Comment on "Can memes explain the birth of comprehension?"

Daniel C. Dennett

Center for Cognitive Studies,
Tufts University, Medford, MA USA

Paweł Grabarczyk explores the shortcomings of my meme theory, concentrating on my rather perfunctory treatment of meaning in *BBB*. He notes that I deem words the best examples of memes and claims that if I don't have an account of the *meanings* of words as memes, I can hardly use memes to explain comprehension. He is right, but I think that I have provided the elements of an account of meaning of words that can be readily united and presented to meet his excellent challenge. In *BBB* I concentrated on phonology and semantics, leaving syntax largely untouched, since I couldn't see through the fog of war among the linguists on this contentious topic, and Grabarczyk also sets syntax aside, but recently my own thinking on how to handle it has been informed by Daniel Dor's remarkable book, *The Instruction of Imagination* (2015)*,* which analyzes language as a "social communication technology," banishing most of the Chomskian innateness dogmas and replacing them with reverse engineering of culturally transmitted habits and dispositions. Dor ignores meme theory and doesn't rely as much as he should on evolutionary processes (and free-floating rationales), but he has clearly set out a reimagined set of specs for language, filling in many details only dimly suggested by my sketchy account. I recommend it to all serious thinkers about the phenomena of language.

In my treatment of both phonology and semantics, I stressed the role of normativity, but not the normativity which Grabarczyk commends, the normativity discussed by Brandom in *Making it Explicit.* It is the *correction to norms* that turn phonemes into digitized elements that can be transmitted with high fidelity. This ongoing dynamic process accounts for the stability or robustness of the phonology of words over long replicative histories. Pronunciations change, but slowly, and the accents of isolated populations can be largely unintelligible to other speakers of "the same language" (a phenomenon nicely akin to ring species) but enough

survives intact to make evolution by natural selection of phonology a reality around the world. With few exceptions, the shifts are not just gradual but unnoticed by the vectors, and almost never the result of deliberate social engineering. This is what permits a seldom-heralded feature of natural language: its ability to transmit messages that are not understood (or more usually, incompletely or imperfectly understood) by the transmitter. This goes a long way to providing the fidelity of transmission of meaning (piggybacking on phonology) through a chain of semi-comprehending speakers and hearers. This phonological stability is also a key element in the utility of speaking to yourself. We learn to hang our inchoate comprehension on familiar phonological hooks that help us return to complex concepts with Google-like ease. "Let's see, ... I *promised* Alf a *salmon* if he lent me his *fishing pole. . . .*" a simple memory that is beyond the cognitive reach of any non-speaking animal or child.

A similar kind of normativity underlies semantic stability. It is dynamic and historical, shifting with the environment, both linguistic and non-linguistic, in which it thrives. As I put it in my essay on Brandom, "The Evolution of 'Why'" (2010):

> There is no way to capture the semantic properties of things (word tokens, diagrams, nerve impulses, brain states) by a micro-reduction. Semantic properties are not just relational but, you might say, super-relational, for the relation a particular vehicle of content, or token, must bear in order to have content is not just a relation it bears to other similar things (e.g., other tokens, or parts of tokens, or sets of tokens, or causes of tokens) but a relation between the token and the whole life—and counterfactual life—of the organism it "serves" and that organism's requirements for survival and its evolutionary ancestry.

In that essay, and in *BBB*, I distinguished Brandom's *social censure* style of normativity from my *engineering* style of normativity (a handy way of emphasizing the difference is to look at the negative terms, "naughty" and "stupid") and argued that mine is the more basic, since Brandom's must be left as a skyhook unless he grounds it in evolution ultimately. It is not hard to see how an evolutionary account of the utility of communicative practices could support the maintenance of Brandom's (and Haugeland's) censoriousness—a fine case of a disposition that people find "natural" without having to understand why it's a good thing—a *free-floating rationale*. I have not read Grabarczyk's 2017 essay about a directival theory of meaning, but the synopsis he provides ("users try to learn the meaning by remembering iconic, prototype circumstances in which the words were presented to them") seems to be another account, like Grice's (see *BBB*, pp. 292–4), where a free-floating rationale is presented as an articulated reasoning process, which it seldom is. (I can recall a few times in my life when I've tried to remind myself of a word meaning by recalling an iconic moment, but in all but a few cases I have source amnesia for the word-learning events in my life.) Mark Richard's new book, *Meanings as Species* (2020), elaborates the details of how the semantics of words shifts gradually, while maintaining enough stability to count as fixed for most practical purposes—just like species. What anchors semantics? Richard speaks of the *interpretive common ground*, which is communally maintained much as Brandom says, in spite of not being grounded in something like hard-edged definitions or Fregean intensions or essences. And of course Ruth Millikan's masterpiece, *Beyond Concepts: Unicepts, Language, and Natural Information* (2017), gives a fine account of the way things in the world, or more specifically, affordances, anchor meanings so people don't actually have to share the same *concepts*. (They each have their own *unicept* for CAT and HOUSE and DEMOCRACY, and they coordinate through

social and environmental interaction to keep these unicepts largely in harmony.) The old quest for Necessary and Sufficient Conditions for being the concept of X can be gracefully set aside in this solidly Darwinian enterprise.

Grabarczyk notes that for words to play a role in generating genuine comprehension they must have a meaning component, not just a phonology, and in fact words are quite complex entities: informational structures in brains that play complicated roles. They may be transmitted perceptually, so they have audible or visible phenotypes, in effect, available to conscious discrimination and recall, but they are otherwise as invisible as genes, and their meaning components are famously hard to enunciate—a task that has engrossed philosophers since Socrates first proposed it. We are fluent users of our words—we have the *knowhow*—but linguists have struggled for decades to expose the structure behind our competences. Most of their tokens (if Chomsky, 2002, pp. 75–77, is right!) are invisible and inaudible but known only to the speaker of a silent soliloquy. We are so used to words that we don't easily recognize how special they are.

Grabarczyk points out that we know more about the mechanics of virus replication than we do of word replication, but we now are gathering the data about the cognitive neuroscience of language learning that will move us beyond the "Mendelian" stage with words and other memes. It is worth remembering that Mendelian genetics thrived as a science for decades with no mechanistic details and provided the specs for the "computational level" (in David Marr's (1982) sense) that led Crick and Watson and the others to home in on DNA. We shouldn't hope for a similarly universal implementation of the software of memes, but we are beginning to sort through the options for the laying down of phonology and semantics in the brain. Jackendoff's recent work (1996, 2002, 2007, 2012, forthcoming) makes some pioneering inroads into this important topic.

Grabarczyk asks "Are there memes that could not be replicated by cognitive systems that are unable to comprehend?" An excellent question which I should have raised and answered explicitly in *BBB*. Yes, by the thousands. They are ways of doing things that have to be explicitly taught by teachers and learned with effort by students and apprentices. Some students understand them deeply and others settle for workaday fluency with techniques they have been trained to use but only dimly understand. These memes are supported by thousands more that we all use competently with various intermediate levels of explicit understanding. Cecilia Heyes' recent book, *Cognitive Gadgets* (2018) is the first detailed examination of this arena and it has a bounty of insights (see Rathkopf and Dennett "Mending Wall", 2019, for a discussion). And, as Grabarczyk says, I anchor comprehension to the practices of explaining and persuading. Brandom is indeed a good source for this. Understanding is both a product of the process of explaining and a source of the capacity to explain. This is no paradox but just the standard Darwinian chicken-and-egg question that introduces the role of the cursive cycle of selection and replication. Grabarczyk claims that "we cannot really explain our reasons as we do not really know them in the first place. The best we can do is that we try to reverse engineer our competences. Problem is, this process is just as difficult as the process of reverse engineering the Alpha Zero." Not quite as difficult, precisely because we can doggedly probe our cerebral deep learning systems with questions we ask ourselves, learning to know ourselves, as Socrates put it, by engaging ourselves in that language-based activity known as philosophy.

References

Brandom, R. (1994). *Making it Explicit: Reasoning, Representing, and Discursive Commitment.* Harvard: Harvard University Press.

Chomsky, N. (2002). *On Nature and Language.* Cambridge: Cambridge University Press.

Dennett, D. (2010). "The Evolution of 'Why'". in B. Weiss and J. Wanderer (eds.), *Reading Brandom: On Making it Explicit* (pp. 48–62). New York: Routledge.

Dennett, D. C. (2017). *From bacteria to Bach and back: The evolution of minds*. New York: WW Norton & Company.

Dor, D. (2015). *The Instruction of Imagination.* Oxford: Oxford University Press.

Grabarczyk, P. (2017). Directival Theory of Meaning Resurrected. *Studia Semiotyczne – English Supplement*, 9, 65–81.

Grabarczyk, P. (2019). Can memes explain the birth of comprehension? *Avant*, 10(3).

Heyes, C. (2018). *Cognitive Gadgets: The Cultural Evolution of Thinking.* Oxford: Oxford University Press.

Heyes, C. (2019). "Précis of *Cognitive Gadgets: The Cultural Evolution of Thinking*," *Behavioral and Brain Sciences* 42, e169, 1–58.

Jackendoff, R. (1996). How Language Helps Us Think. *Pragmatics and Cognition,* 4, 1–24.

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution.* Oxford: Oxford University Press.

Jackendoff, R. (2007). *Language, Consciousness, Culture.* Cambridge, MA: MIT Press.

Jackendoff, R. (2012). *A User's Guide to Thought and Meaning.* Oxford: Oxford University Press.

Jackendoff, R. (forthcoming). Inner Speech: One Linguist's View. In *Inference: International Review of Science*.

Marr, D. (1982). *Vision.* Cambridge, MA: MIT Press.

Millikan, R. (2017). *Beyond Concepts: Unicepts, Language, and Natural Information.* Oxford: Oxford University Press.

Rathkopf, Ch, Dennett, D. C. (2019). Mending Wall (commentary on Heyes, 2019). *Behavioral and Brain Sciences*. 42, 32–33.

Richard, M. (2020). *Meanings as Species.* Harvard: Harvard University Press.