Abstract

Externalist theories in natural language semantics have become the orthodoxy since Kripke is widely thought to have refuted descriptive theories involving internal cognitive representation of meaning. This shift may be seen in developments in philosophy of language of the 1970s - the direct reference “revolution against Frege” (Wettstein 2004, 66). Almog (2005, 493) writes of the “uprising against Frege’s doctrines” that “spread like fire” based on the work of Kripke, Donnellan, Putnam and Kaplan. However, I consider Fodor’s (2004) heretical thought that something has gone “awfully wrong” in this philosophical consensus, perhaps confirming Chomsky’s (1992) view that the whole field of philosophical semantics is “utterly wrongheaded” and “crazy” by virtue of its non-naturalist assumptions and “methodological dualism.” I suggest that the externalist orthodoxy is a kind of cognitive illusion seen elsewhere in philosophy and cognitive science.

Keywords: semantics; externalism; meaning; intuition.

Externalist Orthodoxy

Externalism is widely acknowledged to be the orthodoxy in the theory of mental content and psychological states. However, despite its subjective force, externalism may be undermined by attending to its aetiology and showing how the intuitions evoked arise from deceptive mechanisms. Instead of defending internalism directly, we may ask: Why does externalism seem so convincing? This is a cognitive science of biases and illusions among philosophers.

Kripke (1972) is regarded as having “usher[ed] in a new era in philosophy” (Soames 2005, 1) by refuting a widely held descriptive conception of proper names. In the philosophy of language, this was part of the 1970s direct reference “revolution against Frege” (Wettstein 2004, 66). Frege held that something about the speaker’s cognitive state must explain the difference between sentences such as “Hesperus is Phosphorus” and “Hesperus is Hesperus.” The first is cognitively significant but the second is knowable a priori as necessarily true, even though the substituted terms are co-referential. However, Kripke’s externalist doctrine of “rigid designators” has become the orthodoxy – essentially the view of J.S. Mill that proper names have no meaning other than the name’s denotation, and a name refers to the same individual in any possible situation.

In Putnam’s slogan, the externalist orthodoxy holds that “meanings ain’t in the head” since mental content is individuated by referents in the world. This view rests on intuitions elicited by thought-experiments such as Putnam’s (1975) famous Twin Earth story, characterized as “a sort of paradigm in the philosophies of language and mind” (Segal 2000, 24). On another planet, Twin Earth, the only difference is that the clear, potable liquid in rivers and lakes has chemical structure XYZ rather than H2O. An atom-for-atom replica of an Earth person might have identical internal psychological/brain states and yet not have the same water-thoughts since Twin Earth thoughts are about XYZ. Also influential has been Kripke’s (1979) puzzle about Pierre who believes both that Londres est jolie and also that London is ugly, not realizing that London is the same city as Londres. Kripke says “I know of no answer” to the question “Does Pierre, or does he not, believe that London is pretty?” Kripke regards the puzzle as comparable to the Liar Paradox (1979, 904). On this point, Salmon (2011, 236) endorses Kripke’s “sound methodology” quoting Tarski’s classic discussion of the Liar antinomy and its intellectual challenge.

Kripke’s “primary moral” is that “the puzzle is a puzzle” (1979) and he insists that it can not be resolved by re-describing the problem, but this conception is open to challenge. A re-description need not avoid the problem but rather it may show how a pseudo-problem arises. After all, the indeterminacy of Pierre’s belief about London is not like the contradictory state of Shrödinger’s cat or the quantum wave/particle duality. To be sure, in another case, restating Zeno’s paradox of Achilles and the Tortoise (e.g. with a distance/time graph) is to sidestep the puzzle rather than solving it since the re-description doesn’t expose the flaw in Zeno’s reasoning. Kripke is right to say that talk of ‘what is really going on’ doesn’t answer his original question, but it does show clearly what’s wrong with the original question and why the puzzle isn’t a puzzle, after all. With Kripke, we can point out that “No answer has yet been given” to the question of whether Lois Lane loves Clark Kent, but we understand why.

Or, seeing the Necker Cube on two different occasions, Pierre might not recognize it as the same geometrical figure. Adapting Kripke’s (1979) words, we may ask “Does Pierre, or does he not, believe that the Figure (not the shape satisfying such-and-such descriptions, but the Figure) is facing upwards to the left? No answer has yet been given.” Fodor (2008, 76) pointedly asks “But why on earth should we suppose that the question [concerning Pierre] has a definite right answer when it’s phrased that way? And, once one sees why it doesn’t, why does it matter that it doesn’t?” However, while sharing Fodor’s dismissive attitude, we may go further to ask why the puzzle should have such a firm grip on philosophical imagination.

Thus, Devitt (1984, 385) has made a salutary distinction: “Thoughts are one thing, their ascription another.” Devitt
warns “it is a common practice … to use ‘belief’, for example, where what one means to refer to is belief ascription” (1984, 389). The failure to respect Devitt’s distinction is to blame for Kripke’s puzzle in which we seem forced to describe the hapless Frenchman as holding contradictory beliefs about London. The relevance of Devitt’s distinction should be clear: “a difference in sorts of thought ascription does not entail a difference in the sorts of thought object ascribed” (1984, 389). In this case, the question concerning Pierre’s belief about London involves thought ascription about the thing itself or de re, using our own reference, like Putnam’s thought ascriptions about H₂O and XYZ. The intuition that we can be induced to share is simply the idea that we can ascribe de re beliefs from our own perspective independently of the beliefs of the subject in question. Brandom (1994, 503) explains, “expressions that occur within the scope of the ‘that’ [in de dicto contexts] serve to specify how things are represented by the one to whom the belief is ascribed.”

**Little Choice?**

Significantly, Kripke (1972, 42) acknowledges that he was led by his “natural intuition” to his view of proper names and that there could not be “more conclusive evidence one can have about anything, ultimately speaking.” However, Farkas (2003) characterizes the “deeply rooted” intuitions as “baffling” and a “vexatious problem” that “poses a serious challenge for any attempts to give an internalist analysis.” Accordingly, we may ask why philosophers feel that the “intuitive responses to a certain kind of thought-experiment appear to leave them little choice,” as Boghossian (1998, 273) puts it. Fodor (1987a) has noted that the Twin-Earth Problem is not a problem but “just a handful of intuitions together with a commentary on some immediate implications of accepting them” (1987a, 208). Significantly, he says: “it is very plausible that all these intuitions hang together. The question is: What on earth do they hang on?” (Fodor 1987, 202). I offer an answer that gains a distinctive, if not decisive, strength from the fact that the intuitions in this domain “hang on” the same biases and illusions to be seen operating elsewhere throughout cognitive science.

**Giving Intuitions a Bad Name**

In different guises, under such headings as ‘conceptual analysis’ (Jackson 1998) or ‘conceivability’ (Chalmers 2002), intuitions have played a central role in philosophy (DePaul & Ramsey eds. 1998). Hintikka (1999, 127) suggested intuitions “came into fashion in philosophy” as philosophers’ attempted to “get on the bandwagon of transformational grammar” that they took to provide a methodological model for research into cognition. Hintikka (1999, 127,8) specifically cites Kripke’s (1972) Naming and Necessity as an influential case in point, suggesting “Unfortunately” his doctrines are “apt to give intuitions a bad name.” Even a sympathetic account by Hughes (2004) makes a damaging admission: He confesses “blindness” to Dummett’s (1973) alternative reading of key sentences but takes “comfort” from the fact that the same defect is very widespread among philosophers. However, the Müller-Lyer illusion is very widespread too. As Sosa (2001, 26) notes, the phenomenon of ambiguity is widespread in the English language and the “shiftiness” of linguistic constructions containing modal expressions is akin to lexical ambiguity of words such as “bank.” Closer are the structural ambiguities familiar to linguists and the basis for jokes such as Groucho Marx’s remark: “One morning I shot an elephant in my pajamas.” Failure to appreciate the humour through blindness to the ambiguity is a psychological defect rather than theoretical criticism.

Contrary to Hintikka (1999, 132), Chomsky’s use of intuitions in linguistics has nothing to do with being a “self-acknowledged Cartesian” or innate ideas. Nevertheless, Hintikka (1999, 133) correctly notes, in contrast to linguists’ use of intuition, “philosophers’ intuitions do not pertain to the supposed faculty of intuition itself but to the truths about which this faculty is supposed to provide knowledge.” For an egregious example, Bealer (1998, 202) argues that intuitions have a “strong modal tie to the truth” which he suggests “is a philosophical (conceptual) thesis not open to empirical confirmation or refutation.” In the same vein, Chalmers (2002) challenges the systematic scientific picture asking “Does Conceivability Entail Possibility?” However, we need not agree that “Philosophical intuition is epistemologically useless” (Cummins 1998, 125). If philosophers’ intuitions are taken properly on the model of generative grammar, they may be seen as psychological evidence rather than intimations of truth. The Putnam-Kripke intuitions might be explained like the Müller-Lyer illusion as deceptive in spite of its subjective force.

**Omniscient Philosopher-Narrator**

The model for this kind of inquiry into intuitive judgements is the ‘heuristics and biases’ program of Tversky and Kahneman (1974). This work has demonstrated the systematic unreliability of compelling intuitions resulting in a wide range of cognitive illusions to which we are prone. Seen from this perspective, I suggest externalist theories of reference involve a generic pseudo-explanatory mistake that is not confined to any one domain. For example, Chomsky has explained the need for a fully explicit grammar that avoids the unwitting dependence on the linguistic knowledge of the theorist. Of course, the potential for this error is not unique to linguistic explanation and its very seductiveness means we should expect to find it elsewhere. Generally, it seems difficult to avoid invoking internal representations which have their meaning because we, as theorists, can understand them. This has been the charge against pictorial theories of imagery by Pylyshyn (2003) and was precisely anticipated by Descartes. In this case external representations are taken as a model for internal representations and, therefore, relying on the theorist’s intelligence and invoking the notorious homunculus. Chomsky (1962) notes that a grammar may produce the illusion of explanatory completeness, but in fact have
“serious limitations so far as linguistic science is concerned” because the success of the grammar depends on being “paired with an intelligent and comprehending reader.” Chomsky explains: “Reliance on the reader’s intelligence is so commonplace that is significance may be easily overlooked” (Chomsky 1962, 528). In a different guise of interest here, the theorist posits mental representations based on his own knowledge of the truth rather than the subject’s beliefs. In this case, philosophical intuitions arise from tacitly adopting the perspective of an invisible narrator – the illusion of the omniscient story-teller, the literary device that Mario Vargas Llosa (1975) aptly refers to as the “philosopher-narrator.”

**Residue of Commonsense**

Pietroski (2003) suggests that despite a considerable literature on reference, “no one has shown that names do bear any interesting and theoretically tractable relation to their bearers.” If he is correct, we are owed an explanation of how so many philosophers could have been so misguided. Chomsky characterizes the commonsense conception of semantics as a kind of illusion and points to the kind of diagnostic, aetiological concern I wish to pursue: “Here, I think, philosophers and linguists and others who are in the modern intellectual tradition are caught in a kind of trap, namely, the trap that assumes that there is a reference relation” (2012, 28). That is, “there is no word-thing relations.” This is undoubtedly a shocking remark that flies in the face of the most obvious, taken-for-granted facts about language. Of course, that’s just the point. Chomsky suggests that we may suffer from a “residue of commonsense,” some deeply persuasive, but illegitimate, “distorting” picture of the world (see also Egan 1999, 188). Word-thing relations are “mythical” by contrast with the question of “how the person’s mental representations enter into articulation and perception” (1996, 23), but this is *syntax*. Chomsky (2000, 148) suggests that we can have no intuitions about such questions as whether an identical replica of ourselves uses the word “water” to refer to something, XYZ, which is not H₂O because the key terms such as “extension” and “reference” are technical inventions. In the same way, it would be pointless to explore our intuitions about “tensors” or “undecidability.” However, there can be no doubt that certain intuitions may be consistently induced in philosophers and others by the notorious thought experiments. These are not random in the way that intuitions about tensors might be among the uninformed. The vast philosophical literature attests to the existence of systematic, robust and widely shared intuitions that are at the heart of externalism.

**Who Cares What the Mayans Think?**

Recently, the question has been illuminated from a new angle by empirical inquiries into the cross-cultural variation in intuitions on which philosophers have relied (Machery et al., 2004). These studies have challenged the universality of the evidence on which philosophical puzzles have relied. For example, Segal (2004, 339) says “we should not trust those intuitions” because Putnam and Kripke “mistakenly think that their intuitions are ‘ours’, that they are representative of those of all sensible, reflective humans” (2004, 340). Segal reports studies “designed to tap relevant twin-Earth intuitions among tribespeople” such as the Mayans of the Yucatan in Mexico. The data are mixed, but Segal says “surely these data … should be given considerably more weight than Putnam’s intuitions about Oscar’s “water” concept and Kripke’s intuitions about medieval “unicorn” concepts (Segal 2004, 343). In the same vein, Machery et al. (2004, B7) found that “Chinese subjects tended to have descriptivist intuitions, while Westerners tended to have Kripkean ones” and these data suggest “significant philosophical conclusions.” The authors conclude:

> We find it wildly implausible that the semantic intuitions of the narrow cross-section of humanity who are Western academic philosophers are a more reliable indicator of the correct theory of reference … than the differing semantic intuitions of other cultural or linguistic groups. (2004, B9)

**Competence or Incompetence?**

Devitt (2011) rejects the challenge of cross-cultural evidence to semantic theory because they tested the wrong subjects. The intuitions of ordinary folk are unreliable by comparison with intuitions of “experts,” namely, “metaphysicians and other philosophers.” However, we need not accept philosophers’ intuitions as authoritative divinations to treat them, instead, as diagnostic evidence of illusion among those who suffer from it – data for the development of a theory of ‘tacit knowledge’ or “competence” (i.e. incompetence).

The point has been missed in the ongoing controversy about empirical inquiries into intuitions. Recently, Nagel (2012) has argued that epistemic intuitions do not, after all, vary in ways that pose a challenge, but Stich (2012) has defended such research and its threat to philosophical reliance on intuition. He cites evidence that even the Müller-Lyer visual illusion is not universally shared among all human cultures. Kalahari San foragers apparently do not judge the familiar lines as differing in length. So what? Devitt, Stich and Nagel miss the point that it remains a matter of psychological interest to explain why we do suffer from the illusion. The only difference with the case of *philosophical* intuitions is that we don’t take our visual perceptions as veridical.

That is, it is no help to be told that someone else doesn’t share your puzzlement. Who cares what the Mayans or Chinese think? Their failure to be puzzled doesn’t help resolve our problems. If I am the only one who is guilty of confirmation bias or base rate neglect, I need diagnosis and a cure, not anthropology. Even if it is parochial to Western departments of analytic philosophy, the central problem remains for Kripke and those who do, as a matter of fact, share the intuitions in question. Moreover, the anthropological evidence of cross-cultural variation does not
illuminate the fundamental question because, even if the Kripke-Putnam intuitions were universally shared, their credentials are not thereby established as guides to scientific or metaphysical claims.

**Who is in the Know?**

Putnam (1975, 11) explains that internally identical “water” thoughts are said to have different meaning on Earth and Twin-Earth, although the chemistry of H2O or XYZ may never be discovered by people on either planet. That is, externalism depends on intuitions arising from the theorist's knowing the truth. Indeed, defending externalism, Burge (1988) confirms this diagnosis saying “We take up a perspective on ourselves from the outside.” The conception of an “Omniscient Observer” is explicitly embraced as unproblematic by Donnellan (1974), a perspective Kaplan (2012, 156), too, has endorsed as “description from above.” This is an understanding “in which one surveys another’s thought” from a point of view “independent of whether the subject’s thought corresponds to reality.” These are remarkable confirmations of my diagnosis of the illusion of the “philosopher-narrators” omniscience.

In Crane’s (1996) useful phrase, the question of who is “in the know” is central to untangling the intuitions at the heart of puzzles concerning externalism. Crane’s question recalls Putnam’s (1981, 50) question “From whose point of view is the story being told?” The invisibility of our own role and our own knowledge creates the illusion that it is the relational fact about how the world really is that determines the thought or belief in question. As Crane (1996, 293) notes, “the Twin Earth cases are meant to demonstrate that the world itself can, as it were, fix the meanings of some of our words.” Crane’s apt characterization captures the paranormal or clairvoyant conception of meanings which somehow link the mind directly with its objects in the world.

**Philosophers, autistics & three year olds**

Burge (2012, 119) recently explains the nature of _de re_ belief in terms that are suggestive of other philosophical problems: “One can have a _de re_ belief that is successfully referential and meets all other conditions on being _de re_, which nevertheless fails to count as knowledge.” Consider the case in which someone is looking at a chair which he can see in a certain position apparently in the next room. However, he doesn’t notice that he is looking at a large mirror and, therefore, sees the reflection of a chair that is actually nearer to him in the same room. As it happens, there is an identical chair in the next room behind the mirror, exactly where the reflection appears to be. It is evident that this circumstance is precisely Burge’s scenario of _de re_ belief and it is also exactly the Gettier (1963) case of justified, true belief that doesn’t count as knowledge. Burge doesn’t mention Gettier, but these parallels suggest the Problem has a wider interest beyond the epistemological issues it has been directly concerned with. Accordingly, it is interesting to notice Fodor’s comment about the semantics of mental representations applies to Gettier too: “we need it [broad or externally individuated content] to make sense of the fact that thoughts have the truth conditions that they do” (1994, 50). As if describing the Gettier Problem, in an entirely different context, Fodor gives a diagnosis that is apt for this puzzle:

> It is, to put the point starkly, the heart of externalism that **semantics isn’t part of psychology.** The content of your thoughts (utterances), unlike for example, the syntax of your thoughts (utterances), does not supervene on your mental processes. (Fodor 1994, 38)

In the Gettier case, too, the wide contents of your thoughts construed transparently as knowledge do not supervene on your mental processes, being merely justified beliefs. Fodor had made the same point where he said “truth, reference and the rest of the semantic notions aren’t psychological categories” (1980, 253).

In response to the semantic orthodoxy, Farkas’ (2003) argues that “external features are important only if they are incorporated into the internal cognitive or experiential perspective of cognizers.” Schantz, too, explains, “As far as psychological explanation is concerned, what counts is how the world is internally represented as being, **not how the world really is**” (2004, 23; emphasis added). This is essentially the formula with which Fodor (1998, 20) characterized externalism, the view that “what you are thinking depends on what world you’re in.” This diagnosis of externalist semantic intuitions is precisely appropriate to the notorious Gettier (1963) Problem. In Chisholm’s (1966) classical version, the subject sees a sheep-like bush and acquires a perceptual belief “There is a sheep in the field.” Although this belief is justified by the evidence, it is true only by accident because, unbeknownst to him, there is a sheep elsewhere in the field. The classical criteria for knowledge – justified, true belief – appear to be met, but the belief does not count as knowledge. Hetherington (2012) has recently given an analysis of “Gettiered beliefs”, being cases in which “truth remains essential.” His diagnosis is that philosophers’ intuitions are evidence of their “being infallibilists, without realizing this about themselves.” This seems to be another way of making my point about puzzles that arise from the “narrator’s” omniscience. Putnam’s Twin Earth example, too, is a case in which mental content is ascribed to someone on the basis of truths that are not represented internally by the subject just as in the Gettier Problem.

Schantz’ prescription for psychological explanation – what counts is how the world is internally represented as being, **not how the world really is** – is apt also for capturing the mistaken “theory of mind” in a different domain. We see a striking analogue known to clinical psychologists in the Wimmer and Perner (1983) “false belief” task: Autistics and three year-olds ascribe beliefs to others based on their own knowledge of the truth rather than on the other's justified beliefs. Switching the candy when the character isn’t looking in the experiments of Wimmer and Perner is
analogous to Gettier’s substitution of bushes for sheep, or Fodor’s substitution of shrews for mice in cases of misrepresentation. Putnam’s substitution of XYZ for H₂O, like Dretske’s (1986) disoriented microbes, are various ways that have been devised to make ‘the world go wrong.’ The truth-making facts are unconnected for the grounds for belief which are known only to the philosopher-narrator. By ascribing beliefs in this way, it appears that philosophers make the same mistake that autistics commit and children grow out of by the age of four. Ralph’s belief about Orpwood (Quine 1960) just like Twin Oscar’s thoughts about water (Putnam 1975) and Pierre’s thoughts of London (Kripke 1979) are essentially ascriptions of belief based on the philosophers’ knowledge of the truth (see Slezak 2011).

**Obscurantist Intentional Magic**

“Object-dependent” referential thoughts called *de re* are taken to be “singular thoughts” about a particular object or person that the speaker has in mind. This is the strong intuition expressed by Brian C. Smith that symbols somehow “reach out and touch someone” (1987, 215). Kripke has placed these issues in his framework of ‘rigid designators’ that denote the same individual in all ‘possible worlds.’ However, Stalnaker (2003) emphasizes that Kripke’s claims rest on intuitive grounds, and poses a revealing question: “Doesn’t this presuppose that the same individuals can be found in different possible worlds? Searle (1969, 93), too, argues that if an expression has no descriptive content as Kripke and ‘direct reference’ theorists claim, “then there could be no way of establishing a connection between the expression and the object.” He asks “What makes this expression refer to that object?” Kripke’s preferred answer is that a chain of historical, causal, connections back to a baptismal event fixes the reference. However, this account utterly fails to explain how a particular individual acquires the competent use of a name. The point is precisely analogous to Putnam’s (1967, 18) attempt to rebut Chomsky’s “innateness” claims by citing the normal historical origin of all human languages. But this response fails to address the problem of language acquisition – the question of how each individual child must accomplish the task of becoming a competent speaker. The common origin of all human languages is irrelevant to this question, just as the supposed historical-causal chain is irrelevant to an individual’s understanding and use of proper names. Stalnaker (2003, 178) captures the problem aptly, speaking of the only alternative to descriptive accounts which seems to be “some kind of obscurantist intentional magic.” In Searle’s (1969, 87) suggestive words, the idea that we can mean or intend a particular object and not another inclines us to think “that it is a movement of the soul.” In the same vein, Putnam (1981) suggests that externalist intuitions are a “magical theory of reference” that assumes occult “noetic rays” connecting words with their referents. Indeed, these referential intuitions are suggestive of widely held, compelling misconceptions concerning visual perception that are thought to involve emanations from the eyes – the so-called “extramission theory of perception” maintained by early Greek philosophers. Remarkably, following Piaget, Winer et al. (2002) report evidence that belief in extramission remains widespread, deeply ingrained and resistant to educational efforts. I don’t mean to suggest that such theories are literally believed by philosophers, but the compelling conceptions are very suggestive of intuitions underlying the most widely held externalist semantic theories.

**References**


Devitt, M. 1984. Thoughts and their Ascription, Peter A. French, Theodore E. Uehling and Howard K. Wettstein