meanings"). Their mental or "Intentional" contents, as Searle puts it, are in fact *different*. Putnam's mental content is about water (indexically defined as whatever is identical to this stuff he calls water) and his doppelganger's mental content is about water (indexically defined as whatever is identical to this stuff he calls water). They do not (necessarily) have identical mental contents corresponding to different external environments at all.

Internalism does not seem subject to this objection, or, probably, to any other versions of it. Any example that seeks to show that the physical or social environment of the speaker determines (at least in part) her meanings will merely run up against Searle's indexical definitions again. Like Davidson's extreme externalism that invites attack but seems impervious to it, internalism cannot escape the incommunicability problem but forms a coherent system anyway.

So here is the problem: we have two major theories of meaning, both of which stand up to philosophic scrutiny and have plausible responses to

the objections leveled at them. Moreover, each has a serious intuitive difficulty, extreme externalism in its indeterminacy and internalism in its inability to explain communication. They are fundamentally opposed to each other—both cannot be right—but there is nothing to choose between them, except how seriously one takes their respective intuitive failings. Anyone comfortable with solipsism will find the internalist position more compelling, and anyone comfortable with indeterminacy or uncomfortable with solipsism will find the extreme externalist position more compelling. If there is a real reason to accept or reject one or the other (or both) it will have to come from outside the traditional arguments.

Warner's Moderate Externalism

Davidson's picture is one in which any rational human being with the mental capacity to match utterances to empirical circumstances in a way that makes those utterances true is capable of developing a correct theory of meaning for whatever language he is trying to understand. We have already seen that this picture is a coherent one; we have yet to prove whether it really is possible to determine meanings and even learn a language with those capacities alone.

It does seem to be possible. Verbal autistics exhibit a number of cognitive and linguistic oddities[2] (some of which we will discuss later) but are capable of producing and understanding meaningful speech, including making statements about the world and referring to objects properly. So they clearly understand the truth-conditions of sentences and how to determine them. They can also understand the meaning of commands, following them and uttering their own, which shows an ability to develop theories of meaning based on more than propositions. Many can use logic to solve problems and understand the linguistic equivalents of quantification

[2] The realization of the importance of autistic language characteristics for the study of meaning is original to Warner; the use of the characteristics of autistic language to illuminate Davidson's theory is my own.

The following information about autistic language is compiled from a variety of sources. For a good overview of autism research and a summary of relevant research, see Happe, *Autism: An Introduction to Psychological Theory*; and Frith, *Autism: Explaining the Enigma*.

and logical connectives; some are highly educated and successful professionals in technical fields. By Davidson's own account, language-using autistics have all the tools necessary to develop correct theories of meaning. In fact, verbal autistics report having learned language almost precisely the way Davidson describes theories of interpretation-developing (cf. Fullerton et al., chapter 1). They pay attention to sounds speakers utter and check the speaker's and other listeners' behaviors and surroundings to determine what those sounds must mean; through trial and error, many autistic people achieve perfect grammatical and semantic understanding of their own and even of foreign languages. Few verbal autistics make semantic errors in clinical tests (cf. Tantam, "Asperger").

Verbal autistics, then, demonstrate the sort of language people produce when they learn language as Davidson describes. And, strictly speaking, Davidson would probably be pleased with the results—verbal autistics seem to master literal meaning as he and Quine had theorized. But for one not committed to the extreme externalist project, this demonstration is a bit disturbing. The speech of even the most able autistics is so awkward and strange-sounding to typical ears that it is clear to everyone who works or deals with them that they are not using language the way the rest of us are. Even those who manage to lead self-sufficient lives and hold relatively demanding jobs report they are generally unable to recognize sarcasm or to understand implicit requests (cf. Fullerton et al.; Howlin et al. 10ff.). They are virtually unable to detect moods (like tension or anticipation) that have not been explicitly discussed and don't know when others have taken offense to them until they are told. In general, they misinterpret non-literal or untrue uses of language such as sarcasm, implied requests, pretending, jokes, and white lies even when they understand that the utterances are not literally true (Happe "Advanced"). This cannot be explained as a simple intellectual deficit, however, because autistics can be taught about non-literal uses of language and can thereafter correctly interpret such uses (in the same setting) (cf. Hadwin et al.). "David," an important autistic organizer of autistic support groups, reports having learned the clues that indicate that his listeners are bored with what he is saying or that someone is angry with him (personal correspondence). Temple Grandin, for example, is a famous and highly successful autistic who reports having learned to recognize that certain tones of voice indicate sarcasm, others irritation, etc. In *Emergence, Labeled Autistic* she describes herself as having a huge "database" of information about linguistic

communication, and a lightning-quick photographic memory that can sort through the entire database of experienced conversations in an instant. From information in that database, she is able to decipher why the people around her are acting the way they are, what they are likely to mean by their words, and so on. But she and others like her, in spite of high intelligence and immediate mental and empirical access to all relevant information, continue to misinterpret speakers in atypical ways. They become especially confused in any situation that is not sufficiently like any in the "database."

So the problem is not that autistics are unable to understand that some sentences are not meant literally, nor is it that they are unable to recognize gestures or tones of voice. The problem is also not that autistics process information more slowly than typical speakers; many verbal autistics have no learning disability at all and some are geniuses in other aspects of intelligence. The problem is that simply being able to recognize these factors, and to piece together evidence about language or language use, is not sufficient to communicate the way typical speakers do. Clearly, autistics lack some important ability that typical speakers possess that accounts for their systematic failure to intuitively grasp so many aspects of communicative language. My point here is that they do *not* lack the qualities Davidson specifies as necessary to develop theories of meaning, which suggests that having a theory of meaning is not sufficient, in actual, real-time interactions, to interpret all the communicative utterances of typical speakers. Of course, the extreme externalist position is not in the business of explaining all the ins and outs of non-literal speech in communication, so autistic language does not necessarily serve as a counterexample to that position. But Davidson does recognize communication as the primary purpose of language, and seems to intend for radical interpretation to be possible for non-literal uses of language as well as literal uses, so the discovery that a person who can develop a satisfactory theory of meaning may yet be systematically deficient in interpreting other speakers ought to provide reason for one to be nervous about adopting extreme externalism as a real descriptor of meaning.

On the other hand, these doubts about extreme externalism do not necessarily tip the balance in favor of internalism. The great strength of internalism is its intuitive plausibility: it accords with ordinary intuitions about the location of meaning as in the mental state of the speaker and concludes that there is a fact of the matter about meaning (in contrast with extreme externalism). But it violates another important intuitive insight about

meaning—that it is communicable. Internalism offers no guarantee that the meanings so carefully preserved as "facts of the matter" inside the individual mind are conveyable to other individual minds. The idea of a background purports to explain such conveyance. The background shared by any two speakers assures that, generally speaking, what the speaker means by certain utterances is what the hearer would mean by those same utterances. Given internalism, this is the same as saying that the shared background assures that the mental state of the speaker is relevantly the same as the mental state of the hearer would be were he to emit the same utterance. Understanding another, then, would require only that the hearer be aware of his own mental states. Conversation is starkly individual, with each speaker and hearer in full possession of her own meaning and physically isolated from anyone else's. There is nothing irrational or self-contradictory about this view, and again, I am not trying to conclusively refute internalism. But the picture of communication painted here seems as "alienat[ed] from the human condition and form of life" (Hacker 304) as Davidson's recursivity. It ignores the fundamentally social, in-relation aspect of the human experience and describes meaning as only accidentally communicable.

Where should we locate meaning then? There are difficulties with locating it entirely in the public arena, but there are difficulties with locating it entirely in the private arena as well. Given that the purpose of language is communication, it seems less important that a theory insist that each speaker knows precisely what she means (that is, can fix the extension of all her terms) than that meaning can be communicated. Since meaning cannot be communicated if it is entirely identified with internal mental states, we have reason to conclude that meaning is (at least partially) determined by the speaker's physical and/or social environment.

The next question, if meaning seems to be external to the speaker's mental states, is whether or not that conclusion entails *extreme* externalism. It seems not. Dummett and Burge are two who reject internalism, but also avoid indeterminism and related problems, by describing the speaker as holding himself responsible to language as it is used in the community (cf. Dummett). This moderate externalist picture is based on the belief that a large part of the externalism of meaning comes from the so-called "linguistic division of labor"—the observation that speakers often use words they cannot define, or cannot clearly define. Certain specialists in the language community are responsible for fixing the extension of a term, and the rest of the community hold themselves to the specialists' definitions.

There are also standards of usage developed by the community as a whole, which are somewhat less explicit, and individual members of the language community hold themselves responsible to those standards as well. Where they misuse a word, having mistaken its assigned or standardly-used extension, members of the community consider themselves as having been wrong, and correct their usage. There is, on this moderate externalist account, an absolute standard of meaning to which all language speakers hold themselves responsible (no indeterminacy) but that standard is *socially* determined (no internalism). There are, of course, differences among the major players here, but this is the general picture that moderate externalists offer of language and meaning.

The most important point is that externalism need not be extreme, as Dummett and Burge have shown. But there are difficulties with this model; for one thing, the linguistic division of labor does not disprove internalism any more than the Twin Earth argument did. As Davidson points out, when a speaker uses a term x whose extension he cannot "fix," "what the relevant experts mean by x" may be part of what the speaker means by x ("Social" 3). The speaker does not have to be able to fix the extension of his terms in order to have a determinate meaning in mind. On the other hand, we have already established the coherency of internalism and its relative immunity to externalist attacks; it does not disprove externalism to show that internalism is internally consistent.

More to the point, as Davidson points out, the theory still has to explain *why* the speaker holds himself responsible to the expert's decisions. Dummett and Burge imply an almost ethical commitment here, an obligation the speaker is held to in virtue of being a language-user. And one could construe the situations when language-users "stand corrected" as nothing more than evidence that people do not like to be thought of as different from others, and so will correct themselves to maintain their social status or something ("Social" 4–5). Furthermore, Davidson has shown in many places that communication is possible without convention and that, therefore, being a language-user cannot necessarily involve commitment to a certain set of conventions. The existence of malaprops, for example, shows that meaning can be understood even when the conventions are violated ("Nice Derangement").

I would like to outline a model of meaning developed by C. Terry Warner and show that it is similar to Burge and Dummett's in maintaining a moderate externalism but spells out more clearly exactly how it is that

meaning can be determined partially by the speaker's mental states and partially by factors external to her.

The model I will present is compatible with the more compelling ideas that have arisen in our discussion of indeterminacy and internalism. The most intriguing point I get from extreme externalism is the observation that the point of language is to be understandable to others. Any set of sounds that does not aim to be intelligible to others is not language, whatever else it may be. I agree with Davidson that "the intention to be taken to mean what one wants to be taken to mean is so clearly the only aim that is common to all verbal behavior that it is hard for me to see how anyone can deny it" ("Social" 6).

I also think a useful theory of meaning must incorporate an element that Davidson neglects: the background (in the Heideggerian sense I have discussed here). It is a useful way to account for the fact that we consistently treat as ordered and meaningful the endless barrage of sights and sounds that bombard our senses in every moment. It explains why we do not have to sift through an infinite list of possible utterances in every situation, but that what words or sentences will even occur to us are already sorted through and limited when we enter a situation. Even more, there is some empirical evidence for the necessity of a background. Verbal autistics, whom we discussed earlier, evidently lack some capacity (or capacities) that is necessary for complete interpretation of typical language. They share a host of typical deficiencies and abnormalities, but there are two deficiencies that experts discuss in particular as generating autistic symptoms: a significantly reduced ability to visualize the mental states of others and a significantly reduced ability to generate a coherent whole from isolated facts and events (Frith, *Autism and Asperger* 173–74; Howlin et al.; Baron-Cohen et al., *Understanding*).

The second of these might be stated as the lack of a background. Notice, for example, that the background is invoked to explain why (among other things) we do not generally have to sort through memories of previous experiences in order to figure out how to behave and interpret others in new situations; we respond, as it were, "automatically." But autistics, as pointed out earlier, do in fact consciously sort through their previous linguistic experiences in order to decide what to say or how to interpret. Many of their communicative problems can be explained as arising from the fact that they do not work from an organized network of information gleaned from past experiences but must deliberately think to

"check" a stored catalogue of information (cf. especially. Frith *Autism: Explaining*; she refers to this deficiency as a lack of "global coherence"). This might offer further reason to suppose that a background is an important component of using and interpreting language.

Warner's work in the philosophy of language brings together the importance of being understandable and the necessity of a background in a form of moderate externalism more palatable than any of the other positions we have discussed in this paper. He argues that since language is in the first instance an attempt to make oneself intelligible or understandable to a hearer, what the speaker knows of the other will also form a part of the background against which he speaks.[3] The hearer's history with the speaker (if any), her native language, her level of education, her interests, her desires, her purposes in listening, etc., are all as determinate of what sentences will be able to occur to the speaker in the situation as the other factors described. The hearer becomes, herself, part of the speaker's background. Part of what the speaker can even intend to convey, then, is determined by the person to whom she is speaking.

Now, as Putnam, Burge, and Dummett have been at pains to point out, the necessarily social character of meaning does not require all meaning to be available to any rational third-person observer. Externalism *simplicitur* does not entail complete publicity of meaning. On the model I am expounding here, the meaning of a speaker's words is partially determined by the person or group of person to whom they are addressed. The relationship between speaker and hearer frames, influences, and affects the meaning of the speaker's words. It is still true that the meaning of many sentences, statements of fact, etc., will be discernible to anyone who pays attention to the conditions under which they are uttered. But the fact that the particular hearer is incorporated into the speaker's meaning suggests that meaning exists in the relationship between speaker and hearer. The drive to be intelligible

[3]This idea is similar to Davidson's notion of theories of interpretation; he identifies as "prior" the theories we act upon to determine what/how to say to another, and as "passing" the theory that is changed as interaction with the other progresses. But for Davidson, these theories (though not necessarily "held" in any fashion by the speaker herself but attributed by the observer) are cognitive, recursive models that could in principle be completely described. The background, however, is absolutely and in principle *in*describable.

to the particular other or others to whom she speaks means that the speaker will, without taking thought to do so, use different (and in each case, appropriate) words and sentences to convey a given thought to different hearers or groups of hearers. This means, too, that the true determiner of meaning is not the speaker herself but the hearer for whom she is speaking. What she means is what she is taken to mean by the person to whom she is speaking. It is the particular hearer who has authority to determine what is meant, in virtue of being the person for whom there is meaning at all.

Warner's theory does not dismiss the speaker's intent as a component of meaning, as extreme externalism does, because the speaker's intent is determined in part by the hearer in the first place. And it does not entail indeterminacy: for one thing, indeterminacy depends on meaning's being absolutely public, but by this account meaning is strictly available only to the hearer(s) at whom an utterance is directed. A third-person observer would be able to properly assign truth-values in each case, and could likely guess at the meaning But he cannot know the meaning with any authority because he had no part in its generation. What a theory of meaning needs, besides telling us under what conditions certain utterances are true or not, is a recognition that language is an aspect of human relationships and cannot be divorced from them.

The case of verbal autism again lends empirical support to this model.[4] We mentioned earlier that autistics are deficient at visualizing other persons' mental states. Some are capable of learning *that* other people have their own mental states, and of learning what sort of behaviors tend to reflect what sort of mental states (cf. Hadwin et al.; Howlin et al. 8). A small minority, when thinking carefully, can track true and false beliefs in others (see studies cited next; and Grandin, *Emergence*). But even those are incapable of attributing mental states automatically, as the rest of us do, and are imperfect—often strikingly so—in making attributions when they do it (Happe, "Advanced"). These facts lend credence to the possibility that an automatic, preconscious, background-like understanding of other people's mental states and likely interpretations is a necessary condition of typical language use. As further support, it has been widely established that an increased ability to visualize others' mental states is

[4]This point was made in Warner, Melby, and Patterson "Language, Autism, and Consciousness of Otherness."

correlated with increased verbal skills, but not with any other factor (cf. the studies by Baron-Cohen et al., Eisenmajer, Leekam, and Prior; also Happe, *Autism*; Happe, "Advanced"; and Frith et al., "Theory,"). It is not conclusive, but reasonable to suppose, that a full theory of meaning presupposes a full ability to appreciate and adapt to the internal lives of others. At any rate, this model avoids the pitfalls of internalism and extreme externalism while preserving the important insights of each.[5]

[5] I would like to thank Terry Warner for years of patient friendship and philosophic direction; the strength of this paper depends on his brilliant insights. I would also like to thank Mark Wrathall for invaluable comments on an earlier draft, and K. Codell Carter for sympathetic guidance.

Works Cited

Baron-Cohen, Simon, Alan Leslie, and Uta Frith. "Does the autistic child have a theory of mind?" *Cognition* Oct. 1985: 37–46.

Baron-Cohen, Simon, et al. *Understanding Other Minds*. Oxford: Oxford UP, 1993.

Burge, Tyler. "Other Bodies." *The Twin Earth Chronicles*. Ed. A. Pessin and S. Goldberg. New York: M. E. Sharpe, 1996. 142–60

Eisenmajer, Richard Margot and M. Prior. "Cognitive-linguistic Correlates of 'Theory of Mind' Ability in Autistic Children." *British Journal of Developmental Psychology* 9 (1991):351–64.

"David." Email to the author, 1999.

Davidson, Donald. "A Nice Derangement of Epitaphs." *The Philosophy of Language*. 3rd ed. Ed. A.P. Martinich. New York: Oxford UP, 1996. 465–75.

_____. "Belief and the Basis of Meaning." *The Philosophy of Language*. 3rd ed. Ed. A.P. Martinich. New York: Oxford UP, 1996. 456-64.

_____. "Indeterminism and Anti-Realism" ms 1999.

_____. "Knowing One's Own Mind." *Self Knowledge*. Ed. Q. Cassam. New York: Oxford UP, 1994. 43–64

_____. "The Social Aspect of Language." *The Philosophy of Michael Dummett*. Ed. Brian McGuinness. Dordrecht: Kluer Academic Press, 1994.

Dummett, Michael. " 'A Nice Derangement of Epitaphs': Some Comments on Davidson and Hacking." *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Ed. Ernest LePore. Oxford: Basil Blackwell, 1986. 459–76.

Frith, Uta. *Autism and Asperger Syndrome*. New York: Cambridge UP, 1991.

_____. *Autism: Explaining the Enigma*. Oxford: Basil Blackwell, 1989.

Frith, U., et al. "Autism and Theory of Mind in Everyday Life." *Social Development*. 3:108–24.

Fullerton, A., et al. *Higher Functioning Adolescents and Young Adults with Autism: a Teacher's Guide*. Austin: Pro-ed, 1996.

Grandin, Temple. *Emergence: Labeled Autistic*. Novato, CA: Arena Press, 1986.

Hacker, P.M.S. "Davidson on First-Person Authority." *TPQ* 47 (1997): 285–304.

Hadwin, Julie, et al. "Can We Teach Children With Autism to Understand Emotions, Belief, or Pretence?" *Development and Psychopathology* 8.2 (1996): 345–65.

Happe, Francesca. "An Advanced Test of Theory of Mind: Understanding of Story Characters' Thoughts and Feelings by Able Autistic, Mentally Handicapped, and Normal Children and Adults." *Journal of Autism and Developmental Disorders* 24 (1994): 129–54.

_____. *Autism: An Introduction to Psychological Theory*. Cambridge: Harvard UP, 1995.

Howlin, P., et al. *Teaching Children with Autism to Mind-read: A Practical Guide for Teachers and Parents*. Chichester, England: John Wiley and Sons, 1999.

Leekam, S., and J. Perrer. "Does the Autisric Child Have a Metarepresentational Defect?" *Cognition*. 40:203–18.

Prior, M. et al., "Autistic Children's Knowledge of Thinking and Feeling States in Other people." *Journal of Child Psychology and Psychiatry* 31 (1990): 587–601.

Putnam, Hilary. "The Meaning of 'Meaning.'" *The Twin Earth Chronicles*. Ed. A. Pessin and S. Goldberg. New York: M. E. Sharpe, 1996. 3–52.

Searle, John. "Are Meanings in the Head?" *The Twin Earth Chronicles*. Ed. A. Pessin and S. Goldberg. New York: M. E. Sharpe, 1996. 89–97.

_____. "Indeterminacy, Empiricism, and the First Person." *The Journal of Philosophy* 84 (1987): 123–46.

_____. *Intentionality: an Essay in the Philosophy of Mind*. New York: Cambridge UP, 1983.

Tantam, D. "Asperger Syndrome in Adulthood." *Autism and Asperger Syndrome*. Ed. Uta Frith. New York: Cambridge UP, 1991. 147–83

Warner, C. Terry, Alan Melby, and Kimberly A. Patterson. "Language, Autism, and Consciousness of Otherness." paper presented at the Stamenov workshop on consciousness at the 1999 LSA Summer Institute.