

Convention before Communication

Ernie Lepore and Matthew Stone

Davidson famously declared “there is no such thing as a language, not if a language is anything like what many philosophers and linguists have supposed” (Davidson, 1986, 107). Naturally, Davidson did not intend to deny the existence of language, or even languages, *tout court*. What he was rejecting was a certain conception of language, one that he took to enjoy widespread endorsement. Davidson characterizes this conception in terms of three components: first, the meanings that are involved in communication must be systematic; second, they must be shared; and third, they must be governed by convention (ibid. 93). Davidson rejects just the last of these components—he argues that conventions are not necessary for linguistic communication. In particular, as we shall see, Davidson rejects the idea that the audience can understand the speaker just by a straightforward application of some antecedent grammar. Rather, for Davidson, successful communication is fundamentally a matter of opportunistic inference and creative insight. Davidson’s urgent rhetoric reflects the stark implications he draws from discovering such holistic, open-ended leaps of interpretation at the heart of ordinary communication.

Davidson’s argument is ultimately an argument about how to explain successful communication. What Davidson takes to be the received view explains communication by appeal to convention. As Davidson himself puts the view (Davidson, 1994, 110), “in learning a language, a person acquires the ability to operate in accord with a precise and specifiable set of syntactic and semantic rules; verbal communication depends on speaker and bearer sharing such an ability, and it requires no more than this.” The claim that meaning is conventional can be understood in diverse ways,¹ but when philosophers invoke conventions to explain communication, they rely only on a broad commitment that conventions are regularities that interlocutors are party to that are established in advance of communication. In brief, the rules are given; in speaking and understanding, we abide by them; and this is what enables us to understand one another.²

Ultimately, Davidson rejects the received view because he concludes that communication can be successful even when the meaning a speaker wants to get across is completely particular to that occasion of use—when it is improvised by the speaker and is completely new to the interpreters. Davidson offers a range of

¹ For example, conventions might amount to certain kinds of mutual expectations, as in Lewis (1969); or they might amount to certain kinds of joint commitments, as in Gilbert (1989). They might amount to certain kinds of self-perpetuating patterns, as in Millikan (1984); or they might amount to constitutive rules that generate the phenomena in question, as in Searle (1969).

² Not all articulations of the view that meaning is conventional appeal to conventions to explain communication. For example, we could say instead that because successful communication helps *create* conventions, we inevitably find conventional meanings in communication (Armstrong, 2016). On such views, even though meaning is conventional, the conventionality of meaning presupposes rather than explains our ability to communicate.

examples in support of this idea, each intended as an intuitive basis for thinking that non-conventional meaning is crucial in language: cases of malapropisms, nonsense and neologisms. Davidson's broader writings on meaning, truth and interpretation allow us to develop an extremely precise understanding of how he takes each of these cases to play out. That makes Davidson's article an ideal target of critical investigation, even amid a diversity of current philosophy that continues to defend some notion of improvised, opportunistic or radically particularized meaning.³

In this paper, we undertake just such a critical effort. It leads us to reject Davidson's position, and to defend the old-fashioned idea that convention *is* essential to the meanings involved in communication. In working through Davidson's cases, and rethinking the intuitions surrounding them, we will try to convince you that when you look closely at them, you will *not* find any improvised *meaning*, in any sense that would undermine a crucial explanatory role for convention in communication. Nevertheless, a closer assessment of Davidson's cases reveals that not just any notion of convention will explain the communication interlocutors achieve. In particular, it does not suffice to simply describe interlocutors' mutual expectations for one another's utterances and interpretations, as in Lewis (1969); semantic conventions must specifically acknowledge the practice of naming (Kripke 1972), the division of linguistic labor (Putnam, 1975), and interlocutors' semantic deference (Burge 1979). By sharpening our understanding of convention and communication this way, we hope to draw general lessons about meaning.

Malapropisms

Davidson wants us to abandon the presumption that shared conventions underwrite communication. Instead, he argues, interlocutors regularly succeed in getting ideas across to one another *flexibly* and *inventively*, despite obvious—and not so obvious—differences in their assumptions and expectations. In these cases, Davidson suggests, speakers use words in ways that have no grounding in convention and yet they “get away with it” (Davidson 1986, 98)—that is, they are understood by their audiences as meaning what they intended to mean. In developing this understanding of communication, Davidson puts particular emphasis on the phenomenon of malapropisms—invoking one, “a nice derangement of epitaphs” as the very title of his 1986 paper. Davidson's focus is classical malapropisms,⁴ which involve the deliberate use of a word in a way that suggests the speaker's *ignorance* of its conventional meaning, as with “derangement” or “epitaphs” in Davidson's title. Even though the speaker of a malapropism doesn't use the conventional meaning to plan the utterance, and even though the hearer cannot use the conventional meaning to interpret it, Davidson suggests that communication with malapropisms generally succeeds. In

³ Philosophers might compare Davidson's improvised, opportunistic meanings with Clark's *nonce senses* (1983), Carston's *ad hoc concepts* (2002), Ludlow's *dynamic meanings* (2014), or the radical contextualism of Atlas (2005), Bach (2004), Bezuidenhout (2002), Recanati (2010), and Travis (2000).

⁴ Zwicky (1979) refers to them this way, to distinguish them from various other kinds of speech errors that are also sometimes called malapropisms in research on language production.

particular, as Davidson sees it, Mrs. Malaprop produces an utterance of “a nice derangement of epitaphs” and yet is successfully interpreted as meaning what she intends to mean, namely, a nice arrangement of epithets (Ibid. 103-4). What Mrs. Malaprop means, and what her audience understands her to mean, to be communicating, is best paraphrased as ‘a nice arrangement of epithets.’ As Davidson describes it, when confronted by her utterance, the audience ignores the prior (conventional) theory of meaning, but succeeds in reconstructing and adopting a passing (non-conventional) theory of meaning, where ‘derangement’ means *arrangement*, and ‘epitaph’ means *epithet*. Davidson goes on to base his full understanding of meaning and communication on this apparent success. There are, however, complications to this story.

To confirm that malapropisms are genuine cases of improvised communication, we need to be sure that the speaker has a particular meaning in mind, that the hearer interprets the utterance with the same meaning, and that what enables this success is the interlocutors’ nonce insights into each other’s mental states, rather than a shared regularity they have joint access to. If interlocutors do not arrive at the same meaning, they have not communicated; they have at most influenced one another. If they rely on a shared regularity to do so, their communication is not improvised; it remains an outgrowth of their antecedent knowledge of the rules of language. When we take full stock of what’s required for improvised communication in any particular case, we will argue, fundamental difficulties emerge in trying to substantiate it. Thus, even though we acknowledge Davidson’s contribution in calling attention to this sort of flexibility and inventiveness in linguistic communication, we, nevertheless, believe his description of the phenomena to be incomplete and misleading in important respects; his conclusions are correspondingly flawed.

The problems are general, but we want to elicit them by drawing on a slightly different example than Davidson’s—for which the issues, we believe, are more salient and easier to highlight. Consider (1) ((10) in Zwicky (1979, 344)):

1. What a handsome soup latrine!

Davidson assumes that in such cases the intended meaning is so obvious that it’s quite plausible to assume communication succeeds: in particular, that the speaker started by intending to communicate what we might paraphrase as ‘what a handsome soup tureen’, and then, presented the audience with what was presumed to be a suitable utterance, thereby hoping to enable the audience to recover the original idea. But, of course, we all know that *conventionally* ‘latrine’ does not mean *tureen*; indeed, we naturally imagine this usage to be new to the audience (as we hope it was to our readers), so reconstructing the intended meaning is a matter of nonce insight, and *not* of convention. Or, so a Davidsonian might assume. But let’s look at the phenomenon more closely.

To describe the case in full, we need to be precise about what the speaker means by ‘soup latrine’ and how the hearer interprets ‘soup latrine.’ Apparently, the utterance is

a comment on an elegant bowl that the interlocutors are jointly confronted with. It's unproblematic that that bowl, at least, will fall under both interlocutors' interpretation of 'soup latrine.' But what else will? What is the broader class of which this particular item is a handsome exemplar?

Suppose the hearer approaches this case just with what Davidson calls a *prior* theory, appealing to a conventional satisfaction principle like (2),

2. (x)(x satisfies 'latrine' iff x is a latrine)

On this interpretation, the utterance marks the present bowl as a handsome exemplar of a category that also extends to basins like the one below (imagined full of soup):



James Emery (flickr) at Fort Benning, GA

That can't be right. Presumably, the speaker had no intention of comparing this elegant bowl to a bathroom fixture. Instead, the speaker wanted to say that this was a handsome example of the special bowls used for soups at formal dinners, a category that invites comparison with items such as the two below:



User Daderot (wikimedia) Gardiner Museum, Toronto User Daderot (wikimedia) UW Art Museum

In sum, interpreters conclude that the speaker made an error, failing to say, "What a handsome soup tureen!" Davidson's summary is to describe this inference in terms of a *passing* theory that speakers and hearers improvise in understanding the original utterance. In place of the conventional (2), the passing theory features the nonce principle in (3):

3. (x)(x satisfies 'latrine' iff x is a tureen)

If the speaker intends the utterance to be interpreted according to this principle, and the hearer does interpret it that way, then the malapropism will be no obstacle to successful communication.

Davidson regards it as self-evident that the principles of a passing theory, like (3), do not need to have any basis in convention. But if it is not convention, what is it that does guide the hearer's inference in deriving the principles of the passing theory? On what basis could one infer that (3) is a correct attribution of meaning to the speaker? Answering those questions requires us to link Davidson's account of malapropisms to his broader theory of interpretation.

Davidson believes that what's crucial to meaning and communication is being understood as one intends to be understood. Thus, he proposes that we start from the speaker's communicative intentions, which he takes to be a species of non-natural meaning (Davidson, 1986, 93), close to Grice's fundamental notion of utterer's meaning or speaker meaning. In Davidson's narrative summary of successful (improvised) communication, then, "the speaker says something with the intention that it will be interpreted in a certain way, and the expectation that it will be so interpreted... [t]he interpreter adjusts his theory so that it yields the speaker's intended interpretation...What must be shared is the interpreter's and the speaker's understanding of the speaker's words" (Davidson 1986).⁵

Davidson had long argued, however, that a particular standpoint is required for subjects that aim to achieve shared understanding in the absence of shared conventions. His approach was the theory of Radical Interpretation. A radical interpreter, just like a child learning a first language, cannot presume an existing correspondence between a speaker's meanings and her own, nor can she presume that the speaker is using words with their conventional meanings: the aim is to characterize the speaker's understanding of the words, regardless of whether it happens to match her own understanding of those words or widespread practices in their community. As Davidson says, "the theory is supposed to supply an understanding of particular utterances that is not given in advance, so the ultimate evidence for the theory cannot be correct sample interpretations" (Davidson, 1973). Instead, radical interpretation "is accomplished by assigning truth conditions to alien sentences that make native speakers right as often as plausibly possible, according, of course, to our view of what is right" (1973). This process of "optimizing agreement," then, is what justifies the hearer in characterizing a speaker's communicative intention in terms of the hearer's own understanding, and thus what enables the

⁵ Elsewhere he writes, the "intention to be taken to mean what one wants to be taken to mean is ... the only aim that is common to all verbal behavior" (Davidson, 1994). In other words, all successful instances of communication involve the satisfaction of some communicative intention, and all conveyed (non-natural) meaning must be specified in this corresponding intention.

recognition of a shared understanding of an utterance without appeal to shared conventions.

Spelling out what must be involved, for Davidson, in constructing a passing theory, by putting these ideas together, leads to a picture something like this: when you hear, 'What a handsome soup latrine!', you adopt a passing theory including principle (3), *because* doing so optimizes your agreement with what the speaker would describe as a 'soup latrine'.

There are many reasons to be worried about this sort of reconstruction of the hearer's understanding; for example, it seems to us *prima facie* highly implausible that someone who would call a gorgeous piece of Wedgwood china a 'soup latrine' has the relevant judgments about what is, and what is not, a soup tureen in sundry boundary cases. But what's much more worrisome is that even if the hearer has a very specific idea of what the speaker thinks these objects are, it seems unlikely that the specific choice of *tureen* from among the many candidate options compatible with the data is going to be the unique one that somehow most optimizes agreement between speaker and hearer.

To get a feel for the sort of problem we have in mind, we need only survey the internet. What *is* the full range of alternative objects with which the speaker of (1) invites the present bowl to be compared? Would the following object count?



User Daderot (wikimedia) Indianapolis Art Museum

Perhaps the speaker has in mind a broader category of sealed containers on pedestals, used for various hot liquids as well as for other purposes: an understanding we might better render with the English word *urn*. So should we continue to interpret using (3) or should we instead adopt (4)?

4. (x)(x satisfies 'latrine' if and only if x is an urn)

What about this ancient Vietnamese soup tureen? Is this kind of alternative included in the comparisons the speaker intended?



User Daderot (wikimedia) Vietnam Nat. Art Museum

Perhaps the speaker has in mind a category of ceramic containers with lids, without much regard for their shape, an understanding we might prefer to capture with the English word *crock*. In this case, the interpreter might wonder whether (5) isn't a better improvised passing hypothesis:

5. (x)(x satisfies 'latrine' if and only if X is a crock)

Of course, given the larger picture, maybe it's a mistake to limit ourselves to the meanings of actual words. Wouldn't we have to exploit these same resources in order to interpret novel words (compare Davidson's "foison")? Maybe, the speaker has a concept in mind that does not match any English word, such as a "ceramic vessel sculpted to resemble its intended contents." This turns out to be a kind of thing, but, surprisingly, it seems not yet to have been lexicalized; for the purposes of this discussion, we will opt to baptize objects of this sort as "figureens." Perhaps the speaker specifically intended that the target of the exclamation in (1) only be compared with such other "figureens" as these:



User Daderot (wikimedia) at the Winterthur

In China, one is expected to serve turtle soup in a turtle "figureen."



David Jackson (wikimedia) at the V&A

While in England, one is expected to serve cauliflower soup in cauliflower “figureen.”



LACMA public images

We suppose that it will be obvious to you what one is supposed to serve with this magnificent figureen above from the LA County Museum. In each of these three presentations instead of (3) we might just as well come up with (6):

8. (x)(x satisfies ‘latrine’ if and only if X is a figureen)

Clearly, then, what each of these cases reflects is that there are many useful categories in the vicinity—indeed, astronomically many; and so, given these data, there are many plausible passing theories we might entertain when confronted with an utterance of, “What a handsome soup latrine!” *if* what we were trying to do is merely capture the understanding of the situation that the speaker is intending to get across.

Davidson offers no story whatsoever about how to determine which of these different improvised principles of interpretation is the correct one. Since there are so many equally good options, the story about optimizing agreement in the face of the speaker’s judgments, or dispositions to accept utterances, seems highly implausible.⁶

⁶ Davidson seems sensitive to this worry in some passages; for example, he writes, “... what matters, the point of language or speech or whatever you want to call it, is communication, getting across to

We want to emphasize how important theoretically it is to appreciate the diversity of candidate meanings, all equally good, in situations like this, because, as we have argued extensively (Lepore and Stone 2015), *communication is a coordination problem* (Lewis 1969), where alternative strategies, all equally good, are available, and so, success in communication must result from our having agreed on one option over the others *arbitrarily*. But, in order to achieve this feat reliably, something must have already been in place to make coordination possible.⁷ And it is in this light that we conclude that Radical Interpretation does *not* seem to be operative in guiding us to whatever interpretation we ultimately decide upon.⁸

In fact, the interpretation we actually give the utterance, using *tureen*, seems very much bound up with the explanation we imagine for the how the utterance must have come about. After all, we do kind of understand how people come to make mistakes of this sort, even if it's very fraught to say in what sense they count as mistakes. Indeed, we assume most readers are chomping at the bit to object that alternative meaning proposals like *urn*, *crock*, and *figureen* fail to respond to one of the more salient aspects of the scenario: namely, that "latrine" and "tureen" sound alike. But the only reason we think this datum could be relevant, if it is at all, in the present context has to do with a distinction we are ready to explicitly introduce.

What seems to be going on in Zwicky's scenario is that the speaker has encountered an English word, but, like most of us, has incomplete knowledge about it. In this particular case, the speaker believes the word describes a sort of big fancy bowl associated with soup but also mistakingly believes that it is pronounced something like [luh TREEN]. And what the speaker intends to do is to utter *this very word*. We insist on taking the idea that the speaker is uttering this particular word very seriously; in doing so, we draw heavily on the metaphysics of words more generally, following Hawthorne and Lepore (2011).

someone else what you have in mind by means of words that they interpret (understand) as you want them to. (...) The intention to be taken to mean what one wants to be taken to mean is, it seems to me, 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" (1994b, 120). See also Fodor and Lepore (1993), and Lepore and Ludwig (2006).

⁷ Davidson himself in some passages seems to acknowledge as much; he writes, "[t]he best the speaker can do is to be interpretable, that is, to use a finite supply of distinguishable sounds applied consistently to objects and situations he believes are apparent to his hearer" (Davidson, 1984, 13; emphasis ours). Regularity and repetition are vital, and once we grant that, it's relatively easy to see the conditions for coordination are satisfied.

⁸ Moreover, we can see that the problem of ambiguity is not specific to the Davidsonian project of Radical Interpretation. The problem comes simply from assuming that the speaker's communicative intention involves one specific understanding of a category from the speaker's experience, derived by creative inference, while the hearer's interpretation also involves a specific understanding from the hearer's experience, derived by creative inference. How could two such processes fortuitously arrive at the same outcome from among all the possibilities? In short, it seems to be part and parcel of the intuition that speakers and hearers do communicate, that there is some regularity that aligns the hearer's interpretation to the speaker's. We therefore suspect that any philosophical approach to malapropisms as improvised communication will be vulnerable to criticisms analogous to those we have just leveled against Davidson.

What Hawthorne and Lepore recognize is that words must be distinguished from their articulations; once this is recognized, we realize that the same word can be spoken sometimes, written down others. And indeed, the same word can be spelled in multiple different ways and pronounced in multiple different ways. Our intuitive notion of a word thus builds in the possibility of diverse articulations. “The same word can be spoken or written, indeed spelled in different ways (gaol, jail), or pronounced in different ways (tomayto, tomahto)” (Hawthorne and Lepore 2011). To repeat, expressions must be distinguished from articulations.

Once this basic distinction between expression and articulation has been grasped, it should be obvious that speakers can be *mistaken* about the articulations of words they know and use. Misspellings provide infamous examples, as in “I was not effected.” When someone misspells the word “affect” like this, we do not conclude they are using the “effect” word; rather, we infer they are using the “affect” word, but are spelling it the wrong way. They made a mistake. And so, in a sense, our view (contrary to Davidson) about interpreting malapropisms is that they should be seen as just extravagant examples of the same thing. We might go so far as to say that when a speaker says, “What a handsome soup latrine”, what they mean is what a handsome soup tureen, and they do for a totally boring reason: they actually used the word “tureen”. They just happened to mispronounce it, because they were wrong about its articulation.

Our approach to malapropisms, like Zwicky’s (1979), is very much tied to the cognitive processes involved. This is important in assessing potential criticisms of our view. In what cases does it make sense to say that a speaker is wrong about the articulation of the word they have used? Our explanation applies directly only to classical malapropisms, which corrupt actual words, either through errors of phonological and morphological analysis on the one hand, or through problems in storage and retrieval in the mental lexicon on the other. Such processes need not result in errors that substitute one word for another, however—Zwicky also cites examples, such as ‘assimulated’, ‘laxadaisical’, and ‘vanishment’, which result in apparent lexical innovations. In such cases, it’s particularly transparent that speakers are mistaken about the articulations of the words involved.

We observe, however, that there are many other kinds of possible error in the vicinity. In our opinion, each requires its own philosophical response. To start, speakers often substitute one word for another not deliberately, as in (1), but simply inadvertently—this leads to another kind of malapropism that Zwicky (1979) discusses. Reverend Spooner accuses his pupils not of having ‘wasted the whole term’, but of having ‘tasted the whole worm’. A mathematician finds that ‘two vectors are equivocal’, not ‘equivalent’. Such cases reflect a different kind of disruption between intention and articulation. One suspects, in fact, that the speaker in these cases has a completely ordinary communicative intention, in which the plan is to use the conventional articulation of the words. The execution of this intention is unfortunately disrupted by the performance limitations of the human mind. Action slips of all kinds are

painfully familiar in human action (Norman 1981)—think of throwing the dirty shirt in the toilet rather than the laundry basket. It's precisely because of such errors that we engage with others' actions by attempting to recognize their intentions in the first place. Performance errors should not challenge any theory of meaning.

Two further cases are worth commenting on, in addition, since they will help us to situate malapropisms within the broader purview of apparently improvised language. On the one hand, there are *bogus words* that speakers might use without any grounding in established usage or their own experience, like the infuriating 'supercalifragilisticexpialidocious' of Disney's *Mary Poppins*.⁹ Utterances of such expressions can certainly influence an audience—as the song goes, if you say it loud enough you'll always sound precocious—but we needn't assume that they do so as a matter of meaning. In the next section, we look more closely at nonsense, emphasizing the resources a philosopher has for explaining the effects of nonsense utterances without appeal to talk of meaning or communication.

On the other hand, there are also cases where a speaker invents a word inadvertently, as it were, because they assume this *must* be the word for what they have in mind. You can get the flavor for this possibility from another famous fictional example: when dentist Tim Whatley of the TV show *Seinfeld* passes on Elaine's label maker as a present for Jerry, Elaine inaugurates the term *regifting* to describe the situation. Elaine isn't wrong about the meaning or articulation of any word, and she hasn't offered mere nonsense. But her word is new, as is its meaning. These too are not classical malapropisms, and it seems best to defer them until we can present a broader investigation of neologisms.

Nonsense

Davidson explains nonsense with an extension of his approach to malapropisms; he writes, in malapropisms "the hearer has no trouble understanding the speaker in the way that he intends" (1986, 434). He goes on to say that in a range of cases it doesn't matter whether the speaker uses a real word, which he supports by registering this radical claim: "most of 'The Jabberwock' is intelligible on first hearing" (1986, 434). We personally never had the patience to get deep enough into the poem to savour its intelligibility, so when we scrolled through its Wikipedia entry, we were eventually able to arrive at a stanza that we hadn't read before, namely:

And, as in uffish thought he stood,
The Jabberwock, with eyes of flame,
Came whiffling through the tulgey wood,
And burbled as it came!

⁹ Some famous examples from Monty Python include 'splunge', from the 20th Century Vole skit, and 'zalling', from the North Minehead By-election skit. 'Zalling', asks Michael Palin, "Is it even a word? If it is, what does it mean? If it isn't, what does it mean?"

from *Jabberwocky*, Lewis Carroll (Charles Ludwig Dodgson) 1871

The picture that this reading prompted for us unfortunately bore little resemblance to the scene depicted in the original illustration of the poem.



Original illustration by John Tenniel, 1871.

Still, Davidson invites us to think about interpreting the poem in the same way as always: “The speaker says something with the intention that it will be interpreted in a certain way, and the expectation that it will be so interpreted... the interpreter adjusts his theory so that it yields the speaker’s intended interpretation” (1986, 434). And, so, the interpreter derives a passing theory with improvised principle (7):

7. (x)(x satisfies ‘uffish’ if and only if x is uffish)

Keeping with the doctrine of Radical Interpretation, this derivation is supposed to result from optimizing agreement with how Carroll, or the purported poet, would characterize uffishness. But, of course, interpreters don’t retrieve this sort of interpretation of the poem.¹⁰ Indeed, the best that Alice can say of the poem is, “Somehow it seems to fill my head with ideas—only I don’t exactly know what they are!”

We think what Carroll himself says about uffishness is particularly diagnostic. He writes in a letter to Maud Standen: “It seemed to suggest a state of mind when the voice is gruffish, the manner roughish, and the temper huffish”. What’s obvious from this description is, first, that he doesn’t have a very specific idea of what uffishness is, but, more importantly, that he is inviting us to just listen to the sound of the word, and then, let that sound take us wherever we go. This is characteristic of poetry, and shouldn’t come as a surprise, especially, to anyone in the Davidsonian tradition.

Davidson emphasized in much of his work that language use was not just communication. Indeed, we began our collaboration several years ago (Lepore and

¹⁰ Both Wiggins (1971, n.7) and Fodor (1975, 121) make fun of just this sort of interpretation of *Jabberwocky*. See also Reimer (2004) and Armstrong (2016) for further criticism of Davidson on *Jabberwocky*.

Stone 2010, 2014) in commenting on, and exploring, Davidson's paper "What metaphors mean," where he writes: "A joke or dream or metaphor can, like a picture or a bump on the head, make us appreciate some fact—but not by standing for, or expressing, the fact" (Davidson 1978, 262). "The common error is to fasten onto the contents of the thoughts a metaphor provokes and to read these contents into the metaphor itself" (Davidson 1978, 245). This is just not how metaphor works, according to Davidson. And, it's not how poetry works either, according to us.

The way we like to put it is, "Poems are about their own articulation. They are about their own articulation, because poems ask to be understood poetically, so that the interpreter looks at their articulation for insights into their meaning" (Lepore and Stone, 2016; cf. also, Lepore 2009). And, of course, this is exactly what Carroll is directing us to do in order to understand "uffish". The best we can say about how to imagine someone acting in an uffish manner is to allow our imaginations to be prompted by the sound of the word itself. An active metaphor, like the vocabulary of Carroll's evocative nonsense, is best understood as an *invitation* for the audience to develop the imagery of the metaphor by characteristic kinds of thinking. A metaphor does not create public propositional content. We see our view on interpreting poetry as merely a development of Davidson's view of interpreting metaphor.

In short, nonsense *is* really just nonsense. Postulating a passing theory for Jabberwocky with an interpretive principle like (7) is implausible, and worse, does nothing at all to explain the interpretation of the poem. Jabberwocky is a particularly tough example because its sonic imagery is so compelling that even many of the nonsense words that Carroll had no meaning for himself eventually became part of the English language through the imaginative enrichments of subsequent writers, as in "chortling" and "galumphing". The place of the poem in our culture allows us to forget how important it is that nonsense art really is meaningless. But it's often crucial to the point of nonsense art that hearers appreciate its meaninglessness.¹¹ Few are better masters of this art than John Cleese. A classic online video shows Cleese in a white lab coat pointing to a model of the human brain while delivering a rapid-fire

¹¹ The vocabulary of fiction often leads to borderline cases. We think such cases confirm our perspective on meaning, rather than Davidson's. How ready one is to attribute meaning depends not only on judgments about the intentions of the author, but also on judgments about the collective authority of the community of readers who take the world of the text seriously. Take J. K. Rowling's fantastic game 'quidditch' from her Harry Potter series. Though the books describe some eventful quidditch matches, the details leave many basic questions about the sport unanswered. The readers took over. There's now an International Quidditch Association, publishing official rules; our own university has a competitive quidditch team (apparently, a rather good one). There's a case to be made that Rowling's game, played as it is by wizards flying on magic broomsticks, remains a bit of nonsense—nothing more than an open-ended, incomplete prompt for the imagination, like Carroll's verse. But there's also a case to be made that Rowling's fans have made her game a thing, and that now Hogwarts students and Rutgers students alike play quidditch. The difference, it seems to us, is not answered by assessing Rowling's original, or even enlightened, conception of quidditch, but by assessing the continuity and status of the conventions her fans have established in engaging with her imagined world.

barrage of scientific-sounding terminology, intermixed with half-mumbled explanations of anatomy and physiology and the occasional ludicrous equation.

<http://www.johncleese podcast.co.uk/cleeseblog/2008/08/john-cleese-podcast-33-brain-explained.html>

At one point in the lecture, for example, the key topic appears to be the importance of ‘resumptory frictation’ in brain activity. Of course, there is no thing—but more to the point, even if there were, we wouldn’t know. The superficial similarity of Cleese’s speech to real neuroscience (compare ‘reinitiated inhibition’) dramatizes the abstruseness of technical talk, and the inaccessibility and often inappropriateness of scientific explanations. You certainly cannot appreciate Cleese’s performance by assuming he has his own neuroscience we should try to Radically Interpret. In short, we believe Davidson should have left his discussion of Jabberwocky out of the paper, because his own philosophy offers a rich way of thinking about those kinds of examples without bringing in a passing theory, or addressing more general considerations about how it is that speakers and hearers manage to get their ideas across flexibly in real situations. We turn now to neologisms.

Neologisms

Davidson (1986) considers neologisms only in passing. Here’s what he has to say about them. He starts by describing uses of novel names in terms of his passing theories: anytime you have a new name it “requires a change in the interpreter’s theory, and therefore a change in our description of his understanding of the speaker” (Davidson 1986, 431). Concretely, you encounter a new name for a new individual. That’s something that wasn’t part of your prior theory, but you take it on—the link between name and referent becomes part of the passing theory that captures your understanding the utterance. Davidson then quickly generalizes from new names to other kinds of lexical innovation: “sheer invention is ... possible, and we can be as good as interpreting it ... as we are at interpreting the errors or twists of substitution” (Davidson, 1986, 441) that we saw with classical malapropisms. That’s as much as Davidson says—his focus remains situations where the speaker and hearer already share the relevant concepts, as in malapropisms.

What’s going in these cases, as in our prior cases, is supposed to be an instance of improvisational meaning. For invented words, the radical interpreter aims to construct a passing theory that optimizes agreement with a speaker by recognizing the category that the speaker is trying to name. Once again, as Davidson says, “the speaker says something with the intention that it will be interpreted in a certain way, and the expectation that it will be so interpreted... the interpreter adjusts his theory so that it yields the speaker’s intended interpretation” (Davidson 1986, 434).

Neologisms may be incidental to Davidson, but they constitute a central challenge to anyone who would claim—as we do—that meaning must be conventional. How could the novel meaning of a novel word be a convention? Nobody has used the word

before, so it seems impossible that its meaning could be a matter of established precedent, shared knowledge, or mutual agreement among interlocutors. The problem is particularly acute because of how philosophers have tended to parse their intuitions about the meaning of neologisms: the received suggestion is that the meaning of a neologism just is whatever understanding its inventor intended.¹² So understood, it's simply part of the phenomenon of lexical innovation that speakers can freely associate new words with meanings of their choice. This seems about the clearest case of improvised meaning there could be.

As with malapropisms, we believe this description of the phenomena to be incomplete and misleading in important respects; philosophers' conclusions are correspondingly flawed. So we will begin by looking more closely at one neologism in particular, once again to make sure that the speaker has a particular meaning in mind, that the hearer interprets the utterance with the same meaning, and that what enables this success is the interlocutors' nonce insights into each other's mental states, rather than a shared regularity they have joint access to. The conclusions we draw, however, are quite different from the ones Davidson and neo-Davidsonians like Armstrong (2016) would have us draw. As with malapropisms, we believe the results undermine the attribution of non-conventional meanings to neologisms—although, here too, we believe that philosophers have made an important contribution in calling attention to the flexibility and inventiveness seen in such cases.

We will consider the recent and vivid coinage, "bromance", first used by Dave Carnie in the 1990s in an article in his street skating magazine *Big Brother* in recognition of the fact that skaters who spend a lot of time together related to one another with a mix of trust, intimacy and affection that felt new, valuable and, for straight males, slightly transgressive.

"Bromance," of course, is a neologism that succeeded. It's not just a description of skateboarders anymore. As popular culture would say it, bromance is "a thing."¹³ Use of the word seems to reflect a change in how men relate to each other as friends across a wider range of communities and settings: describing a relationship as a bromance acknowledges the reality of (potentially surprisingly) intense emotions and affections in certain kinds of men's friendships. This is possible, of course, because attributions of a bromance between any two men is either true or false, and so,

¹² We can cite Armstrong (2016) as a particularly clear example of this principle. His account of neologisms presumes that what makes the communication successful in these cases is that hearers can identify speakers' intentions, "which requires audience members to identify the same (or some sufficiently similar) content to the one the speaker intended to convey". Of course, he notes, "...audience members cannot look inside a speaker's head and identify her intentions", so with neologisms, Armstrong suggests, real-world salience fills the gap, allowing audience members to pinpoint a meaning that "stands out', or is sufficiently distinguished, to allow audience members to identify that meaning as the one the speaker intended."

¹³ See Alexander Stern, "Is that even a thing?" <https://opinionator.blogs.nytimes.com/2016/04/16/is-that-even-a-thing/>

“bromance” picks out a kind of relationship that turned out to have a broader application than *Carnie* recognized.

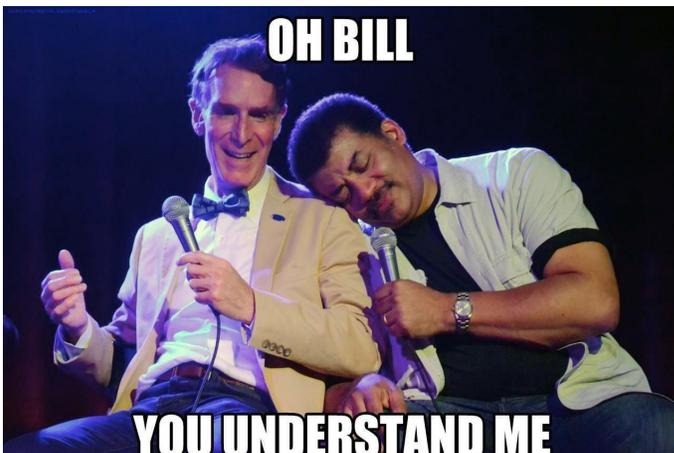
And so, here is a typical bromance meme, from the TV show *Scrubs*:



Donald Faison and Zach Braff in *Scrubs*

As is evident, there’s an intensity to the engagement these men have with each other that is not part of the traditional picture of masculinity, but is also not how skateboarders relate to each other either.

We needn’t view bromance just as a description of existing relationships: thinking of bromance as a possibility might open up space for men to relate to one another in new ways. And so, in the wake of climate change denial, we have learned that people can bond in a bromance even over the fragile attractions of science:



Bill Nye and Neil DeGrasse Tyson

Or, conversely, we can come to use bromance as a foil that helps us to recognize the “toxic” elements of traditional masculinity, as in the ironically incongruous “bromance” of this viral photoshop image:



Vladimir Putin & Donald Trump.

What these various memes show us is that our community has come a long way in *understanding* what bromance is, as the word persists. And, in particular, the fact that Carnie’s word has been adopted into common currency is a clear indication that these new relationships were part of a broader phenomenon—one whose characteristics, extent and implications were a matter of general interest. The obvious conclusion to draw is that Carnie’s own take on these relationships—on bromance—turned out to be neither exhaustive nor authoritative—despite his having coined the term.

Different philosophers take different approaches to understanding and placing the meaning of terms through such different trajectories. But a reasonable view, one that attracts us, is that there is but a single meaning at play: “bromance” means bromance, in a traditional Tarskian way. That meaning is passed from one speaker to another like links in a chain, and it is the whole network of speakers that work together to lock onto its meaning. In other words, once the community acknowledges that bromance is a thing, then the community as a whole is implicated in settling what bromance is, in triangulating its nature and delimitations, through processes of discovery and “litigation” (Ludlow 2014), in a way that accords with their experiences and values. The original speaker, who understood and named the phenomenon, just offers the starting point. In fact, the process can, and often does, lead retrospectively to an understanding of the meaning that could not have been attributed correctly to the original speaker. If we want to figure out what bromance is, then, we do not just look inside the head of the first speaker; rather, we engage with the world and our own understanding. In short, despite what Davidson seems to suggest about “glory” (Davidson, 1986, 86), even a word that you have invented and defined need not, in Humpty Dumpty’s famous phrase, “mean just what you choose it to mean—neither more nor less.”

The view we are suggesting is one that is familiar from philosophical discussions of externalism in semantics (Kripke (1972), Putnam (1975), Burge (1979), etc.). A word means what it means in someone’s mouth, on this model, in virtue of its being related in the right way to whatever falls under it, and *not* straightforwardly in virtue of the

person's knowledge of those objects, or through a disposition to call things with a certain expression. Normally, the one and only way to acquire a concept is by appropriate exposure to, and participation in, a community that names things with this word (either through historical, causal or social connections) and whose aggregate behavior can be understood to establish the information, standards and boundaries that settle membership in the category (to the extent that it is determined).

Externalism, however, undermines the idea that audiences interpret neologisms in virtue of their insight into the speaker's mental states, rather than through a regularity they and the speaker have joint access to. For an externalist, Carnie's utterance of "bromance" communicates successfully as long as his audience interprets his utterance of "bromance" with a concept that picks out bromance. However, for an externalist, such a concept can *only* be acquired if it is linked in the appropriate way to Carnie's act of naming bromances "bromances." So there is already a regularity in place, of a certain kind, for an externalist, inaugurated simply by giving something a new name. And the audience exploits that regularity in interpretation. They choose to interpret a new name with a new concept that shows its debt to the practices of the wider community: in effect, they think "I intend to use bromance to pick out whatever the speaker from whom I learned it intended to pick out with their use." Such chains *sinew* their way right back to its introduction (either through historical, causal or social chains).

In short, successful communication doesn't depend on Carnie's audience recognizing or mirroring Carnie's *understanding* of bromance—which is good, because the possibilities are too open-ended to expect such a fortuitous alignment. Instead, successful communication depends simply on the audience's recognizing and exploiting Carnie's *naming* of bromance. As Kripke taught us long ago, language gives us a lot of freedom to give new things new names. But a speaker who coins a new word is *doing* something that *adds* to the language, and not just verbalizing a personal understanding. The connection they have established, which will ramify and come to be cemented in a network of speakers, is what links the new word to its content, *not* their intention. On this understanding, creating, recognizing and exploiting a new name is no challenge to the idea that meaning is conventional: rather, it offers just another familiar illustration of the way we understand one another, in speaking and understanding, by drawing on rules that have been established in specific ways.

To sum up: our description of straightforward neologisms aligns neatly with our description of malapropisms. Both are cases where speakers use words in apparently novel ways. However, despite appearances, both are cases where meaning depends on an antecedent regularity that enables interlocutors to coordinate on a meaning. Neither is a case where meaning arises purely through the speaker's intention to evoke a particular concept. Neither is a case where coordination on meaning succeeds purely based on interlocutors' accurate inference about one another's mental states.

And so, we have thus completed our critical reaction to the examples Davidson explores in “Nice Derangement...” in his effort to establish the existence of improvised non-conventional meaning. And we now have a way of seeing how natural it is to resist the idea that any of these cases involves a speaker creating improvised meaning, and a hearer just recognizing, through processes grounded in, or motivated by, the picture of Radical Interpretation, a match between their own understanding and the speaker’s understanding of a situation. We will end by returning to more general questions of bridging philosophical insights about meaning with cognitive questions about the language faculty.

Broader Lessons

Malapropisms and neologisms may look like they involve improvised meanings, but nevertheless, we have argued that speakers’ meanings in such cases are, in fact, established in advance in key respects. Counter to first appearances, malapropisms exploit the links between words and their ordinary meanings. To get clear on how this works, however, we must be precise about what words speakers use, and appreciate the status of words as abstract types, distinct from both their meanings and their articulations. Counter to first appearances, neologisms are ordinary additions to the language, not just spontaneous enrichments of interlocutors’ passing theories. To get clear on how this works, however, we must be precise about the way meanings are propagated through a community, and appreciate the corresponding distinction between giving something a name and using that name.

In all cases, we have discovered that communication is based on regularities linking form and meaning. These regularities are conventional in an intuitive sense: they are social constructs that involve an arbitrary choice, selected by human speakers from among a set of equally workable alternatives. These features go hand in hand. It would be surprising if we could coordinate on meaning if there weren’t something in place in advance to do it, given that there many alternative strategies, all equally good, yet we still do somehow manage to agree on one in particular. Coordination represents a central challenge to accounts like Davidson’s involving improvised meaning. Following Armstrong (2016), you might try to offer a theory of salience, or of focal points, or of other principles that might enable us to coordinate without antecedent knowledge, just through our nature as people, but, as our ‘latrine’ and ‘bromance’ examples bring home, we’re skeptical that such explanations could work. Davidson certainly doesn’t have one. To the extent that these really are coordination problems with multiple alternatives, Davidson has failed to grapple with the difficulty of orchestrating matching strategies across agents.

Although meaning is conventional on our view, we want to emphasize again that speakers often use language for other purposes besides communicating meanings. There are many other ways to get our points across, including not only metaphor and poetry, but also irony, sarcasm, humor, hinting—we catalogue a range of options in Part 3 of Lepore and Stone (2015). However, following Davidson (1978), we don’t think of these effects in terms of meaning, truth, interpretation and communication.

One of the advantages of this Davidsonian perspective is that all of these resources are available. So, what, then, is our take-home lesson? It is that, as semanticists, we should *reject* Davidson's explanations in "Nice Derangement." There is *no* reason to believe in Davidson's passing theories, i.e., in improvised meanings!

Although we claim that communication rests on conventional meaning, it is not easy to articulate a precise definition of convention that straightforwardly fits the regularities we have considered. For example, one influential way of thinking of conventions of meaning is due to Lewis (1969). He defines a convention as a strategy for solving a coordination problem whose regular use is common knowledge among a group of agents. He applies the definition to language by explicating meaning in terms of conventions of truthfulness and trust. In place of semantic principles like:

'W' is true if and only if W

Lewis proposes that language users share a convention of using 'W' if and only if W (truth), and of reacting to utterances of 'W' the way one would if W (trust).

In offering such definitions, Lewis's goal is to account for semantic conventions in terms of quite general features of agents' abilities to act together. That very generality, however, makes it difficult to accommodate the insights we have offered in this paper. The most glaring problem is that Lewis expects conventions to ground out in the common knowledge of agents. But externalism says that what matters is social, causal or historical facts, not what agents know. When you pronounce 'tureen' as [luh TREEN], it's your *acquaintance* with *that word* that enables you to use it, not your knowledge. When you wonder about a bromance, it's your *acquaintance* with bromance, not your knowledge, that gives your question its meaning.

Perhaps one could find a general definition of convention that extends to conventions with this kind of grounding.¹⁴ However, our suspicion is that the most promising approach is to sidestep the problem. Our preference is to describe the competence that allows interlocutors to coordinate on meanings directly, by appeal to suitable models of the language faculty. We gesture at some of the considerations that should inform such an account in Chaps 13 & 14 of Lepore and Stone (2015), looking at how linguistic competence might provide for such features of meaning in conversation as causal-historical referential chains, the dynamic conversational scoreboard, and possibilities for clarification and argument.

Such a project in fact dovetails with the broader concern of Davidson's "Nice Derangement", which is to reconcile philosophical approaches to language with a contemporary understanding of the language faculty. Researchers in the Chomskyan tradition regard the language faculty as an internal psychological process to be characterized individually, and, maybe even, solipsistically, in terms of computations

¹⁴ Millikan's (1984, 1998) understanding of convention in terms of self-perpetuating patterns, for example, is one of the few proposals that doesn't ground convention in agents' mental states.

and representations. This is a natural match for the subtle and idiosyncratic ways grammars seem to vary across people. As Davidson points out, in principles-and-parameters approaches to syntax, even one slight change in settings might lead to different grammars that different people acquire without any evidence: the difference may lead to potential misinterpretations further down the road when that one sentence occurs that tweaks the difference between these two idiolects, as it were. This is already enough to undermine the idea that familiar languages like “English” have any psychological reality.

Linguistic methodology normally avoids investigating whether language offers ways of making meanings public and sharing them across people—but these hidden differences in grammatical parameters do not in themselves foreclose the possibility. For example, we are keen to describe words both as linguistic objects that are represented, acquired and constrained by the language faculty, and as types that can be shared across a community through causal, historical or conventional links. However, the fact that the conventions around these words are much more complicated than anyone has spelled out in full detail underscores the work that still remains to be done in reconciling Chomskyan approaches to the language faculty with philosophical understandings of meaning and communication.

Even if interlocutors share words, of course, they need not share a language. In particular, we do not think you need a passing theory or Radical Interpretation to be worried about the fact that the network of words that any one speaker uses is idiosyncratic and connects that speaker to lots of different communities and lots of different networks of individuals and causal relations and referential potential. Understanding, in light of this complexity, how one person manages to pinpoint the specific meanings they share with another is a complicated project that bridges philosophy and cognitive science, and doesn’t necessarily lead to the simple answer that interlocutors just look up the words they hear in their mental lexicon.

We are not here to say that there are languages, but just to say if you want to show that there are no languages—and there are real worries about whether there could be—going forward we should argue in a different way than Davidson did.

Bibliography

- Atlas, J. D. 2005. Logic, meaning, and conversation: semantic underdeterminacy, implicature, and their interface. Oxford: Oxford University Press.
- Armstrong, J. 2016. The Problem of Lexical Innovation. *Linguistics and Philosophy* 39:87–118.
- Bach, K. 2004. Minding the gap. In C. Bianchi (Ed.), *The semantics/pragmatics distinction* (pp. 27–43). Stanford: CSLI Publications, Stanford University.
- Burge, T. 1979. Individualism and the Mental. *Midwest Studies in Philosophy*. 4: 73–121.
- Carroll, L. 1871. *Jabberwocky*.
- Carston, R. 2002. *Thoughts and utterances. The pragmatics of explicit communication*. Hoboken: Wiley-Blackwell.
- Clark, H. H. 1983. Making sense of nonce sense. In G. B. Flores d'Arcais & R. J. Jarvella (Eds.), *The process of language understanding* (pp. 297–331). London: Wiley.
- Chomsky, N. 1980. *Rules and Representations*. New York: Columbia University Press.
- Davidson, D. 1973. Radical Interpretation. *Dialectica*. 27: 314–28.
- Davidson, D. 1978. What Metaphors Mean. *Critical Inquiry*, 5(1): 41–7.
- Davidson, D. 1986. A Nice Derangement of Epitaphs, in E. Lepore (ed.). *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Oxford: Basil Blackwell: 89–107.
- Davidson, D. 1994. Social Aspect of Language. In Brian McGuinness & Gianluigi Oliveri (eds.), *The Philosophy of Michael Dummett*. Kluwer Academic Publishers. pp. 1—16.
- Fodor, J. 1975. *Language of Thought*, Thomas Y. Crowell Co
- Fodor, J. and E. Lepore. 1993. "Is Radical Interpretation Possible?" in *The Philosophy of Donald Davidson*, edited R. Stalker, Kluwer Press, 1–23.
- Gilbert, M. *On Social Facts*. New York: Routledge, 1989.
- Hawthorne, J. and E. Lepore. 2011. *Journal of Philosophy* 108 (9): 447–485.
- Kripke, S. 1972. Naming and Necessity. In Gilbert Harman and Donald Davidson (eds), *Semantics of Natural Language*, 253–355. Dordrecht: Reidel.
- Lepore, E. 2009. The Heresy of Paraphrase: When the Medium Really Is the Message. *Midwest Studies in Philosophy*. 33:177–197.
- Lepore, E. and M. Stone. 2010. Against Metaphorical Meaning. *Topoi*, 29(2): 165–80.
- Lepore, E. and M. Stone. 2014. The Poetic Imagination. In John Gibson and Noel Carroll (eds), *Routledge Companion the Philosophy of Literature*, London: Routledge. Pp.
- Lepore, E. and M. Stone. 2015. *Imagination and Convention*, Oxford University Press: Oxford.
- Lewis, David K. 1969. *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.
- Lewis, David. 1979. Scorekeeping in a Language Game. *Journal of Philosophical Logic*, 8(3):339–59.
- Ludlow, P. 2014. *Living Words*. Oxford University Press: Oxford.
- Millikan, R. *Language, Thought and Other Biological Categories: New Foundations for Realism*. Cambridge, MA: MIT Press, 1984.
- Millikan, R. 1998. Language conventions made simple. *Journal of Philosophy*, 95(4),

161–180.

Norman, D. A. 1981. Categorization of action slips. *Psychological Review*, 88(1), 1-15.

Putnam, H. 1975. The Meaning of "Meaning." In *Mind, Language and Reality*, 215–71.

Cambridge: Cambridge University Press.

Recanati, F. 2010. Truth-conditional pragmatics. Oxford: Oxford University Press.

Reimer, M. 2004. What malapropisms mean: A reply to Donald Davidson. *Erkenntnis*, 60, 317–334.

Searle, J. 1969. Speech acts. Cambridge: Cambridge University Press.

Travis, C. 2000. *Unshadowed Thought: Representation in Thought and Language*.

Cambridge, MA: Harvard University Press.

Wiggins, D. 1971. "On sentence-sense, word-sense and difference of word-sense.

Towards a philosophical theory of dictionaries." in Danny D. Steinberg and Leon

A. Jakobovits (edd.) *Semantics: An Interdisciplinary Reader in Philosophy,*

Linguistics and Psychology [Cambridge: Cambridge University Press], 14-34.

Zwicky, A. 1979. Classical malapropisms. *Language Sciences*. 53 1(2):339-348.