

A Defense of Derangement
Paul M. Pietroski, McGill University

In a recent paper, Bar-On and Risjord (henceforth, 'B&R') contend that Davidson provides no good argument for his (in)famous claim that "there is no such thing as a language."¹ And according to B&R, if Davidson had established his "no language" thesis, he would thereby have provided a decisive reason for abandoning the project he has long advocated--*viz.*, that of trying to provide theories of meaning for natural languages by providing recursive theories of truth for such languages. For he would have shown that there are no languages to provide truth (or meaning) theories of. Davidson thus seems to be in the odd position of arguing badly for a claim that would undermine his own work.

I think Davidson may well have undermined *a* philosophical project that he once advocated. But his remark has been taken out of context. His conclusion in 'Derangement' is that "there is no such thing as a language, not if a language is anything like what many philosophers and linguists have supposed."² B&R ignore the crucial qualifying remark, which has a quite precise meaning in the context of the whole essay. And one must understand this qualification in order to interpret Davidson charitably. His account of interpretation leads to scepticism about the possibility of providing a *complete* theory of understanding. But, I claim, such scepticism undermines neither the attempt to provide recursive theories of "truth-in-L," nor the hypothesis that speakers deploy such theories. The substitution instances of L, however, will not be languages in the 'Derangement' sense. For according to Davidson, there are no languages in that sense. And I think *this* conclusion is both important and defensible.

1. The Arguments

B&R's paper is exegetically complex. Their main focus is "the question of whether a practitioner of the Davidsonian program is committed to the claim that there is no such thing as a language;" and they attribute to Ramberg two arguments that purport to establish such a commitment.³ But B&R grant that Ramberg does not explicitly offer the first, and that Davidson does not offer the second. Nonetheless, discussion of these arguments may serve to make Davidson's actual position clearer and more plausible.

1.1

The first argument B&R consider is the following:⁴

(1) The successful application of one and the same theory of truth is the only criterion for the identity of a language. (2) There could always be more than one successful truth theory interpreting an individual's speech, and there is no fact of the matter as to which is the correct theory. (The indeterminacy of interpretation.) *So* (3) there could always be more than one answer to the question of which language a given individual speaks, and there is never a fact of the matter as to which language the individual speaks. *But* (4) if there is such a thing as a language, there ought (*sic?*) to be a fact of the matter as to which language an individual speaks. *Therefore*, (C) there is no such thing as a language.

I agree that this is a poor argument. But B&R offer no textual support for attributing it to Davidson, whose explicitly stated views about indeterminacy preclude him from making such an argument.

Indeterminacy of meaning or translation does not represent a failure to capture significant distinctions; it marks the fact that certain apparent distinctions are not significant. If there is indeterminacy, it is because when all the evidence is in,

alternative ways of stating the facts remain open...[for example] if the numbers 1, 2, and 3 capture the meaningful relations in subjective value between three alternatives, then the numbers -17, -2, and +13 do as well. Indeterminacy of this kind cannot be of genuine concern.⁵

Davidson also offers the analogy of measuring temperature. The fact that we can correctly describe the freezing point of water in more than one way—*e.g.*, '32 degrees (Fahrenheit)' or '0 degrees (Celsius)'--does not show that *temperatures* are somehow indeterminate. And just as numbers can capture all the empirically significant relations among weights or temperatures in infinitely many ways, so one person's utterances can capture all the significant features of another person's thoughts and speech in different ways.⁶

Davidson takes a truth theory to be an empirical hypothesis about the linguistic competence of a speaker. There may be alternative (correct) ways of stating the facts about linguistic competence. But this indeterminacy of truth theories does not suggest that "the states of mind of the speaker or thinker thus captured are somehow vague or unreal;" and to suppose the contrary is like supposing that the "difference" between the properties of "being a yard long and 36 inches long in a yardstick were a difference in the yardstick itself."⁷ Pressing this analogy, instead of (2) above, consider

(2*) There could always be more than one successful temperature scale that measures a range of temperatures, and there is no fact of the matter as to which is *the* correct scale.

One cannot infer from (2*) that

(3*) There could always be more than one answer to the question of what the

temperature is, *and* there is never a fact of the matter as to what the temperature is.

The second conjunct of (3*) does not follow, given the possibility that *both* scales can be used to correctly describe the same phenomenon--*viz.*, the temperature. And on the very pages B&R cite as the location of Davidson's views about indeterminacy, he says that "Quine's own views undermine the idea that ontology can be relativized;" all that gets fixed by a relativization to some language is "the way we answer questions about reference, not reference itself."⁸

Thus, one can grant premise (2) and the first conjunct of (3) in the argument B&R consider. But the second conjunct of (3) does not follow; and this is what leads to (C). In short, indeterminacy (in Davidson's sense) as to which of two languages, L and L*, an agent A speaks would hardly show that there are *no* languages. For we could correctly say *both* that A speaks L, and that A speaks L*; although this would just be a dramatic (and potentially misleading) way of saying that there are two correct ways to characterize the same phenomenon--*viz.*, A's linguistic competence. Perhaps Davidson is wrong to think that even this weak kind of indeterminacy ever arises. But even if it does, "linguistic eliminativism" does not follow. Davidson certainly does not *think* it follows, nor is there is any reason to think he should.⁹

1.2

Before turning to the second argument B&R consider, I want to rehearse the main points of 'Derangement,' and to emphasize that its conclusion is more modest than it appears when taken out of context. But the modest conclusion is of considerable importance. And I find the argument compelling.

Davidson's primary concern is the fact that speakers sometimes use words in nonstandard ways, but are understood nonetheless. The examples he focusses on are--with an important

exception I discuss below--malapropisms. Consider Mrs. Malaprop's utterance, 'There's a nice derangement of epitaphs.' Despite her "linguistic error," she is understood as saying *there's a nice arrangement of epithets*. But as Davidson notes, the main problem is *not* that of explaining how the hearer succeeds in interpreting malaprops. And he offers a distinctly Gricean explanation from the interpreter's perspective:

The absurdity or inappropriateness of what the speaker would have meant had his words been taken in the 'standard' way alerts the hearer to trickery or error; the similarity in sound tips him off to the right interpretation.¹⁰

Here talk of error and standards is to be cashed out in terms of claims about what a good dictionary would (or better, ought) to say, or by polling experts who the speaker trusts. So what *is* the issue?

Davidson takes as given that we must maintain a firm distinction between speaker's meaning and literal, or what he calls "first," meaning. First meaning is usually that which is found in a good dictionary. It is also first "in order of interpretation." For at least typically, the interpretation of poetry, metaphor, irony, *etc.* requires that we first recognize the "literal" meaning of the relevant utterances. Another way to get at first meaning is to consider the order of intentions in the speaker. The ironist presumably intends to evoke a certain response in the listener *by* using words that have a certain first meaning, but not *vice versa*. No doubt the notion of first meaning requires further investigation. But the project in 'Derangement' is to (a) show that sharing a language cannot by itself explain how successful communication is possible (while maintaining the distinction between speaker and first meaning), if languages have the features

that languages have traditionally been thought to have, and then (b) sketch an alternative account of communication.

Davidson exposts what he takes to be the traditional account of language by stating three theses about first meaning. First, it is *systematic* in the familiar sense. The interpreter can, on the basis of learning finitely many words and composition rules, come to understand novel sentences; and "there is no clear upper limit to the number of utterances that can be interpreted."¹¹ Davidson still holds that only a recursive theory with a finite base can provide an adequate model of the interpreter's ability in this respect. But as always, he stops short of saying that the interpreter *knows* (or even believes) such a theory. Talk of the interpreter (or speaker) using such a theory is, for Davidson, a shorthand way of saying that the interpreter (or speaker) makes use of a competence that is correctly described by the theory. But semanticists, at least, need a recursive theory to describe this competence.

The second thesis is that first meanings, both of words and the sentences they make up, are *shared*. This claim is close to tautological, once we assume that communication is ever successful. For successful communication presumably requires shared meanings; although Davidson grants that success here can be a matter of degree. Talk of "sharing" can be cashed out as follows: The interpreter uses her theory to understand the speaker; and the speaker uses the same theory--or at least an *equivalent* theory (since Davidson allows for indeterminacy)--to guide her speech. This is why the problem posed by malaprops is not merely that of explaining how the interpreter manages to understand, but also that of how the speaker manages to "get away with it." For the speaker and interpreter must share a theory. Mrs. Malaprop's interpreter may think that 'epitaph' means *epitaph*, and that Mrs. Malaprop has made a linguistic error. But Mrs.

Malaprop does not think she has deviated from standard use; and while she might say (and in some sense think) " 'epitaph' means *epitaph*," she also thinks that (as we theorists would put it) 'epitaph' means *epithet*.

Nonetheless, Davidson maintains that first meanings are both systematic and shared, and that this is not incompatible with the understanding of malaprops. But to accept these two principles is not yet to suppose that the speaker and interpreter share a *language*, unless one supposes that a language just *is* whatever theory the (communicatively successful) speaker and interpreter share. And it is plausible to think that first meanings--and thus the shared, recursive theories--are governed by *conventions* that are learned *prior to* the communications whose success is to be explained. For meanings do indeed seem to be conventional; and it is hard to see how the sharing of a theory could explain successful communication, unless the theory was shared prior to the communication. But Davidson rejects this plausible thesis about meaning. And his claim that there are no languages must be understood in this context. Indeed, this claim amounts to no more (or less) than the following: There is no x, such that (i) we can provide a finite recursive theory of x, *and* (ii) communication is successful--when, and to the degree that, it is successful--in virtue of speakers and interpreters sharing x, *and* (iii) x is governed by conventions that speakers and interpreters learn prior to their communication.

It remains to be seen why Davidson takes malaprops to show that theories of interpretation are not governed by prior convention. Here it is crucial to note that Davidson assumes that Mrs. Malaprop's utterance has as its first (literal) meaning: *there's a nice arrangement of epithets*. I think this is correct. And the puzzle is that, from our standpoint, a homophonic translation of her utterance would be mistaken; hence, Mrs. Malaprop does not

speaking *our* language. So it is hard to see how she uses her utterance to communicate with us. But as Davidson notes, malaprops are not the only examples of nonstandard (in the dictionary sense) first meanings. Alfred MacKay once charged that Keith Donnellan's views about reference amounted to the claim that words mean whatever we want them to mean.¹² The central example concerned Humpty-Dumpty's claim that his words 'There's glory for you,' could mean *there's a nice knockdown argument for you*, despite Alice's having no clue as to what Humpty meant.

While rejecting Humpty's general semantic views, Donnellan did make the following claim:

If I were to end this reply to MacKay with the sentence 'There's glory for you' I would be guilty of arrogance and, no doubt, of overestimating the strength of what I have said...but given the background...I would be understood.¹³

Donnellan also claimed, and Davidson agrees, that those words would have literally meant *there's a nice knockdown argument for you*. But it may be best to construe Davidson as challenging others to explain the Malaprop/Donnellan phenomenon *without* (i) making the same assumptions about the first meanings of the relevant utterances, or (ii) doing damage to the first-meaning/speaker-meaning distinction. And again, any explanation of this phenomenon must explain how Malaprop and Donnellan (but not Humpty) "get away with it," in addition to explaining the interpreter's success.

If Mrs. Malaprop's utterance literally means *there's a nice arrangement of epithets*, then an adequate theory of interpretation for Mrs. Malaprop must have this as a consequence. But an interpreter encountering Mrs. Malaprop for the first time will not *enter* the conversation with a theory according to which Mrs. Malaprop's utterance literally means *there's a nice arrangement of epithets*. Whatever "prior theory" the interpreter has will have to be adjusted, if it is to yield

that correct interpretation of Mrs. Malaprop's utterance. But--and this is the crucial point--the interpreter's theory must be adjusted in light of the very utterance that is to be interpreted. Thus, we cannot restrict the interpreter to theories that she had prior to the utterance. The (adjusted) theory that the interpreter actually uses to interpret the speaker is what Davidson calls the interpreter's "passing theory." The speaker's prior theory is what she believes the interpreter's prior theory to be; and the speaker's passing theory will be the theory she intends her interpreter to use. Unless Mrs. Malaprop comes to recognize her error about her interpreter's prior theory, Mrs. Malaprop's passing theory will continue to include the claim that 'epitaph' means *epithet*. (But recall that talk of *her* theory is really talk of *our* theory of her linguistic competence.)

The same points would hold, *mutatis mutandis*, of Donnellan's first use of 'There's glory for you.' to mean *there's a nice knockdown argument for you*. No (normal) interpreter would come prepared with a theory that would have delivered this interpretation of Donnellan's words; and Donnellan is unlikely to expect his interpreter to come prepared with such a theory. But given the context, Donnellan might well intend that (say) MacKay use a theory that yields this interpretation; and if MacKay also makes the corresponding adjustment to *his* prior theory, then he and Donnellan can share a passing theory. Note that the context alone (*i.e.*, in the absence of the utterance) would not lead MacKay to adjust his prior theory as required. For the passing theory that yields the correct interpretation of Donnellan's words is a response to those very words. Of course, once an interpreter becomes familiar with Malaprop or Donnellan, she will be prepared. It may well *become* part of an interpreter's prior theory of Mrs. Malaprop (but not of others) that utterances of 'epitaph' mean *epithet*. Or what comes to the same thing, the interpreter may come to believe that Malaprop-utterances of 'epitaph' mean *epithet*. And similarly, *mutatis*

mutandis, for interpreters of Donnellan. But there will still be many other opportunities for interpretation that are not anticipated by the interpreter's prior theory.

Davidson's account of the malaprop phenomenon is that communication is successful to the degree that speaker and interpreter share a *passing theory*. And since the interpreter constructs her passing theory after the speaker has made the utterance to be interpreted, understanding malaprops is consistent with this hypothesis. But, Davidson says,

A passing theory is not a theory of what anyone (except perhaps a philosopher) would call an actual natural language. 'Mastery' of such a language would be useless, since knowing a passing theory is only knowing how to interpret a particular utterance on a particular occasion. Nor could such a language, if we want to call it that, be said to have been learned, or to be governed by conventions. Of course things previously learned were *essential to arriving at the* passing theory, but what was learned could not have been the passing theory.¹⁴

There are several important points here. While passing theories are not themselves theories of languages in any traditional sense, they do not spring into existence *ab initio*. What speakers have learned will be relevant—indeed, *essential*—to the passing theories they construct. But precisely because passing theories are a response to the actual utterance to be interpreted, they cannot be determined in advance by convention.

An interpreter's prior theory of a speaker (or the speaker's prior theory of the interpreter) is not a candidate for being a natural language in the traditional sense either. For it can contain a great deal of idiosyncratic information—*e.g.*, that Malaprop-utterances of 'epitaph' mean *epithet*. Moreover, speakers and interpreters will not always share prior theories; and when they do, it

will be (at least to some degree) fortuitous. But most importantly, because prior theories are "unadjusted," we cannot account for the malaprop phenomenon if we hold that successful communication is a matter of sharing prior theories. Commonalities of prior theories may help explain how the speaker and interpreter come to share a passing theory. And in many cases, the deliverances of passing theories that agents come to share may not differ (or not differ noticeably) from those of their prior theories. But according to Davidson, the sharing of passing theories is that in *virtue of which* communication is successful.

A Chomskian (and I am one) will think--quite independently of the issues discussed here--that we should reject the idea that English and French are natural languages, and instead take the notion of an idiolect as basic; where an idiolect is a more-or-less stable "adult state" of a (largely innate, though environmentally affected) linguistic system that comprises (i) a set of principles that govern the elements, acquisition, and subsequent expansion of a *lexicon*, and (ii) a set of recursive *rules* that, given a lexicon, determine the set of grammatical utterances for the idiolect. It may be that all humans share the "initial state" of such a linguistic system; and most "English-speakers" will share considerably more than this. But idiolects will rarely, if ever, be wholly shared. Davidson describes this as the idea that speakers share a "general framework." Of course, providing a model of communicative success is not the Chomskian's goal. And in any case, speakers can share a framework while failing to understand one another. But if speakers do share a framework, one would expect this to play at least *some* (presumably causal) role in the explanation of communication. And Davidson does not deny this possibility. Indeed, he says:

The general framework or theory, whatever it is, *may be a key ingredient in what is needed for interpretation*, but it can't be all that is needed since it fails to

provide the interpretation of particular words and sentences as uttered by a particular speaker.¹⁵

It is far from obvious that a non-human could *ever* interpret Mrs. Malaprop's utterance. For without a prior *grammar* of a (roughly) Chomskian sort, it may well be that the very project of interpretation could not get off the ground. Of course, this is in some sense an empirical question. But Davidson does not assume that "pure reason" suffices for interpretation. Moreover, if Davidson's Gricean explanation of how the interpreter manages to understand malaprops is correct, then at least some (and perhaps much) of the "interpretive work" is done by the prior theory. The conventionally determined interpretation of a malaprop is that which "alerts the hearer to trickery or error;" and the differences between the interpreter's passing and prior theory may well be small when compared to the similarities. If Chomsky is right, at least some (and perhaps much) of the interpretive work done by a prior theory is done by an innate system of rules. But again, the sharing of such rules is not that in virtue of which communication is successful.

Davidson suggests that linguistic ability is the ability to construct a passing theory that covers with the passing theory of another. But he grants that this proposal is "so nearly circular that it cannot be wrong: it comes to saying that the ability to communicate by speech consists in the ability to make oneself understood, and to understand."¹⁶ But, one might think, if speaker and hearer at least begin their communication with the ability to construct convergent passing theories, perhaps therein lies a sense of "knowing a language" that is wholly antecedent to communication. Davidson is sympathetic with this idea. But as he says, it erases "the

boundary between knowing a language and knowing our way around in the world generally." Moreover, there are no *rules* for arriving at passing theories, just "rough maxims."

A passing theory really is a theory at least in this, that it is derived by wit, luck, and wisdom from a private vocabulary and grammar, knowledge of the ways people get their point across, and rules of thumb for figuring out what deviations from the dictionary are most likely. There is no more chance of regularizing or teaching the process of creating new theories to cope with new data in any field-- for that is what this process involves.¹⁷

In short, there is only one "thing" that an interpreter has prior to communication, such that when coupled with the interpreter's prior theory, it suffices for yielding an interpretation of the speaker's utterance: General Intelligence. (In more Davidsonian terms we might speak of *rationality*, or *charity in interpretation*.) But there is no hope of providing a recursive theory of General Intelligence (or rationality). Indeed, it may not even make sense to speak of providing a recursive theory here.

Once the distinction between prior and passing theories has been drawn, I take it that the conceptual distinction matters even if the deliverances of prior and passing theories are the same. Not modifying a prior theory is, on my reading of Davidson, just as much an interpretive decision as making a modification; though in the former case, the "decision" will typically not be associated with any conscious processing or feelings of "dissonance." But successful communication is always a matter of converging *passing* theories; and general intelligence is always implicated here, if only by giving "tacit approval" to the deliverances of prior theories. Moreover, even if such tacit approval is often granted, one cannot speak of rules *governing*

passing theories. For the "rules" could always be overridden in cases of the Malaprop/Donnellan sort; and one cannot capture the extent of these cases formally or in advance.

To sum up: Sharing passing theories suffices for communicative success; but passing theories are neither governed by convention nor shared in advance of the communication. Prior theories and "frameworks" can be conventional and/or shared in advance; but sharing them does not suffice for communicative success. Being intelligent may (in conjunction with a prior theory) suffice for communicative success, and this property may be shared by speakers in advance of their communication. But if being intelligent is not simply a matter of following rules, *a fortiori* it is not a matter of following conventional rules laid down in advance. And perhaps more importantly, there is no hope of stating a recursive theory of intelligence. So Davidson concludes that there are no languages in the traditional sense.

1.3

The second argument that B&R consider is attributed directly to Ramberg:¹⁸

- (1) The construction of a truth theory for a natural language is an incompletable process.

So

- (2) we never apply the exact same truth theory to any two speakers.
- (3) Linguistic convention can play a role in explaining communication only if a single truth theory can be applied across speakers or over time. *So*
- (4) linguistic conventions play no role in explaining communication.
- (5) If the notion of a language has an essential role to play in explaining linguistic communication, then language must consist in a regular pattern of shared conventions explaining the possibility of linguistic communication. *Therefore,*

(C) the notion of a language plays no essential role in explaining the possibility of linguistic communication.

It is assumed that (C) is an adequate gloss of Davidson's "no language" thesis.

B&R devote much of their discussion of this argument to Ramberg's understanding of (1), and why (1) fails to provide reason for (2); although, as B&R themselves note, Davidson would defend (2) by citing the malaprop phenomenon, not (1). Nonetheless, Davidson does accept (2) in some sense--at least if we substitute 'rarely, if ever' for 'never'. But Davidson does not actually speak of "applying truth theories;" and this locution is potentially ambiguous in this context. In one sense, a hearer H applies a truth theory Θ to speaker S at time t, if and only if Θ is H's passing theory of S at t. That is, Θ would have to be a set of axioms and recursive rules (of the Tarskian sort) that provide a model of the competence (as adapted to S at t) that actually produced H's interpretation of S at t. In this sense, H applies her prior theory to S only if H's prior and passing theories of S are equivalent. And only rarely, if ever, will the same truth theory be applied in this sense to different speakers. For it is safe to assume that most speakers (or time slices thereof) will have some linguistic idiosyncracies; and even though passing theories are driven by the need to handle *particular* utterances, any Davidsonian truth theory must interpret--*i.e.*, assign truth conditions to--the indefinitely many utterances that a speaker of the "language" in question might produce.

There is, however, another perfectly good sense of 'apply' in which hearers apply prior theories that subsequently get modified. Again, it is because of such application that interpreters are alerted to trickery or error. Moreover, nothing Davidson says rules out the possibility, or even the likelihood, that an interpreter applies a particular set of axioms and rules \mathfrak{R} to different

speakers, S1 and S2, in the following sense: \mathfrak{R} is a (perhaps improper) subset of the interpreter's *prior* theories of both S1 and S2; and by making suitable adjustments to these prior theories, the interpreter arrives at distinct passing theories, Θ_1 and Θ_2 , of S1 and S2; where \mathfrak{R} may or may not be a subset of Θ_1 and/or Θ_2 . \mathfrak{R} might include, for example, a rule for determining the truth conditions of a declarative utterance as a function of the component noun and verb phrases; and this rule may well go unmodified in the passing theories Θ_1 and Θ_2 . We might think of \mathfrak{R} as a "fragmentary" truth theory. And the claim that interpreters apply (in the sense that prior theories are applied) such fragmentary theories to speakers is perfectly consistent with Davidson's remark that what interpreters learn prior to the communication (*e.g.*, grammatical rules) will be essential to the passing theories they form.

The point is that Davidson is committed to (2) in the argument above, only in the more restrictive sense of what it is to apply a truth theory to a speaker--*viz.*, the sense in which applying prior theories (and/or fragments) does *not* count as applying truth theories. But he is not committed to (3) in this restrictive sense. Convention may help explain why interpreters hold certain prior theories (and/or certain fragments thereof). Thus, convention may *partially* explain why interpreters form the passing theories they do. A conventionally determined set of rules may not *suffice* to explain successful communication. But this claim does not entail (3); and without (3), (4) and (C) do not follow. I return to these issues below. But as with his "no language" thesis, Davidson's rejection of convention must be understood in the context of the whole paper.

B&R make a similar mistake when they say that, according to Davidson, "nothing interesting can be said about this ability to construct passing theories." They infer this from Davidson's claim that passing theories are derived by "wit, luck," *etc.* And this is at least one

source of B&R's concern that Davidson is undermining his earlier work. For if linguistic competence is the ability to construct passing theories, and nothing interesting can be said about such construction, then why try to state recursive rules that yield T-sentences for object language sentences involving adverbs, intentional verbs, definite descriptions, quantification, and so on? But there is a difference between (a) scepticism that a complete theory of some phenomenon can be provided, and (b) the claim that nothing interesting can be said about the phenomenon. We know that gravity has something to do with the fact that leaves fall, and that the attraction between molecules has something to do with the flow of water in a stream. But at least for now, we have nothing like a complete theory of either phenomenon. And even if we have nothing to say about how passing theories are constructed from prior theories, there may be much we can say about the construction of (the relatively non-idiosyncratic portions of) prior theories.

Moreover, if there are \mathfrak{R} 's that interpreters apply--in the more inclusive sense described above--to a relatively wide range of speakers, then Davidson's semantic program is still of considerable interest. At any particular time, semanticists work on a particular fragment (*e.g.*, intentional contexts) of what we ordinarily call a language; though, of course, a fragment of an actual complex language is a simple possible language. But if the argument in 'Derangement' is correct, we will never be able to draw together all the "fragments of English" that (in the fullness of time) semanticists work out, and get something that meets all three conditions for being a language in the traditional sense. Indeed, it may turn out that no passing theory will ever include all and only such fragments. That is, there may never be, even for a moment, a perfectly "ideal" speaker of English. But this leaves open the possibility that the theory for each "fragment of

English" figures in a significant number of actual passing theories. If Davidson is right, we should not hope for *more* than this possibility. But this would not be nothing.

Let me end this section by noting that Fodor advances a structurally similar thesis for similar reasons.¹⁹ Simplifying matters somewhat, Fodor thinks the mind is composed of a number of *modules* (e.g., visual, auditory, and linguistic) and a *central processor*. The former are characterized primarily by the fact that, as a consequence of their design, they (i) fail to take into account all the information available to the system as a whole, and (ii) respond to the inputs they do receive in a pre-ordained reflex-like fashion. The central processor, on the other hand, can (at least in principle) access all the information provided by all the modules; and it is where "intelligence" resides. The central processor does not operate according to pre-established principles. For it can, and often does, produce novel solutions to cognitive problems. There are costs associated with the "information encapsulation" of modules. The persistence of optical illusions is a paradigm example. But modules are fast and typically efficient at their relatively narrow cognitive tasks. If we had to infer what was in our visual field by applying "pure reason" to retinal images, we would die before making a move. And the very idea that such inferences are possible--even "in principle"--may be a philosophical fantasy.

Fodor thinks we can provide theories of modules, precisely because they are informationally encapsulated. But once we turn to the "smart" central processor, the prospects for theory are dim. Fodor does not conclude that cognitive science is hopeless, or that there is nothing interesting to say about the mind. But he does think that theoretical success will continue to track the degree to which the portion of the mind being investigated is modular.

I think the similarities between Fodor's claim about the organization of the mind and

Davidson's claim about the nature of interpretation run deep; because I think that significant portions of our prior theories are the result of largely innate, modular processing. But I will not press this issue here. For my point is simply that Davidsonians are free to hold (along with Davidson) that the qualified rejection of language in 'Derangement'--and the associated scepticism about providing a *complete* theory of interpretation--in no way suggests that the familiar semantic project associated with Davidson's early work is irrelevant to understanding our linguistic competence. And here Fodor provides a useful analogy: One can learn a lot about the mind by studying modules, and one can learn a lot about interpretation by doing formal semantics; but a complete theory of cognition or interpretation would require a theory of what it is to be intelligent.

2. In Search of an L

Thus far, I have defended Davidson's "no language" thesis as a more modest thesis than it might initially appear. But one might think that even this modest rejection of language is implausible and/or incompatible with other aspects of Davidson's semantic program.

2.1

B&R challenge the assumption that Mrs. Malaprop's utterance literally means *there's a nice arrangement of epithets*, suggesting the obvious alternative--viz., that the interpretation of malapropisms may be "parasitic on the expectation of general adherence to linguistic conventions shared across speakers or at least over time," and that

...contrary to the impression Davidson gives, in practice, we hold individual speakers responsible to the norms of speech of their linguistic community, and only rarely resort to ascribing completely idiosyncratic uses of language. The

most common reaction to malapropisms, as well as slips of the tongue and Spoonerisms, in the speech of our interlocutors is correction.²⁰

Here B&R ignore the Donnellan example, in which *correction* is not at issue. But in any case, Davidson never denied that we hold speakers responsible to community norms. And he grants that Mrs. Malaprop has made a linguistic error in the dictionary sense. But correction in this sense--you mean *epithet*, not *epitaph*--is possible only if one has understood what the speaker meant, despite her flouting of the convention. Davidson wants to know how understanding is possible in such cases, given that understanding presupposes that the speaker and hearer share a theory. B&R offer no suggestions as to what such a theory might be, nor how it could be governed by prior convention.

Dummett has suggested that speakers can share but *imperfectly grasp* their public language.²¹ And this is indeed an alternative to Davidson's proposal. But one wants to know how an agent can be mistaken about her own language, and how a community of agents--none of whom perfectly knows the language--establish a set of conventions that determine the proper use of language. Moreover, any alternative to Davidson's proposal must avoid the patently false consequence that there is *nothing* wrong with interpreting Mrs. Malaprop as saying *there's a nice derangement of epitaphs* (and similarly for Donnellan); yet it must not lapse into Humpty's view that good interpretation is merely a matter of grasping the speaker's intention or "point" in making utterances. I do not say that Dummett's (or any other) alternative to Davidson is hopeless. But without a defense of some such proposal, the claim that there *may* be a convention-based account of the malaprop/Donnellan phenomenon lacks force.

B&R go on to say that, even if we allow that some word(s) are used idiosyncratically in a malaprop, our understanding of the malaprop depends on our antecedently fixed interpretations of other words in the sentence (and antecedently fixed composition rules). But suppose that instead of 'There's glory for you.', Humpty had uttered, 'Ouch! Go away!' Then according to Davidson, Donnellan's inscription of 'Ouch! Go away!' (in the context of his reply to MacKay), would have literally meant *there's a nice knockdown argument for you*. Nothing B&R say challenges this claim. And in any case, they are again mistakenly assuming that if *anything* speakers learned prior to the conversation figures in the explanation of successful communication, then Davidson is wrong.²²

2.2

B&R rightly ask what truth theories are theories *of*, once we accept the picture of interpretation in 'Derangement'. But if 'truth theory' just means *passing theory*, we can say that truth theories are theories of languages, as long as we recognize that such "languages" are ephemeral and not governed by convention. The terminology, however, is irrelevant. A passing theory is a theory of an agent's linguistic competence--or better, an agent's linguistic competence as adapted to a particular communicative situation. On the other hand, if prior (and fragmentary) theories count as truth theories, such theories need not be theories *of* anything at all. Pressing an earlier analogy, we might think of truth theories in this sense as abstract constructs against which utterances are *measured*--or better, the *prior* constructs that, when modified, become *passing* constructs. But I think Davidson must hold that truth theories are either passing theories, or at least models of "subcompetences" that speakers deploy in the course of forming passing theories. For as B&R note, Davidson elsewhere says that:

The question whether a theory of truth is true of a given language (that is, of a speaker or group of speakers) makes sense only if the sentences of that language have a meaning that is independent of the theory (otherwise the theory is not a theory in the ordinary sense, but a description of a possible language)...If the question can be raised...*the language must have a life independent of the [truth] definition.*²³

This passage is perfectly consistent with a "no language" thesis according to which passing theories are theories of nontraditional languages. For an utterance of any given "passing language" L will indeed have a meaning independently of any particular theory of L. The hearer can misinterpret the speaker; and a third party overhearing Donnellan's remark to MacKay might be mistaken about both the speaker and (intended) hearer. Mere formal constructs, on the other hand, can be thought of as merely possible passing theories. But this is not yet an objection to appealing to such constructs in semantics, as long as we do not suppose that such constructs are themselves descriptions of languages (even passing languages) that anyone speaks. For it is still an empirical question whether a given interpreter deploys a given set of axioms and rules in the course of constructing passing theories. That is, a given set of axioms and recursive rules \mathfrak{R} may describe some *part* of a given agent's semantic competence. And if we abandon Davidson's hesitancy to say that an agent actually believes a theory of meaning that correctly characterizes her competence, we can say that even formal constructs have a "life of their own" in the following sense: interpreters may, in virtue of having the prior theories they do, *believe* that such formal constructs can be used to interpret certain speakers. In the light of new linguistic evidence, I may give up the supposition that a given prior theory (or some fragment thereof) can

be used to interpret Mrs. Malaprop. But the fact that a belief is subject to revision hardly shows that it is not a belief.

B&R would object that this kind of response renders passing theories unverifiable given Davidson's holism. One might also worry that holism precludes any role for theories of "fragments" in the interpretation of actual speakers. So I conclude with some remarks intended to assuage such concerns.

2.3

One aspect of Davidson's holism is captured by his commitment to the following thesis: to know the truth conditions of an utterance u (in a language L) is to have a competence that is modelled by a theory from which a sentence providing the truth conditions of u can be derived in the familiar way. But the rules of such a theory will be recursive and instantiable by *any* syntactically appropriate lexical item or phrase (of L), not just those of which u is composed. So the very theory that interprets u will interpret infinitely many other utterances (of L).²⁴ Thus, when Davidson says that knowing a passing theory is "only" knowing how to interpret a particular (dated) utterance, I take him to be making a point about *actual* utterances: Since each (actual) utterance provides the interpreter with an opportunity to revise her most recent prior theory, each (actual) utterance requires a new passing theory; although in most cases, the "new" theory will be "constructed" by giving tacit approval to the most recent prior theory. But if Θ is a passing *theory* of a speaker S at t , then a hearer H who knows Θ knows a lot; for Θ will be correct, only in so far as Θ correctly interprets all of S 's *potential* utterances to H at t , even if only a small portion of that knowledge is applicable. Competence, I take it, goes well beyond actual performance.

B&R grant this, but still think that the move to passing theories robs Davidsonian semantics of its empirical content. For they assume that the only point of contact between such a theory and a speaker is the very utterance to be interpreted; and if this is correct, it is hard to see how a hearer could ever *confirm* the claim that a passing theory Θ interprets some utterance (where Θ is at all complex). But it is significant that, when it comes to epistemology, Davidson focusses primarily on the justifiability of our interpretations, rather than the justification of our actual interpretive methods. Given that we can say what utterances mean, he asks two questions:

What could we know that would enable us to do this? How could we come to know it?.... The second question, how we could come to have knowledge that would serve to yield interpretations, does not, of course, concern the actual history of language acquisition. It is thus a doubly hypothetical question: given a theory that would make interpretation possible, what evidence plausibly available to a potential interpreter would support the theory to a reasonable degree?²⁵

Moreover, while passing theories have a very small range of application, we have seen that they do not spring into existence *ab initio*. Indeed, it is nearly tautological to say that such theories are rational improvements upon prior theories; since a passing theory is (at least for a normal hearer) the result of "doing the intelligent thing" given the deliverance of a prior theory in some context. So as long as H's prior theory of S is--in the light of prior evidence plausibly available to a *potential* interpreter--a fairly justifiable candidate for interpreting S, H's passing theory will be a slightly better candidate in light of the new evidence. If prior theories of speakers are not at all justifiable in the light of prior evidence, passing theories are unlikely to be much better off. But even if such scepticism is motivated, the problem hardly lies in the

distinction between prior and passing theories; and the problem is hardly peculiar to Davidson. B&R think that Davidson's earlier work²⁶ provides an attractive proposal about how a theory of interpretation can (in principle) be justified; and "after 'Derangement'," we can read that work as a proposal concerning how a prior theory of S can be a reasonably well justified candidate for interpreting S. I will not here try to defend Davidson's (or any other) particular views about how the epistemological problems posed by "radical interpretation" are to be solved. For my point is only that the prior/passing theory distinction is not likely to be in tension with any attempt to show how our interpretations could be justified. On the contrary, the appeal to passing theories allows for "on the spot" improvement to whatever prior theory a potential hearer might deploy.

Of course, hearers can and often do interpret strangers (more or less) correctly upon hearing their first utterances. In such a case, H will have no direct evidence that justifies her prior theory of S. And in general, we can ask whether (and how) actual hearers are *justified* in interpreting as they do. Some version of the principle of charity may well play a role here: in the absence of evidence to the contrary, assume that others are like yourself. I have also hinted that "general frameworks" may be indispensable to interpretation; and if these are innate, the (tacit) assumption that other members of the species share my framework will be reliable. Moreover, Davidson can allow for appeals to convention, once the question is what knowledge we actually deploy in the face of which evidence. One can grant that hearers *do* make significant assumptions about linguistic conventions when interpreting, and that conventions are often what make *justified* interpretation possible, while rejecting (in the strongest terms) the idea that appeals to convention are *essential* to justifying inferences from speakers' behavior to claims about what their utterances mean. Thus, Davidson can hold the modest thesis I have attributed to

him in 'Derangement', while also holding that "convention does not help explain what is *basic* to linguistic communication." For the fact that "people tend to speak much as their neighbors do...throw[s] no light on the *essential* nature of the skills that are thus made to converge."²⁷ But as Davidson himself says,

This is not to deny the practical, as contrasted with the theoretical, importance of social conditioning....We do not have the time, patience, or opportunity to evolve a new theory of interpretation for each speaker, and what saves us is that from the moment someone unknown to us opens his mouth, we know an enormous amount about the sort of theory that will work for him....Knowledge of the conventions of language is thus a practical crutch to interpretation, a crutch we cannot in practice afford to do without--but a crutch which, under optimum conditions for communication, we can in the end throw away, and could in theory have done without from the start²⁸.

I would emphasize the role of biological endowment more than social conditioning. And this last remark about what we could have done "in theory" is a little glib. Such heavily idealized counterfactual claims are notoriously hard to defend. Davidson is, in my view, on firmer ground when arguing that knowledge of conventions is not sufficient to account for our actual interpretive ability. Moreover, the counterfactual claim is an unnecessary rhetorical flourish. For the main point is that appeal to convention can partially explain how we do (with some justification) form interpretations that are justifiable. But such appeals can provide *only* partial explanations of our actual practice; and more importantly for Davidson, they do not get to the heart of what makes an interpretation correct and/or justifiable.

Davidson is also a holist in so far as he holds that (i) we cannot ascribe a language to an agent independently of ascribing beliefs and desires to her, and (ii) the meaning of any element of a language depends in some sense on the meanings of other elements. Thus, he holds that agents can be correctly described by more than one language/belief/desire-hypothesis, given the possibility of compensating for one adjustment in the system with further adjustments elsewhere. And even if Davidson's version of indeterminacy is less worrisome than Quine's, the most radical version of holism would have the consequence that these interdependencies are so sensitive, that two agents can share a passing theory only if they share all the same beliefs. Similarly, a hearer modifying her prior theory would thereby alter her interpretations of all the speaker's potential utterances; and hearers could never apply the *same* fragmentary truth theory to two speakers (or two time-slices of the same speaker) with different beliefs.

Thus, I take radical holism to be incompatible with successful communication. For I cannot see how two truth theories could assign "distinct but similar" interpretations to all of a speaker's potential utterances: How can the proposition *p* be more similar to *q* than *r*, unless *p* and *q* are the *same* proposition?²⁹ But note that prior/passing distinctions *per se* are irrelevant here. The same difficulties would arise on the assumption that sharing prior theories suffices for communication. Indeed, appeal to passing theories does not even aggravate the difficulties for radical holism. For when a word or phrase

temporarily or locally takes over the role of some other word or phrase (as treated in a prior theory, perhaps) the entire burden of the role, with all its implications for logical relations to other words, phrases, and sentences, must be carried along by the passing theory. Someone who grasps the fact that Mrs. Malaprop means

'epithet' when she says 'epitaph' must give 'epithet' all the powers that 'epitaph' has for many other people.³⁰

Let me put this (important) point another way. Borrowing a device due to Sellars, let '*x*' refer to the class of linguistic objects that share those aspects of the "functional role" of 'x' in the relevant language/belief/desire system that are relevant to determining the meaning of 'x'. Sellars proposed to cash out the claim that (the public language items) 'red' and 'rot' (in German) mean *red*, by saying that 'red's and 'rot's--*i.e.*, utterances/inscriptions of 'red' and 'rot'--are *red*s.³¹ Adapting this device to Davidson's proposal, a holist could say that my prior theory of Mrs. Malaprop is that her 'epitaph's are *epitaph*s, while my passing theory is that her 'epitaph's are *epithet*s. If meaning is radically holistic, and assuming that no one else shares *all* my beliefs, my own 'epithet's will be the only actual *epithet*s. And I do not see how a Malapropish 'epitaph' could be almost-but-not-quite an *epithet*. But the problem has nothing to do with passing theories or nonstandard usage. The problem is to say how *anything*, including "normal" uses of 'epithet' by anyone but me, could be almost-but-not-quite an *epithet*.

Luckily, holism need not be radical. Consider an analogy. Let the "margin" of an exam E be the difference between the score on E and the class average. The margin of each exam will depend on the other exams; and given a precise measurement of the class average, a change in any one exam score will change the margin of every exam. But we often round off to the nearest whole number. And we can compute the margin of each exam on the basis of the *rounded* class average. The "rounded margin" of an exam will still depend on the other exams. But in a large class, any particular exam score could change considerably without affecting the rounded margin of any other exam. Note that an exam's rounded margin can be measured in many different

ways--in base eight, using roman numerals, *etc.*--thus yielding Davidsonian indeterminacy. The property of *having a nonzero margin* is also "anatomic" in Fodor and Lepore's sense:³² It can be instantiated only if there are at least two exams, just as the property *being a sibling* can be instantiated only if there are at least two siblings. The analogy works, I think, because rounding has the effect of mapping (infinitely) many distinct numbers within a given range onto the same number. And any plausible holistic conception of meaning must map (perhaps infinitely) many possible agents with distinct beliefs onto the same language. I take Davidson's commitment in 'Derangement' to shared meanings to entail a commitment to this kind of "coarse grained" holism. And while this is not the place to provide a theory of meaning and/or belief, there are at least two (mutually compatible) ways to achieve this coarseness of grain.

First, one can *idealize* in semantics as elsewhere. Two agents with different beliefs will be disposed to actually use a given term--say 'rabbit'--in different circumstances. But their usage may be the same, described at a suitable level of abstraction, other things being equal; just as two objects may actually fall to earth at slightly different rates, even though other things equal, both objects (qua having mass) fall at the same rate (*viz.*, 16ft/sec²).³³ Second, I can ensure that the propositional objects of our beliefs and utterances are the same by *using* the objects of my thoughts and utterances to describe yours. Consider one last time the "measurement analogy." Objects have weights in virtue of their mass and the local environment; and to ascribe a weight to an object is to relate the object to a number for purposes of comparing objects along a common metric: A weighs twice as much as B and half as much as C. Similarly, perhaps agents have beliefs, desires, and a language in virtue of their psychological makeup (and the local environment); and we ascribe these to agents, thus relating them to propositions and theories,

precisely for the purpose of comparing agents along a *common* metric: A believes (and/or says) what B believes and C denies. Ascribers would use the contents of their own mental states and utterances as the propositions that form the basis for such comparisons.³⁴ So I would say, for example, that Donald believes what I believe, but Jerry denies this. If either or both of these strategies are workable, Davidson can have holism without the radical consequences. And we can accept the claim that seems radical, but isn't: traditional semantics does not require languages in the traditional sense.³⁵

1. Dorit Bar-On and Mark Risjord, 'Is There Such a Thing as a Language,' *Canadian Journal of Philosophy*, 22 (1992) 163-90. The quote is from Donald Davidson, 'A Nice Derangement of Epitaphs,' in Ernest Lepore, ed., *Truth and Interpretation* (Oxford: Basil Blackwell 1986) 433-46, 446. I shall refer to the latter article as 'Derangement'.
2. Davidson, 446.
3. Bar-On and Risjord, 164. They discuss Bjorn Ramberg, *Donald Davidson's Philosophy of Language: An Introduction* (Oxford: Basil Blackwell 1989).
4. Bar-On and Risjord, 165-6.
5. Donald Davidson, *Inquiries into Truth and Interpretation* (Oxford: Clarendon 1984), 154 (my emphasis).
6. Donald Davidson, 'What is Present to the Mind,' in J. Brandl and W. Gombacz, eds., *The Mind of Donald Davidson* (Rodolpi 1989).
7. Ibid. Of course, we can (in principle) reduce claims about heat to claims about molecular motion; and like Davidson, I think the semantic is irreducible to, but supervenes on the nonsemantic. But I see no reason to be concerned about notational variation even if reduction seems unlikely; though indeterminacy arguments may serve to keep us from making indefensibly fine-grained semantic distinctions. (See also note 8 below.)
8. Davidson, *Inquiries*, 228 and 239. Cf. Bar-On and Risjord, 166 note 5. Of course, we can fix one way of stating the "physical" facts and still correctly describe A's linguistic competence in several ways. But this is also true of molecular motion and temperatures; so "second order" notational variation is no cause for ontological concern. Quine holds that indeterminacy is second order *underdetermination* of theory by data. (See, e.g., 'On the Reasons for Indeterminacy of Translation' *Journal of Philosophy* 67 (1970) 178-83.) But unless non-reductive supervenience is untenable, one can reject reductionism, verificationism, *and* the claim that semantic properties somehow "float free" of the physical as Quine suggests.
9. 'A speaks L' can be true, without L being a *thing* that A speaks. But while Davidson does not reify meanings or languages, there is a more important sense in which he rejects the notion of language. B&R think that defense of (2) must rest on dubious versions of the principle of charity. But while Davidson thinks that charity and (in)determinacy are connected, I do not think his acceptance of (2) hangs on this.
10. Davidson, 'Derangement', 434.
11. Davidson, 'Derangement', 437.

12. Alfred Mackay, 'Mr Donnellan and Humpty Dumpty on Referring', *The Philosophical Review*, 77 (1968), 197-202. Nonstandard first meanings should not be confused with "standard second meanings"--e.g., the intended (and recognized) intention of a speaker making an utterance in an ironic tone of voice.

13. Keith Donnellan, 'Putting Humpty Dumpty Together Again', *The Philosophical Review*, 77 (1968), 213.

14. Davidson, 'Derangement', 443 (my emphasis).

15. Davidson, 'Derangement', 444 (my emphasis).

16. Davidson, 'Derangement', 445.

17. Davidson, 'Derangement', 446.

18. Bar-On and Risjord, 177-8.

19. Jerry Fodor, *The Modularity of Mind* (Cambridge, MA: MIT Press 1983).

20. Bar-On and Risjord, 185.

21. Michael Dummett, 'A Nice Derangement of Epitaphs: Some Comments on Davidson and Hacking,' in Lepore, 459-76.

22. See Bar-On and Risjord, 185, note 5.

23. Bar-On and Risjord, 187 (their emphasis and ellipsis). The quote is from Donald Davidson, 'The Structure and Content of Truth,' *Journal of Philosophy* 87 (1990) 301.

24. The significance of this point depends on the language in question. Suppose it is fully characterized by the following theory Θ : tokens of $[_N \text{ Tom}]$ refer to Tom; tokens of $[_V \text{ runs}]$ are satisfied by all and only things that run; verbs of the form $[_V[_{V_1}\dots]]$ and $[_{V_2}\dots]$ are satisfied by all and only things that satisfy $[_{V_1}\dots]$ and $[_{V_2}\dots]$; and utterances of the form $[[[_N \dots] [_V \dots]]]$ are true iff the referent of $[_N \dots]$ satisfies $[_V \dots]$. Then Θ interprets utterances of the form $[[[_N \text{ Tom}] [_V \text{ runs}]]]$, $[[[_N \text{ Tom}] [_V [_V \text{ runs}]]]$ and $[[[_V \text{ runs}]]]$, etc.; and there are no other utterance forms to interpret. If we construe Θ as a fragment of a more complex language CL that has more lexical items and/or syntactic devices, Θ interprets infinitely many--though not all--utterances of CL. But nothing yet

follows about the correctness of such interpretations. And it is unlikely that Θ will be equivalent to any passing theory that an actual speaker would use; although the application of even such a simple theory requires enough "prior theory" to classify acoustic strings utterances of [[_N Tom] [_V flies]].

25. Davidson, *Inquiries...*, 125.

26. See especially, 'Radical Interpretation,' in *Inquiries...*, 125-39.

27. Davidson, *Inquiries...*, 280 and 278, my emphases.

28. Davidson, *Inquiries...*, 278-9.

29. Cf. Ned Block, 'Advertisement for a Semantics for Psychology,' in French *et. al.*, eds.,

Midwest Studies in Philosophy, v. 10 (Minneapolis: Univ. of Minnesota 1986). But I allow that passing theories can agree in their interpretations of many, but not all, of the potential utterances of (a time-slice of) a speaker. In this sense, understanding need not be an all-or-nothing matter; and we should not expect perfect overlap between the speaker's and hearer's passing theories.

30. Davidson, 'Derangement', 443.

31. Wilfrid Sellars, "Meaning as Functional Classification," *Synthese* 27 (1974) 417-437, 428. It is worth bearing in mind that the "functional role" of a lexical item can determine, for example, its causal relationships to the environment. And Sellars explicitly allowed for "language entry rules."

32. Jerry Fodor and Ernest Lepore, *Holism: A Shopper's Guide* (Oxford: Blackwell 1992).

33. I discuss some of the general issues raised here in 'Prima Facie Obligations, Ceteris Paribus Laws in Moral Theory,' *Ethics*, 103 (1993) 489-515.

34. What other propositions could they use? For discussion, see especially Davidson, 'What is Present to the Mind'. But see also Davidson, 'On the Very Idea of a Conceptual Scheme,' in *Inquiries...*, 183-98.

35. My thanks to Susan Dwyer and two anonymous referees for helpful comments and suggestions.