I find the account of metaphor offered in Donald Davidson’s “What Metaphors Mean” fascinating for a number of reasons. The overall argument, that metaphors mean nothing other than what they mean literally, strikes me in many ways as absolutely right, and corrective of a certain tendency both in the humanities and in more popular forms of criticism to use the word “meaning” where it doesn’t apply. Of course, the other M word also has its history of abuse, whereby it often becomes a sort of last resort for closing down all discussion or counterargument: “You’re missing the point—I was only being metaphorical.” In these contexts “literal-mindedness” seems to take on a pejorative flavor, as though appealing to some clichéd stereotype of the math nerd whose every attempt to read a poem leaves him dumbfounded. While it is certainly true that metaphoric usages of words are more difficult to understand than their non-metaphoric counterparts, that there is something “extra” required in order to make sense of them, it is easy to fall into thinking that when we do succeed in making sense of them we are making use of a second species of meaning. Davidson’s proposal is that we correct this sort of false picture by, as it were, uncrossing the wires; directing our attention to literal meaning is perfectly appropriate because it gives us a way of understanding more fully what metaphor is not.

In what exactly this extra machinery might consist is of course the impossible question, on which Davidson remains understandably silent. One would suppose his model to fit more or less exactly with a picture of language that takes meaning to be the currency of compositional
semantics, and that nameless “something else” (‘use,’ or “implied meaning,” or whatever) to be the currency of pragmatics. Let metaphors “mean” only what it is that they mean regardless of their context, and leave the question of what it is that they “do” to your preferred theory of implicature or performative acts, or relevance. Fair enough. We will return later to how this larger perspective leads his theory to be cast in the particular way that it is. What he does end up saying about this mysterious something extra is along these lines: metaphors invite us to make a comparison between two things; they direct our attention to a similarity between them, in much the same way as similes. This is obviously not the first time that similes have been invoked as a way of explaining metaphor; indeed, the formula “simile—‘like’ = metaphor” has even made it into the grammar school curriculum. And how does this work, exactly? Let us imagine the most banal of examples: The Bard writes, “Those are pearls that were his eyes,” and we are invited to make a comparison between eyes and pearls, noticing all sorts of similarities; roundness, whiteness, immense value, etc.

Further comparison between metaphors and similes yields the interesting observation that all similes are literally true, whereas their metaphoric counterparts are literally false. Our heavily schematized example, “His eyes are (like) pearls,” seems to hold up to this claim. The idea appears to be that “similar” is really a trivial or vacuous predicate, because it is true of any pair of things that they are alike in some way. The way Davidson puts it, “This is like that—Tolstoy is like an infant, the earth is like a floor. It is trivial because everything is like everything, and in endless ways.” The problem is the following: when I say to you, “His eyes are like pearls,” I am not really saying anything, because everything is like everything. All I am saying is that there exists some similarity between his eyes and pearls, when it is a given that there exists a veritable infinitude of rapports between any pair of things.

Davidson’s claim about similarity opens onto a fundamental problem, for which he doesn’t try to propose any solution. On the one hand he appears to have hit upon an important insight, one that for instance explains why genre theory’s rhetoric of “repetition and yet difference” really isn’t saying anything. The idea that “there exists a relation” is a predicate that can be true of some pairs of things but not others, though extremely awkward, is a prevalent one in many schools of art criticism and aesthetics, particularly as regards questions of hermeneutics. In the face of certain texts it is sometimes claimed in all seriousness that there is no relation between one unit and the next; that, for instance, there is
no relation between certain pairs of words in the plays of Tristan Tzara or Roger Vitrac. But has this century not proven the critic capable of building a viable exegesis out of virtually any juxtaposition? In cases where the text does not bother to make any meaningful contributions of its own, the only possibility is to place the burden of interpretation on the reader.

On the other hand, can “AreSimilar(x,y)” really be vacuous predicate? For it would seem to be the lynchpin of many a philosophical claim. Are we to conclude that, for example, the analogy in §11 of the *Philosophical Investigations* between the functions of tools and the functions of words is merely an inconsequential statement? That would certainly be a pity, for it is one of the book’s most illuminating remarks on the range of ways in which we use the term “meaning.” It must be doing something of value. So then should we reject the idea that all similes are literally true on the ground that it unfairly trivializes the simile? I don’t think we have to do that either; the fact is that we use similes meaningfully all the time, and what the observation that everything is like everything should reveal is that rather than pointing out the mere existence of a similarity (which is all they do literally), similes draw our attention to a specific similarity (or set of similarities, possibly).

To turn once again to the simulative half of our admittedly boring example, the question then remains about how exactly the one significant or important similarity between pearls and his eyes is to be extracted from the mere declaration that similarities between them exist. On this matter all Davidson has to offer is that to which likeness our attention is directed depends upon the context of the uttered comparison. And so common sense furnishes us with two possible explanations for the functioning of similes: a) When I say “His eyes are like pearls,” I am thinking of one correspondence in particular, say that they are white, and your job is to reconstruct my thought, and b) When I say “His eyes are like pearls,” a specific relation between the two (that they are white; let’s say that in this example we were discussing the dynamic range of a photograph) can be made to fit the context in which I said it, and your job is to compute the best fit. The metaphoric half of the example should work in nearly the same way, except that where the simile is true, the metaphor “His eyes are pearls” is false. The search for the relevant similarity between his eyes and pearls, however, should be the same. (a) holds the listener to rather unreasonable expectations; (b) seems to work, but it relies on a theory of meaning that separates truth-conditional content, which is ahistorical and context-independent, from the use to
which that meaning is put. I will not attempt to overthrow this view, as it certainly has its attraction, but only to suggest a possible alternative.

There is a third explanation, which I find appealing because it brings out an underlying tension within this picture of metaphor and points to a larger overhaul of the theory that can resolve the tension while preserving its insights. The late Wittgenstein leads us straight to it. 4

§85 of the *Investigations* offers us the dilemma of the signpost:

A rule stands there like a sign-post.—Does the sign-post leave no doubt open about the way I have to go? Does it shew which direction I am to take when I have passed it; whether along the road or the footpath or cross-country? But where is it said which way I am to follow it; whether in the direction of its finger or (e.g.) in the opposite one?—And if there were, not a single sign-post, but a chain of adjacent ones or of chalk marks on the ground, is there only one way of interpreting them?

The implications of what “stands there like a signpost” might mean are spelled out in §454: “How does it come about that this arrow → points? Doesn’t it seem to carry in it something besides itself? . . . The arrow points only in the application that a living being makes of it.” What is it exactly that we bring to this splotch of ink in order to determine which way it points, or even that it points? An interpretation, perhaps? As we soon find out, that temporary solution has its problems.

The rule-following passages go against a certain idealization of language (possibly Tractarian), one that assimilates the rule to its expression (which is probably why it could just as well have started off by saying, “The expression of a rule stands there like a signpost” 5). When I write, “1, 4, 9, 16, 25” I have an understandable temptation to say that the rule to which I am adhering is $y = x^2$. But is “$y = x^2$” the rule itself or only an expression of it? A second glance reveals that it in itself is not the rule, because anything with which we cannot say that a new input chosen at random either will or will not be in accord has to lose candidacy for being a rule. We would like to think that there is only one way of understanding the formula $y = x^2$, that it describes a continuum of moves already made in advance; but the fact is that we only arrive at an idea of what “$x^2$” might mean through our everyday experience with examples of quantities squared. And it is clear, from §185 or any
argument of the “quus” persuasion, that those examples in themselves
describe not merely one continuum, but as many as we please.

The idea that the expression of a rule “just stands there like a sign-
post,” as Kripke suspects, poses a threat not only to the act of writing
numbers in a series, but of meaning anything at all. For clearly it is not
only expressions like “x^2” that we learn to use through example, but
also expressions like “is a cherry.” Language acquisition data attests to
the fact that the situation in §185 can reproduce itself in the case of
learning any word; a child will seem to have grasped its meaning, then
suddenly an unforeseen situation will lead him to use it incorrectly. So
how is it that we ever mean anything?

Is this not exactly the sort of dilemma that Davidson raised with regard
to metaphor? The purpose of a metaphor is to draw our attention to
a similarity between two things, but on the question of what similarity,
the metaphor in itself tells us very little; in a way it does seem to stand
there like a signpost. I say, “Those are pearls that were his eyes,” but
there is absolutely nothing encoded in that statement that can direct
me to how those pearls “were” his eyes. To use the language of §201, no
one similarity can be determined by a metaphor, because every similar-
ity can be made out to accord with it. This is how the skeptical paradox
presents itself, but of course Wittgenstein doesn’t stop here, and neither
should we. In practice, of course words have meanings; expressions of
rules do not just stand there like signposts. We read them all the time
and are able to grasp their sense; able to determine what accords with
them and what conflicts.

Wittgenstein’s solution to the signpost problem comes in §201, which
I will present slightly out of order. At the end of the passage we have an
attractive redefinition of the very idea of interpretation: “But we ought
to restrict the term ‘interpretation’ to the substitution of one expres-
sion of the rule for another.” Innocent as it may seem, this remark is
rather helpful in avoiding confusions around terms like “interpretation,”
“meaning,” “understanding,” etc., which frequently have a way of being
thrown around as synonyms. Here “interpretation” is given a much more
specific meaning (something closer to “glossing,” it seems) whereby my
interpreting the word “crimson” consists not in understanding it in a
certain way, but merely in replacing the word with another expression,
like “deep red.” Since an expression of a rule also stands there like a
signpost, making recourse to an interpretation in order to grasp the
meaning of an expression is only replacing one signpost with another;
if we claim that this is the only path towards understanding, we have
indeed led ourselves into an infinite regress. The solution comes at the beginning of the passage:

This was our paradox: no course of action could be determined by a rule, because every course of action can be made out to accord with the rule. The answer was: if everything can be made out to accord with the rule, then it can also be made out to conflict with it. And so there would be neither accord nor conflict here.

It can be seen that there is a misunderstanding here from the mere fact that in the course of our argument we give one interpretation after another; as if each one contented us at least for a moment, until we thought of yet another standing behind it. What this shews is that there is a way of grasping a rule which is not an interpretation, but which is exhibited in what we call ‘obeying the rule’ and ‘going against it’ in actual cases.

So there is a way of grasping a rule that is not an interpretation; a way of grasping it that is not a substitution of one expression of the rule for another. This notion of interpretation helps us to see the difference between the written formula “y = x²” and the rule that it represents, just as it helps us to see the difference between “deep red” and what that constituent of words means. Substituting one expression of a rule for another helps to make clear how we are understanding it, but the expression itself is not to be confused with our understanding. It isn’t as if the moment when I know how to go on is the moment when I think of the formula “y = x².”

And what is this way of grasping a rule that is not an interpretation, but which is exhibited in what we call “obeying the rule” and “going against it”? §198 suggests that we should think of it as a custom: “. . . I have further indicated that a person goes by a sign-post only in so far as there exists a regular use of sign-posts, a custom.” So it is a custom that allows us to make actual use of signposts, rather than merely leaving them to stand where they may; it is a custom that determines whether I am pointing in the direction of my fingertip or my wrist. To this idea John McDowell adds the following coloring:

I have been urging that we should avoid the threat by not letting the regress start—by not letting it seem that the concept of interpretation must be in play if the concept of accord is to be secured its application. [§198] suggests that we can avoid that appearance by insisting on a bit of common sense about following a sign-post. When one follows an ordinary sign-post, one is not acting on an interpretation. That gives an overly cerebral cast to such routine behavior.
What this passage argues explicitly is that an appeal to custom is an appeal to training in a technique, but what it also suggests to me is that following a custom can also be thought of as following a set of habits (which habits are put into place by earlier training). If claiming that we need to make use of interpretations in order to follow signposts is a sort of over-intellectualization or over-theorization of the way in which we actually use them, then perhaps what we want to say is that the process of following a custom is something habitual or automatic. The notion of the custom as habit fits rather well with commonsense intuitions about how meaning actually works; when parsing a typical sentence I do not have to reflect on it critically in order to understand it; rather, my grasping of it is almost a reflex. This face of the question will come to have its worthwhile ramifications.

If the problem of interpreting the signpost is at all analogous to the problem of reading a metaphor, the obvious next question is whether Wittgenstein’s solution to the regress of interpretations also applies. After all, metaphor is burdened with its own problems of infinite regress, as evidenced by the frequent complaint that accounts of metaphor themselves are never anything but metaphorical, or more generally that in philosophy we can never escape metaphor. I would say the possibility of bringing in the appeal to custom looks promising. But first we have to sort out a tangle that has been threatening to germinate ever since we began noticing the similarities between the workings of meaning and the workings of metaphor. What of Davidson’s central thesis, that there is no specially metaphoric meaning? Will noting the simile/signpost parallel bring us forcibly to the conclusion that there is nothing left at all to distinguish metaphor and meaning?

Yes and no. It is not necessary to resort to so strong a conclusion, but there are notable counterexamples to Davidson’s hypothesis that metaphoric usage is located strictly outside the domain of meaning; namely, intermediate cases—usages that are particularly colorful and suggestive, but for which it is difficult to decide firmly whether we are dealing with metaphor. Think, for instance, of the various equivalents for “kill” and try to decide whether they are metaphors or synonyms: dispose of, do (someone) in, liquidate, terminate, take care of, eliminate, etc. When I say, “I took care of Bob,” am I speaking poetically, or am I just saying that I killed him? I almost want to say that there is a tendency in any skilled writer to produce texts that teem with such borderline cases. Consider the following passage:
As we all know our own faults, and know them commonly with many aggravations which human perspicacity cannot discover, there is, perhaps, no man, however hardened by impudence or dissipated by levity, sheltered by hypocrisy, or blasted by disgrace, who does not intend some time to review his conduct, and to regulate the remainder of his life by the laws of virtue.\textsuperscript{9}

Are “hardened,” “sheltered,” “dissipated,” and “blasted” being used literally or as metaphors? It is difficult to say.

\textbf{II}

At the moment we seem to be in a bind, but all that is required to extricate us from it is an ordinary fact about diachronic semantic change: that over time, metaphoric usages become meanings. A few examples for the unconvinced: French \textit{feuille} “leaf, sheet of paper” is derived from “leaf,” French \textit{entendre} “to hear” from “to understand,” French \textit{fermer} “to close” from “to fix, make firm or fast,” English \textit{chill} “to calm down” from “to cool,” and English \textit{stud} “good-looking man” from “a male animal used for breeding.”\textsuperscript{10} Indeed, it is revealing that so many of the terms used traditionally to describe lexical semantic change are the tropes of classical rhetoric (metaphor, metonymy, synecdoche, litotes, hyperbole, etc.).

The phenomenon referred to as “grammaticalization” is perhaps a more extreme example. Lyle Campbell characterizes it as follows: “. . . where an independent word with independent meaning may develop into an auxiliary word and, if the process continues, it ends up as a grammatical marker or bound grammatical morpheme.”\textsuperscript{11} A noun, let’s say, might start off with an ordinary sense, begin to widen its meaning into something more metaphorical, and eventually become purely grammatical. For example, the French “pas” originally meant only “step.” Next it began to populate expressions like “not one step,” which through repeated usage came to mean “not one bit.” Finally, it was absorbed completely into the morphology of the language’s negation, today meaning “not.” Here the term in question really seems to have gone from one extreme to the other.\textsuperscript{12}

So it seems as though we have good reason to think of metaphors as meanings in the making; as “proto-meanings.” What I am suggesting is that we change the flavor of Davidson’s suggestion by adding to “whatever metaphors are up to apart from their literal meaning is not
itself a variety of meaning” a modest “... yet.” What had its debut as a synchronic theory about which metaphorical reading best fits the context of the metaphor’s utterance now becomes a story about how new meanings are spawned. And the late Wittgensteinian picture of language turns it into a fascinating one.

What §201 can bring to these considerations is the idea that a metaphor is the beginning of a practice; that it lays the groundwork for the eventual blossoming of a custom. It is the kernel of a usage which, if it catches on, will solidify into an institution; a custom into which one can be initiated by training. If, on the other hand, it is not put into continual use, it will remain a metaphor. The sense in which a custom is habitual is useful here, because it explains why metaphors require more work in order to be understood. The more consolidated the custom, the more habitual it becomes; therefore the less the exertion required to make sense of an expression that relies on it. My suspicion is that this is exactly why formulations of the sort proposed by Dan Sperber and Deirdre Wilson rely on notions such as “processing effort”; under their view humans automatically aim at maximal relevance, which is maximal cognitive effect for minimal processing effort. Metaphoric usages, in a way, strive to be non-metaphoric; though they require more in the way of mental labor, they strive to come as close to meaning as they can (given that meaning involves almost no processing effort, at least comparatively). The principle of relevance could be said to be the driving force behind the shift from metaphor to meaning.\footnote{13}

There is a tension in Davidson’s original account: on the one hand there seems to be a natural inclination to explain metaphor in terms of “a second meaning,” but on the other hand whatever it is that metaphors seem to accomplish beyond their literal meaning really can’t be understood as any kind of meaning at all. His thesis does a good job of explaining what some would take to be the defining property of a metaphoric usage: its inability to be paraphrased (i.e., if a metaphor has a second meaning, it should be perfectly glossable). But incorporating Wittgenstein’s solution to the infinite regress preserves this explanation, recognizing what distinguishes metaphor from meaning (or at least what distinguishes it for the time being), while also giving us a way of understanding why it is so tempting to think of metaphoric usage as a second meaning. It isn’t quite a second meaning, but something rather like it: a second meaning in the offing.

Davidson of course insists on the distinction between meaning and use, a distinction that would appear to be very much at odds with the
Wittgensteinian notion of meaning, which is nothing other than “the use of the word in a language” (§43). I will do my best, however, to hold back from fully contrasting the two, for it is neither clear to me whether Davidson intends his term “use” to mean anything like Wittgenstein’s, nor whether Wittgenstein is categorically opposed to anything we might want to call compositional semantics (though passages like the end of §49, and, if Cora Diamond is right, the austere conception of nonsense do make it seem that way). As far as I can tell, focusing on the fact that metaphors evolve into meanings does not necessitate our adopting one of these models of meaning and rejecting the other, so I suppose that for now I will leave the question of which is more useful up to the preference of the reader.

Interestingly enough, Davidson makes an allusion to the scenario I am considering, under the rubric of the “dead metaphor”:

Once upon a time, I suppose, rivers and bottles did not, as they do now, literally have mouths . . . when ‘mouth’ applied only metaphorically to bottles, the application made the hearer notice a likeness between animal and bottle openings. (Consider Homer’s reference to wounds as mouths.) Once one has the present use of the word, with literal application to bottles, there is nothing left to notice. There is no similarity to seek because it consists simply in being referred to by the same word.

This is a particularly pregnant example. As if a galaxy filled with stars at different stages in their life cycle, “mouth” is a constellation of three usages, none of which has vanished entirely from the cosmos: (a) its first literal meaning (mouth of an animal), (b) a second literal meaning which was once a metaphor (mouth of a river or bottle), and (c) a usage that is still metaphorical (the Homeric mouth-wounds). The data are all here; each meaning has been preserved side by side. The language of “death” is interesting here; in what sense has the metaphor died? We may be inclined to picture the following scenario: in the era during which “mouth” had only one meaning, I say to you “Meet me by the mouth of the river.” My saying this to you invites you to make a comparison between some part of the river and mouths. There are an infinite number of possible correspondences; was I talking about the part of the river that “talks,” or the part of the river that “eats,” or the part of the river that “breathes,” etc.—which similarity did I mean for you to notice? Suddenly, something allows you to grasp what I was probably talking about: the part of the river that opens onto the sea. What was this mysterious force that allowed you to grasp it? In the end it wasn’t mysterious at all; it was the same instinct
that allowed you to determine which way my finger was pointing, only much more faint. The metaphor continues to be put into use; to weave itself into everyday practice, and once its usage has become so definitive that the search for similarities shrivels to (to borrow once again from the lexicon of the *Tractatus*) an extensionless point, it passes from metaphor to meaning. What was once a faint inclination has now become a reflex or (we might say subconscious) habit.

Davidson’s next remark is a bit strange: “Novelty is not the issue. In its context a word once taken for a metaphor remains a metaphor on the hundredth hearing, while a word may easily be appreciated in a new literal role on the first encounter” (p. 499). Not only is this a rather abrupt dismissal of the idea that a metaphor lays the rudiments for a future meaning, but it doesn’t make sense. How can a metaphor be uttered hundreds of times in the same context, any more than Heraclitus can step into the same river on a hundred separate occasions? Will each utterance not of necessity usher in a new context? Also, I don’t see how it can be anything but unintuitive to say that a word may be easily appreciated in a new literal role on first encounter. Certainly there exist such things as neologisms, but it just seems too indefensible to hold that they can come into use full-blown, as we say, from the head of Zeus.

Neologisms need a significant degree of outside help before they can be understood, and even then it probably wouldn’t be accurate to call the act of computing their meaning “easy,” at least not in the sense that understanding “I tie my shoes” is easy. Consider “yahoo,” for instance, or “nostalgia.” There had been a place waiting for “yahoo” within the larger cultural myth of the explorer who hits upon a land of savages long before the writing of *Gulliver’s Travels*. The idea, more or less, was already there; this was only a new name for it. “Nostalgia,” originally a medical term, was coined by Johannes Hofer in 1688 as part of a dissertation on his patients’ unfathomable desire to return to their native land. But as I understand it (I’ve not read the document), he doesn’t just begin using the term out of the blue with no clarification; it is surrounded by a fair amount of explanation regarding the phenomenon it is supposed to be designating. These are two of many ways in which new words might be kneaded into a language-game (or the game kneaded around them).

Similarly, new metaphors can only be intelligible with the help of the customs or habits that are already in the process of circulating. When Harold Arlen sings about “that same old witchcraft when your eyes meet mine” for the first time (supposing that he was the first, in any case), we only have access to the proto-customs that allow us to understand how
“witchcraft” is being used here in virtue of other expressions already in use, ones which draw our attention to the parallels between romance and bewitchment that have been propagating through Western consciousness since Shakespeare and probably before. Likewise, the groundwork for making sense of something like “love is a volcano” will have a difficult time being brought into general use unless something on the order of “love is a flame” has already been in use. It is as if I am encountering a new style of arrow, and can only begin to think about which way it’s pointing by reflecting on my previous experiences with arrows.

§199 seems to be relevant here: “It is not possible that there should have been only one occasion on which someone obeyed a rule. It is not possible that there should have been only one occasion on which a report was made, an order given or understood; and so on.” An interesting remark, and one which seems to set up a sort of Catch-22 dilemma for the neologist or coiner of clever metaphors. And what of hapax legomena, words of a language that are attested only once? If there cannot only be one occasion on which someone obeyed a rule, how can anyone have obeyed it for the first time? Does this mean that a group of people must have obeyed it simultaneously? That sounds unlikely. The only explanation that comes to mind is the sort of thing I have been suggesting; that the customs or institutions governing the use of existing words and metaphors must make a place for it in advance, as though letting a new vehicle onto the highway.

University of Pittsburgh

1. Here referring to similes of the “x is like y” rather than of the “as x as a y” variety.
5. John McDowell made this suggestion during a seminar on Wittgenstein in 2003.


