

Explaining Reference: A Plea for Semantic Psychologism¹

SANTIAGO ECHEVERRI

There is a traditional opposition between two camps in the theory of reference: descriptivism *vs.* referentialism or Fregeanism *vs.* Russellianism. Before one chooses which view—if any—one favors, a more fundamental issue should be addressed: Can we *explain* reference? Whereas ‘full-blooded’ theorists answer ‘yes,’ so-called modest theorists answer ‘no.’²

Since the notion of *explanation* can be understood in different ways, there is no single full-blooded/modest contrast that describes all possible views. One could be a full-blooded theorist in relation to one of these concepts of explanation but a modest one in relation to the others.

One might mean by ‘explanation’ some form of reductive *conceptual analysis*. So the explanatory task would consist in offering a non-circular analysis of key concepts like reference, denotation or truth (Dummett 1975, 1976, 1991; Engel 1989, 1994).

The notion of explanation might also amount to an *ontological reduction* of semantic properties to non-semantic properties like causation, information, biological functions, and so on (Field 1972, 1978; Dretske 1981; Millikan 1984).

¹ Jérôme Dokic read a previous version of this paper, and made some insightful comments. Work on this article was funded by research grant no. 100015_131794 of the Swiss National Science Foundation.

² Dummett (1975, 1976, 1991) first defended (a version of) the full-blooded theory of meaning. ‘Modest’ theories are associated with Davidson (1984) and McDowell (1977, 1987, 1997), among others.

One might also drop reductionism, and conceive of the explanatory program as the enrichment of a purely extensional semantic framework with intensions, characters, possible worlds, etc. (Engel 1989).

Still, one might take the explanatory program as a meta-semantic one seeking to uncover the *historical* and *social facts* that explain why some expressions have the meaning they presently have (Devitt 1981; Kaplan 1989; Almog 2005).

Some of these ways of cashing out the full-blooded program bear interesting relations to each other. If you are a conceptual reductionist about reference, you will probably reject historical accounts of reference such as Kripke's (1980), which uses notions like meaning intentions to account for the transmission of reference (Dummett 1973: 148-9; McKinsey 2009). It is however desirable to keep these different concepts of explanation separate before one makes up one's mind on whether one should try to explain reference.

Pascal Engel has rejected various forms of modesty during his prolific career. Yet, given the different ways in which one can construe the notion of explanation, it would be misleading to qualify him as a full-blooded theorist. He has been sympathetic to Davidson's anomalism of the mental, which implies that intentional properties cannot be reduced to physical properties, and has also expressed serious doubts on teleosemantic accounts of semantic content (Engel 1996). Although he has sometimes presented the full-blooded program as some version of conceptual analysis (Engel 1989, Ms.), I would be reluctant to describe him as favoring some form of non-circular factorization of the concept of reference. Still, Engel has been a critic of the sort of modesty championed by philosophers like McDowell and some Wittgensteinian philosophers (Engel 1996, 2001). So, if Engel is a full-blooded theorist, his 'full-bloodedness' does not apply to all the notions of explanation that have been used to characterize full-blooded positions.

An important strand of Engel's critique of modesty can be found in his works on the philosophy of logic. In his seminal book *La norme du vrai: Philosophie de la logique* (1989), he used the debate between Davidson and Dummett on the nature and scope of a theory of meaning as the bedrock to introduce a number of topics in the philosophy of logic. In that book, he granted some of Davidson's constraints on a good theory of meaning, and defended a weak form of psychologism. In the next years, he pursued these two issues in more depth. In his thèse d'état *Davidson et la philosophie du langage* (1994), Engel offered a systematic reconstruction of Davidson's program, arguing that it is compatible with a psychological account of language mastery in the spirit of Chomsky's theory of tacit knowledge. Two years later, he published *Philoso-*

phie et psychologie (1996), a sustained attack on the dogma that psychologism must be avoided at any cost. This led him to defend a 'healthy psychologism,' and reject the pervasive idea that the realm of norms is wholly disconnected from the realm of causes.

All these works constitute an attempt at finding a middle way between some form of modesty and what I would call 'psychological full-bloodedness.' Engel offers a nice formulation of his proposal within Frege's distinction between the world of psychological facts (world 2) and the world of objective truths (world 3). According to Engel (1996: 121-ff.), it makes good sense to introduce a world 2½, i.e. a place where normative and natural properties are in contact. The notion of a world 2½ expresses his conviction that any theory of norms should accommodate two seemingly conflicting data. On the one hand, the psychology of reasoning has shown that subjects may reason in ways that contradict the rules of classical logic. On the other, it is clear that the same subjects who seem to reason in a non-classical way developed classical logical systems, and used them to evaluate and criticize their own patterns of reasoning (Engel 1989: Chapter XIII; 2006). Hence, world 2½ expresses the idea that, although logic does not *describe* the actual procedures used by ordinary subjects in normal reasoning, it is not part of a disconnected 'third realm' of objective truths. Engel sees world 2½ as enabling us to specify the kinds of processes that ought to govern the mental life of humans. But the idea goes beyond that: It suggests that the links between the psychological and the logical are intimate but less direct than some forms of psychologism might take them to be. Hence the label 'healthy psychologism.'

The topic of this paper is not the philosophy of logic but the theory of reference. Yet, my project is tightly related to Engel's lucid defense of world 2½. I do not have any reductionist agenda, nor do I believe in the program of conceptual analysis. Moreover, I will not give arguments for the enrichment of semantics with intensions or possible worlds, nor for a historical account of reference. Still, I will defend a healthy psychologism in the theory of reference. Thus, I might be taken to be pursuing a full-blooded program. I call this brand of the full-blooded program 'semantic psychologism.' Semantic psychologism is the conjunction of a negative and a positive claim:

Negative claim: Pursuing the program of truth-conditional semantics is insufficient to account for semantic competence, for it does not provide a sufficiently illuminating account of referential abilities.

Positive claim: There are good reasons to supplement the program

of truth-conditional semantics with a *psychological* account of referential abilities, i.e. an account that is framed at an intermediary level of description between the personal level and the explanations provided by neuroscience.

As I understand it, semantic psychologism is orthogonal to any historical or social account of reference. Whereas the latter deal with reference *qua* property of expressions in a public language, semantic psychologism seeks to provide an account of the abilities needed to make use of a language. In other words, one does not need to see historical or social accounts of language as *competitors* to semantic psychologism. Its real opponent is the anti-psychologism that dominates various forms of philosophy of language that construe it as a branch of the philosophy of logic or reject any conception of psychology as an explanatory discipline.

My defense of semantic psychologism will have two parts. First, I will present some of the main reasons why the program of truth-conditional semantics is modest in a relevant sense: it does not offer the explanations we want when we are interested in offering an account of referential abilities (sections 1-2). Second, I will respond to some influential considerations against a psychological explanation of reference (sections 3-8).

1. The Explanatory Limits of Truth-Conditional Semantics

Frege's doctrine of *Sinn* is the first formulation of a truth-conditional semantics. Some philosophers reject the notion of *Sinn*, however, because of its dubious ontological pedigree. Yet, one can partially circumvent this problem by pursuing a 'functional' approach. On a functional view, the notion of *Sinn* is introduced by its *theoretical role*. The *Sinn* of an expression *E* is construed as a *solution* to some of the key questions that arise in the study of *E*.³

The central role of a theory of *Sinn* is to provide a truth-conditional account of the meaning of whole sentences. On this approach, the *Sinn* of a whole sentence *S* (a *Gedanke*) expresses its truth-conditions, and the *Sinn* of its constituent expressions $E_1, E_2, \dots, E_{n-1}, E_n$ is their contribution to the truth-conditions of *S*.

How does the truth-conditional role of *Sinn* relate to the more general debate that opposes full-blooded to modest conceptions of meaning? My first claim is that, even if the truth-conditional aspect is a necessary ingredient of a

³ This approach is implicit in Burge (1977: 59) and McDowell (2005).

full-blooded theory, it is not sufficient to explain reference. There are at least two reasons why offering a truth-conditional account of language is not sufficient to explain reference. First, a theory that only specifies truth-conditions would leave out some central *explananda* in the theory of meaning. Second, truth-conditional semantics exploits a very thin notion of explanation, i.e. a notion that is not sufficiently illuminating for philosophical purposes. I develop these two ideas in the next sub-sections.

The Determination of Reference

If one takes the notion of *Sinn* as spelling out the contribution of any expression-type E_i to the truth-conditions of sentences in which E_i may occur, it seems possible to frame Frege's theory of *Sinn* as an axiomatic theory.⁴ The main textual support for this reading can be found in § 32 of *Grundgesetze der Arithmetik*, where Frege makes clear that his stipulations of the meanings of names convey their *Sinn*.⁵ Building on this minimal conception of *Sinn*, one can specify the contribution of singular terms and predicate expressions by means of a list of axioms. Let us consider the following two examples:

(A1) 'Hesperus' stands for (or denotes) Hesperus.

(A2) 'x is agile' is true of something if and only if it is agile.⁶

These axioms satisfy the first characterization of the notion of *Sinn*. Yet, there is a sense in which they are not explanatory. Given that 'stands for' and 'is true of' are *used* in the axioms, these specifications do not explain *how* 'Hesperus' refers to Hesperus, nor *why* the predicate 'x is agile' is true of some things but not of others. Dummett (1975) offers an influential argument in favor of this claim: one could imagine a subject who understands quotation devices but is unable to *use* 'Hesperus' to refer to Venus, and could not *identify* Venus in the sky. So, even if the first axiom articulates the contribution of the name 'Hesperus' to the truth-conditions of the sentences in which it may occur, it is not sufficient to explain what it is to *understand* the word 'Hesperus.' (See also Engel 1989)

⁴ See McDowell (1977), Evans (1981a, 1981b), and Sainsbury (2005). This notion fits Lewis' (1980) characterization of 'semantic value': "If they [semantic values] don't obey the compositional principle, they are not what I call semantic values." (35)

⁵ See Frege (1893: § 32, 50-1). For a recent reading in this sense, see Kripke (2008: 182-ff.).

⁶ See McDowell (1977, 1997). I set aside some complications with predicates. One should allow their argument places to be saturated either by names or variables. I also omit the qualification that these axioms are relative to a particular language, in this case, English.

A highly influential way of understanding the contrast between modest and full-blooded theories of meaning is therefore to claim that, for modesty, it is not possible to provide a more fundamental characterization of reference than the one provided by axioms that use semantic expressions like 'denote,' 'refer,' 'stand for,' and so on. If one thinks that truth-conditional specifications are all there is to explain the reference of the primitive vocabulary of a language, one holds a modest view. By contrast, if one thinks that one can provide an informative account of the notions of denotation, reference, satisfaction, etc., one will hold a full-blooded view.

A popular way of cashing out this distinction is to say that Frege's notion of *Sinn* not only plays a truth-conditional role but also more substantial roles. One of these roles is to offer a *method* by which speakers can determine the semantic value of linguistic expressions.⁷ Thus, the notion of *Sinn* can be seen in the service of another question: What connects the name with the referent? Or, as Almog puts it: What is the chemistry of the bond between a singular expression and the referent?⁸ If one thinks these questions are meaningful, one must find the axioms stated above too austere. In order to provide explanations, one cannot just *state* the axioms of the semantic theory; one has to *explain* how the names mentioned and used in these axioms connect to entities in the world.⁹

Although this approach to the non-sufficiency claim is widespread, it leaves a number of questions open. First, the equation of the notion of *Sinn* with a method has some verificationist connotations, and might convey an unfaithful idea of the referential use of words. When I use a telescope to watch the craters in the moon, I certainly use a method to single out the craters. But it is unclear whether language or concepts work as a method in this instrumental sense. Second, the equation of the notion of *Sinn* with a method raises questions concerning the links between language and thought. If one assumes that the axiomatic theory is a model of *public* language, one shall probably need to distinguish the *public Sinn* articulated in the axiomatic reconstruction of a language from the *private means* speakers use to determine the semantic values of words. As far as proper names are concerned, some theorists have held that speakers use *different routes* to determine their semantic values, even though

⁷ See Dummett (1973: 93, 95; 1978: 119-20). For a critical discussion, see Evans (1982: 17-ff.).

⁸ See Almog (2005: 498) and Wettstein (2004: 110).

⁹ I am assuming that the notion of explanation is psychological. As indicated in the introduction, this is not the only relevant notion of explanation. One might require reductive, historical or social explanations. For reasons of space, I will not deal with these other options here.

there would be a 'shared' or 'public' *Sinn* of those words.¹⁰ I won't explore these possibilities here.¹¹

Truth-Conditional Semantics as Instrumental Explanation

Another way of showing that the truth-conditional approach is not sufficient to explain reference is to focus on its underlying notion of explanation, and show that it is not sufficiently illuminating from a philosophical perspective. When one lays down the axioms specifying the semantic value of an expression-type E_i , one is assuming that the axioms have 'projectible predicates' in Goodman's (1965) sense. These predicates occur in inductive statements, so they express counterfactual-supporting generalizations.¹² Thus, contrary to what might be initially thought, when one lays down the axioms of a fragment of a language, one is not merely *describing* it. If the axioms are compositional, they constitute a system of interrelated principles (Davies 1987).

The explanations provided by the axioms can be controlled for their correctness. On the one hand, theorists can check whether their assignments *cohere* with speakers' intuitions on the truth-conditions of sentences. On the other, those assignments enable theorists to make some *predictions* on the behavior of expression-types. Hence, one can compare competing axioms by their 'explanatory potential'; one has only to examine their intuitive adequacy and the predictions they make. This approach is instrumentalist because semantic theories work as predictive devices. The axioms capture projectible *regularities* in the behavior of expression-types.

Normally, issues of validity provide the main tests of these investigations: One asks which inferences are made valid under a semantic assignment, and

¹⁰ As a result, Burge (1990) distinguishes objective sense from the subject's grasp of sense. Since a speaker might have a partial (or mistaken) grasp of sense, this would lead to a further distinction between the idiosyncratic and the public aspects of language. For a related distinction, see Higginbotham (1998). In a different perspective, Millikan (2005) distinguishes the conventional meaning of words from subjects' conceptions, which correspond to the methods a speaker uses to track the referent.

¹¹ See Frege (1918-1919). On the basis of this text, Kripke (1979, 2008) urges that, as far as proper names are concerned, Frege is committed to idiolects (not public languages) as the main units of analysis.

¹² There is a stronger conception of *projectible* predicates as those that enter into *laws*. But it would be controversial to assume that truth-conditional semantics formulates laws in any non-trivial sense of that term. It is less controversial, however, that truth-conditional semantics treats expression-types as *natural kinds*.

which turn out to be invalid.¹³ Crucially, one can reject a semantic proposal if it fails to accommodate intuitively valid patterns of reasoning or if it counts as valid some pieces of reasoning most competent speakers intuitively count as invalid. The following inference illustrates this strategy:

P1 Jones believes that all men are created equal.

P2 Smith doubts that all men are created equal.

∴ Thus, there is something that Jones believes but Smith doubts.

The instrumentalist theorist is interested in laying down semantic rules capable of specifying the behavior of 'that'-clauses in such a way that inferences of this sort go through. Crucially, one could discard some accounts because they cannot accommodate this intuitively valid inference. An influential solution is to treat the 'that'-clause as a sentential operator. When it is conjoined with a sentence, it forms a singular term that designates a proposition. This view yields a valid inference. Let us use 'S' as a variable ranging over persons, 'B' as a belief operator, 'D' as a doubt operator, and 'p' as a variable for propositions. So we have:

P1 B(s, p)

P2 D(s, p)

∴ $\exists x: x$ is a proposition p & (B(s, x) & D(s, x))¹⁴

If this sort of explanation is all there is to explain reference, one may count as a modest theorist. I assume Engel would agree with this. In a recent review of LePore & Ludwig's *Davidson's Truth-Theoretic Semantics*, he praises the development of specific proposals within Davidson's semantic program. Nevertheless, he also expresses "some nostalgia for the pioneering efforts of the 1970s." (Engel 2007) As philosophers, we may want to understand how such concepts as reference and truth fit within a natural world. If we are interested in linguistic competence, we may also want to tell a story about the referential abilities underlying the use of linguistic expressions. In the next section, I develop this point in some detail.

¹³ Other authors have drawn similar distinctions. Rorty (1979: Chapter 6) calls 'pure semantics' a project similar to what I am calling instrumentalism, and contrasts it with what he calls 'impure semantics.' Barwise & Perry (1983: lxxv) term 'thin semantics' a project similar to instrumentalism. Recently, Predelli (2005: Chapters 1 and 4) mounted a defense of *thin* semantics.

¹⁴ I have borrowed the example from Salmon (1986: 6). This kind of reasoning is pervasive in contemporary philosophy of language.

2. The Psychological Program

Let us take stock. Spelling out the contribution of an expression-type E to the truth-conditions of sentences in which E may occur might be a necessary component of a theory of reference.¹⁵ Yet, it is insufficient. First, the truth-conditional account would leave out the determination of reference and, second, it would rely on a very ‘thin’ conception of explanation as prediction. If we are interested in linguistic competence, we might be interested in elucidating the psychological abilities underlying the mastery of a language.

In this section, I focus on the expansion of the instrumentalist program by means of a psychological account of referential abilities. I will remain neutral on the exact form of such a psychological program. I will simply show that the instrumentalist notion of explanation at work in truth-conditional semantics presupposes the possession of more fundamental referential abilities that may require a psychological account.

One can clarify the notion of explanation proper to semantic psychologism by examining the requirements to *determine* the referent of any expression-type. In the context of a truth-conditional reconstruction of a fragment of a language, there is a sense in which the semanticist determines the semantic value of an expression-type such as a proper name ‘NN.’ In this case, to determine the semantic value of ‘NN’ is to furnish a mathematical function that maps occurrences of ‘NN’ onto its semantic value, e.g. a referent. This sense of reference determination is implicit in some of Frege’s remarks on the notion of *Sinn*. Consider the following excerpt from “Funktion und Begriff”:

It is thus necessary to lay down rules from which it follows, e.g., what ‘ $\odot + 1$ ’ stands for, if ‘ \odot ’ should (*soll*) stand for the Sun. What stipulations we lay down is indifferent; but it is essential that we should do so—that ‘ $a + b$ ’ should always get a *Bedeutung*, whatever signs for determinate objects may be inserted in the place of ‘ a ’ and ‘ b .’ (Frege 1891: 19-20; translation modified)

According to Frege’s example, reference determination can be a stipulation. In a formal system, one can determine the referent of ‘[F080?]’ by providing a rule that assigns to it one and only one referent. To determine a referent of a sign is an *obligation* any theorist incurs when she is laying down the axioms

¹⁵ It would take us too far afield to examine the reasons why most theorists take the truth-conditional program as a necessary component of a systematic theory of meaning. My main concern here is with the non-sufficiency claim.

of the system (Frege 1893: XII; see also Heck 2002: 4). This obligation is also incurred when one tries to formalize a fragment of a natural language. The stipulative sense also captures some of the ways in which Kripke illustrates his notion of 'reference fixing.' He grants that, at least in some cases, one can perform some ceremonies in which one stipulates the meaning of a proper name, as when a policeman in London declares: "By 'Jack the Ripper' I mean the man, whoever he is, who committed all *these* murders, or most of them." (Kripke 1980: 79)¹⁶

In order to stipulate the referent of an expression, one needs an *independent way* of referring to the target referent. Thus, when Frege stipulates that '[F080?]' stands for the sun, he assumes that the audience has independent means to single out the sun, e.g. iconic memories of the sun and the word 'sun.' Similarly, in order to stipulate that 'Jack the Ripper' refers to the man, whoever he is, who committed all these murders or most of them, the policeman must have an independent way of referring to that murderer. This clearly shows that explaining the determination of reference does not stop when the semanticist (or the policeman) 'fixes' the referent of a name in the stipulative sense. We still need an account of the *prior* referential abilities that enable them to make some stipulations, and the audience to understand those stipulations. This asks for a different kind of explanation.

This remark suggests that something along Engel's lines could be correct in the theory of reference. Recall that Engel advocates a healthy psychologism, and believes that the notion of tacit knowledge employed in cognitive science is more respectable than its Wittgensteinian critics assume. Yet, one might wonder whether the previous remark would be sufficient to *vindicate* the project of providing a *psychological* account of referential abilities. After all, from the fact that semantic stipulations presuppose prior referential abilities it does not follow that one can *explain* those abilities. Whereas a semantic psychologist might be optimistic about the prospects of the explanatory task, the semantic anti-psychologist might be inclined to take those referential abilities as primitive. In what follows, I defend semantic psychologism. To this end, I respond to an influential series of considerations against it. I examine the arguments in order of increasing relevance and plausibility.

¹⁶ See also Evans' (1982: 50) example of Julius, the inventor of the zipper, and Kripke's (1980: 55) remarks on the standard meter in Paris.

3. Our Present State of Ignorance

In his influential paper "Index, Context, and Content," Lewis describes a grammar "as part of a systematic restatement of our common knowledge about our practices of linguistic communication." (1980: 21) Later on, he makes clear that he does not envisage providing a psycholinguistic theory:

The subject might be differently delineated, and more stringent conditions of adequacy might be demanded. You might insist that a good grammar should be suited to fit a psycholinguistic theory that goes beyond our common knowledge and explains the inner mechanisms that make our practice possible. There is nothing wrong in principle with this ambitious goal, but I doubt that it is worthwhile to pursue it in our present state of knowledge. (Lewis 1980: 24)

Lewis' remark suggests that our future state of knowledge might enable us to offer an explanation of the inner mechanisms "that make our practice possible." Still, he also thinks that our present state of knowledge does not warrant any psychologically oriented formulation of a theory of a language.

There is a sense in which Lewis' remark is right: we lack the required *empirical* evidence to formulate a detailed psychological account of referential abilities. Nevertheless, this is not sufficient reason to set aside semantic psychologism. After all, any explanatory program requires prior conceptual ground clearing, and philosophers are particularly good at that task. Moreover, science requires the formulation of theories capable of organizing the available findings and generating new predictions. So, it would be a mistake to 'wait' until we gather more data before we decide to pursue semantic psychologism. If our goal is to offer a psychological account of referential abilities, it is already necessary to make some conceptual work to guide empirical research.

The philosophically interesting question is: What general form could a future cognitive theory of referential abilities take? There are at least two ways of answering this question. According to optimism, there will be a *psychological* theory of reference. According to pessimism, there will be no such thing as a psychological theory of reference. The only cognitive account we will get, if we get anything properly called 'cognitive,' will be provided by neuroscience.

It is the latter possibility that threatens semantic psychologism, for it undermines it on principled grounds. And, if there cannot be a psychological account of referential abilities, psychological modesty is not only reasonable

but also mandatory. In what follows, I respond to some in-principle arguments against semantic psychologism.¹⁷

4. The Quietist Stance

A prominent line of attack derives from a quietist conception of philosophy. On this view, providing a psychological account of reference is a bad idea because philosophy is not a theoretical enterprise. This outlook is implicit in Kripke's version of the historical account of reference. He insists that reference is maintained if speakers *intend* to use names with the same referent. Hence, his picture does not explain the required referential intentions. Crucially, this is a problem only if one evaluates the historical view as a theory. He dismisses this reading, though: "You may suspect me of proposing another theory in [the] place [of the cluster theory of names]; but I hope not, because I'm sure it's wrong too if it is a theory." (Kripke 1980: 64)¹⁸

I am unable to see how I could grant a conclusion based on overarching premises on what philosophy is (or ought to be). But maybe some philosophical issues cannot be adequately tackled by constructing theories. What is special about reference that might preclude the formulation of psychological theories thereof? An extreme view says that the problem is not related to reference *per se* but to the project of formulating a theory of the *mind*. On this view, psychology is not a science of the mind in the same way as astrology is not a science of destiny. There is a version of this claim in some of McDowell's seminal remarks in favor of a modest account of reference:

There is no merit in a conception of the mind that permits us to speculate about its states, conceived as states of a hypothesized mechanism, with a breezy lack of concern for facts about explicit awareness. Postulation of implicit knowledge for such allegedly explanatory purposes sheds not scientific light but philosophical darkness. (McDowell 1977: 180)

A less extreme view holds that, although there is a psychological science of the mind, there is no *mechanistic explanation* of referential abilities within a science

¹⁷ The following discussion is congenial to Engel's (1996) lucid defence of a healthy psychologism.

¹⁸ I owe this point to McDowell (1977: 198). In his paper "Speaking of Nothing," Donnellan (1974: n 3) is quite clear that he "wants to avoid a seeming commitment to all the links in the referential chain being causal." As Wettstein points out, we should distinguish the idea of a chain of communication from the more committal idea of a causal theory of reference.

of the mind. Interestingly, a prominent cognitive scientist, Zenon Pylyshyn, holds this view. He thinks that the mind requires some form of *direct reference* analogous to an index. Yet, he also thinks that there is no cognitive account of how indices refer. This leads him to hypothesize that a correct account of how indices work falls “under an architectural or neuroscience vocabulary.” (Pylyshyn 2007: 39, 82)

Unfortunately, Pylyshyn’s view does not lend support to McDowell’s more radical statement, which impugns the very idea of hypothesizing mental mechanisms. McDowell (and many others) has a general picture of the mind that prevents him from approving the development of psychological accounts of reference. The next sections explore the most prominent ways of defending this form of semantic anti-psychologism.

5. Reference and the Vehicle-Content Distinction

In subsequent work, McDowell presents a more specific attack on the program of explaining reference. The attack does not have the form of a rigorous argument but is offered as a collection of suggestive considerations.

Here is the main line of thought. In order to formulate the problem of reference, one has to introduce a dichotomy between two aspects of meaningful entities: a *physical* and a *semantic* aspect. Once a sharp line between these two aspects is drawn, we create a gap that cannot be bridged. Since the gap cannot be bridged, we are left with the feeling that reference is a very deep problem that lacks any intelligible solution. If we reject the underlying dichotomy, however, the problem of reference should not arise. Thus, instead of trying to *solve* the problem of reference, we should try to *dissolve* it by rejecting the dichotomy presupposed in its formulation. We find these considerations in McDowell’s comments on Putnam:

Putnam has often expressed suspicion of the idea that there is good philosophy to be done by grappling with questions like ‘How does language hook on to the world?’ It ought to be similar with questions like ‘How does thinking hook on to the world?’ Such a question looks like a pressing one if we saddle ourselves with a conception of what thinking is, considered in itself, that deprives thinking of its characteristic bearing on the world—its bearing about this or that object in the world, and its being to the effect that this or that state of affairs obtains in the world. If we start from a conception of thinking as in itself without referential bearing on the world, we

shall seem to be confronted with a genuine and urgent task, that of reinstating into our picture the way thinking is directed at the world. But if we do not accept the assumption that what thinking is, considered in itself, is a mental manipulation of representations in Putnam's sense [as vehicles], no such task confronts us. The need to construct a theoretical 'hook' to link thinking to the world does not arise, because if it is thinking that we have in view at all—say being struck by the thought that one hears the sound of water dripping—then what we have in view is *already* hooked on to the world; it is already in view as possessing referential directedness at reality. (McDowell 1992: 288)

The problem of explaining reference arises from a *dualistic* understanding of the distinction between the vehicles of representation and their contents. On this view, one could eventually have a physical vehicle, let us say the ink mark '**Aristotle**,' which could fail to refer to the famous philosopher.¹⁹ This makes the problem of reference puzzling. If that ink mark is *just* a physical pattern with no referential power, how can it refer to Aristotle? If one understands the relation between vehicle and content as intrinsic, however, the question appears to be empty.

Consider an analogy. One could use a piece of bronze to make a statue but also to create many other things. Still, there seems to be a difference between one's having a piece of bronze and one's having a statue. How is it possible that a piece of bronze can also be a statue? This sort of problem seems to rest upon a mistaken assumption on how the piece of bronze is related to the statue. Although the latter is made of bronze, the constitution of the statue *qua* statue does not depend on its matter alone but also on its design. Since a description of the piece of bronze as a statue belongs to a specific level of description, it makes little sense to ask a question about the constitution of the statue by remaining at the lower level of description that considers the statue as a mere piece of bronze. By parity of reasoning, if you focus on the purely physical side of any sign, you will be unable to explain how it manages to refer to something in the world. If you focus on the vehicle *qua* bearer of meaning, however, the problem of reference should not arise.

Howard Wettstein develops a similar line of thought. He compares the problem of reference to Descartes' puzzlement about the locomotive capacities of some bodies:

¹⁹ I am using '**boldface**' quotation to refer to types of physical entities.

Descartes says that he found it amazing that bodies, mere pieces of nature, could move themselves. If locomotion can seem miraculous, what about reference? That mere pieces of nature can mean, or symbolize, or stand for something really seems extraordinary. (Wettstein 2004: 104)

According to Wettstein, Descartes' puzzlement is based on a dualistic assumption. It is certainly extraordinary that animals are capable of locomotion. Still, if we describe their bodies as *merely extended entities* (like stones or pieces of clay), there is nothing we could 'add' to explain how they self-move. That is probably why one might be led to posit a mysterious soul to explain their locomotive capacities. But positing a soul merely explains the obscure by the more obscure.

Similarly, if we describe sound patterns as *mere* physical vehicles, how can we explain reference without adding something similar to a soul? Only by adding something intrinsically significant, such as a Fregean sense or a description, can we explain their reference. But this strategy will be explanatory only if we already understand what it is for an immaterial entity to be endowed with *intrinsic meaning*. And the same will occur if we try to derive reference from mental states. Our account will only work if we are prepared to swallow the idea of an *intrinsic* intentionality (Searle 1983). But very few people seem to understand what it means to be intrinsically intentional (see, e.g., Clark 2005). Wettstein exploits these considerations to promote philosophical modesty:

Why not leave things where we found them? This suggests—and this might seem at least mildly depressing (but not to worry, it grows on one)—that perhaps the best we can do is to describe our ways with language, making no attempt to go beyond or behind. (Wettstein 2004: 106)

One might wonder whether this argument *shows* that semantic psychologism is hopeless. I do not think so. It merely *suggests* that, given a particular metaphysics of mind and language, one cannot provide an intelligible account of reference. But this falls short of undermining semantic psychologism, or so I shall argue.

Most contemporary philosophers certainly think that there is a vehicle-content distinction in the mental realm (but see Sedivy 2004, for criticism). This might lead one to think that this distinction is a necessary presupposition

of the problem of reference. Nevertheless, one can still formulate the problem of reference without presupposing a conception of the mind that takes that distinction for granted. Consider first the case of public language. It is an uncontroversial fact that public signs are *conventional*. Because signs are conventional, it is legitimate to draw the vehicle-content distinction to characterize them. One can use different physical marks such as '**red**,' '**rot**,' '**rouge**,' '**rojo**' to denote one and the same property: REDNESS. This clearly suggests that signs are *arbitrarily related* to their content. Meaning is not an intrinsic property of any physical shape. Hence, trying to explain this relation, as full-blooded theorists try to do, is legitimate. For any word '*W*' one may ask: How did the physical pattern '*W*' come to denote the entity it denotes (or express the property it expresses)? What conditions must obtain for a physical pattern to be meaningful?

Similar remarks apply in the cognitive realm. Even if one rejects the application of the vehicle-content distinction to the mental realm, there is something analogous to the arbitrariness of signs in the psychological structures that underlie our mastery of a language. Imagine a pair of biological twins, one of them raised in England, and the other in Spain. Even though they may be physical duplicates, one of them will learn to use '**red**' to denote RED, while the other will learn to use '**rojo**' to denote the same property. Which specific referential abilities each of them acquired is, to some extent, accidental. One and the same physical substratum could be used to realize different referential abilities. It is therefore legitimate to ask: How can the same types of physical substrata embedded in different environments come to realize different referential abilities?²⁰

Moreover, from a developmental perspective, it makes sense to describe infants' first encounters with words like '**red**' as their bare recognition of sound patterns. Even though they might have expectations that '**red**' be meaningful, they might fail to experience '**red**' as having a determinate content, as occurs when a monolingual English speaker listens to Chinese or Russian. So there are psychological questions that look perfectly adequate even for those who may be reluctant to apply the vehicle-content distinction in the mental realm: How do children identify the content conveyed by physical vehicles like '**red**'? How should we characterize children's incorporation of new words in their

²⁰ To be sure, one might claim that the acquisition of a language produces changes in brain structures. So twins raised in different environments will not be physically identical. As far as I can see, this does not affect the main point of the argument: that there is something conventional in referential abilities.

cognitive life? How do these words interact with children's prior conceptual capacities?

The intuitive plausibility of these questions can be invoked in favor of semantic psychologism. McDowell and Wettstein might be right when they try to undermine some dualistic assumptions. After all, sharp dichotomies usually create unbridgeable gaps. Still, their remarks do not establish the truth of semantic anti-psychologism. It might be that the vehicle-content distinction is not adequate to theorize on the mental. It is however a good policy to introduce the vehicle-content distinction to mark the conventionality of natural languages. Besides, there is a similar conventionality in referential abilities. And none of the considerations introduced above suggests that one can explain reference only by adding a mysterious entity (i.e., intrinsic meaning or intentionality).²¹

6. The Argument from Confinement

The previous reply suggests that semantic anti-psychologism is based on a more substantial view of the mind and its place in the world. In the next sections, I explore some arguments that spell out this view. The first one, the 'argument from confinement,' traces back to early modern philosophy.

Here is a familiar line of reasoning. One starts by depicting minds as having ideas, intentionality, representations, concepts, a language, etc. Next, one assumes that these descriptions pick out some *constitutive* properties of minds. As a result, the question: 'How are ideas, intentionality, representations, concepts, a language, etc. 'hooked' to the world?' becomes problematic. After all, in order to solve the question, one has to assume that the relation between minds and those properties is not constitutive.

This problem has an epistemic counterpart. As theorists, we have minds. In order to answer questions about the relation between our constitutive properties and the world, we should be able to abstract from those properties. When we formulate the problem of reference, it is as if we had to get 'outside' ourselves or 'verify' how the reference relation between our constitutive properties and the world obtains. But, given that the link is constitutive, we cannot perform those feats.

²¹ The vehicle-content distinction is one of the most widely *used* in cognitive science. There is, however, little *discussion* on the ways of drawing it. As far as I know, the clearest and most recurrent version of the distinction is based on the model of public language. Marr's (1982) computational theory of vision introduces the distinction by means of the notion of a code. This has remained a common practice among cognitive scientists and philosophers (see Block 1995).

This line of argument can be seen as a reaction to our previous reply. As the prior example of the twins suggested, it is not a constitutive fact about the twins that they use 'red' or 'rojo' to refer to RED. Those are conventional facts in Lewis' (1969) sense: for each twin, there are alternative ways of picking out RED that would have been equally effective. Still, this does not undermine the intuition that drives the modest philosopher. After all, before each twin learned how to use 'red' or 'rojo,' she was already able to refer to the world. So providing an account of how they managed to incorporate 'red' or 'rojo' in their linguistic repertoire does not really explain their more basic referential abilities. Thus, there is a sense in which we are 'confined' to the realm of intentionality even when we tell a story about the twins' acquisition of the words 'red' or 'rojo.'

There are classical versions of this problem in Berkeley's (1710) arguments for the claim that *esse est percipi* and also in Kant's (1781/1787) arguments for the unknowability of things in themselves. And their force remains intact in some circles. So Rorty (1979) rejects the tendency of analytic philosophers to believe that they could occupy a neutral point of view on nature, i.e. a standpoint independent of any empirical theory. Similarly, Searle (1983) justifies the modesty of his semantic analysis of intentionality on the ground that it is impossible to get 'outside' the intentional circle:

In my view it is not possible to give a logical analysis of the Intentionality of the mental in terms of simpler notions, since Intentionality is, so to speak, a ground floor property of the mind, not a logically complex feature built up by combining simpler elements. There is no neutral standpoint from which we can survey the relations between Intentional states and the world and then describe them in non-Intentionalistic terms. *Any explanation of Intentionality, therefore, takes place within the circle of Intentional concepts.* (Searle 1983: 26; see also: 79)

It is difficult to resist this rhetoric of confinement. Human beings cannot occupy any neutral, external or transcendent point of view to survey the way natural language expressions, mental representations, or intentional states are related to the world. But we should be suspicious of this imagery. If we had skeptical proclivities, we should take very seriously the idea that we are confined to contemplate our ideas, representations or language. When we are theorizing on reference, however, we are already assuming that skepticism is not a live option. We are assuming that we *can* refer to the world. Crucially,

intentionality, representations, and languages are *parts* of the world. Having the ability to refer to intentional states, representations, and languages is the only thing we need to theorize on them!

The semantic psychologist is not forced to occupy any point of view external to intentionality, representations or language. What she rejects is to stay at the level of a *commonsense* understanding of mind and language. Granting that we must see the world through the prism of intentionality or language, we can still provide explanations *from* that perspective. This is what scientists do when they build models. Certainly, nobody has observed elementary particles with the naked eye. One can use, however, macroscopic objects as models of their structure. By parity of reasoning, even though we cannot place ourselves outside intentionality or language, we can use models available to language users to explain reference.²²

The prior reply might look simplistic to some readers. If it does, it is likely that there are more substantial commitments in the argument from confinement. In what follows, I show that *some* of the arguments one might use to rebut the previous line of reply do not undermine semantic psychologism.

7. The Challenge of Internal Realism

When one asks questions concerning how a representation (public or linguistic) refers to an object, one is tacitly assuming that objects are self-standing entities that are intelligible independently of the relation of reference. In other words, one is assuming that there are two different realms: the referring realm (constituted by minds, ideas or representations) and the referred realm (constituted by objects, events, properties, etc.). The problem is to explain how these two realms relate to each other in the asymmetric and normative way proper to reference. Some philosophers have challenged this assumption, though. Examples include transcendental idealism (Kant 1781/1787) and Putnam's (1988, 1990) more recent defense of internal realism.

One of Putnam's arguments goes as follows. The theory of reference presupposes an epistemic-free notion of an object. But there are reasons not to hold an epistemic-free notion of an object. So, in order to engage in a theory of reference, one has to presuppose something we have good reasons not to presuppose.

I will not examine Putnam's defense of internal realism. Instead, I will argue that, even if internal realism is correct, it does not undermine the program

²² I am indebted to Sellars (1956: 94-6; 1964: Chapter 1). See also Engel (1996: 239-ff.).

of providing an informative psychological account of referential abilities. To this end, let me start with the skeptical challenge he formulates on the way we *count* objects:

Suppose I take someone into a room with a chair, a table on which there are a lamp and a notebook and a ballpoint pen, and nothing else, and ask, "How many objects are there in this room?" My companion answers, let us suppose, "Five." "What are they?" I ask. "A chair, a table, a lamp, a notebook, and a ballpoint pen." How about you and me? Aren't we in the room?" My companion might chuckle. "I didn't think you meant I was to count people as objects. Alright, then, seven." "How about the pages of the notebook?" (Putnam 1988: 110-1)

Putnam's argument has the form of a skeptical challenge. For any portion of reality R (e.g. a room), there is a *salient* answer to the question: How many objects are there in R ? Still, one can always force one's opponent to grant that this salient answer is not compulsory, for it tacitly presupposes an arbitrary classificatory principle. Just by modifying the relevant classificatory principle, one can shift the estimation of the cardinality of objects. If the classificatory principle is 'non-living material object,' the response is 'five' but this answer is inadequate if one shifts to 'material object' as a classificatory principle. If classificatory principles are expressed by sortals $F_1, F_2, \dots, F_{n-1}, F_n$, for any portion of reality R , the number of objects it contains is relative to a sortal F_i . Since sortals are relative to our minds, epistemic equipment, conceptual schemes, languages, etc., there is no epistemic-free notion of an object we can rely on in order to formulate the problem of reference.

There are a number of replies available to the defender of the epistemic-free notion of an object (or 'metaphysical realist' for short). Nevertheless, for the sake of the argument, I will grant Putnam's claim that the notion of an object is not epistemic-free. I will also grant that *many* theorists of reference seem to presuppose some form of metaphysical realism. Nevertheless, it does not follow that the problem of reference will disappear just by pointing out that the notion of an object is epistemic-dependent.

We can defend this point by examining an extreme version of the claim that the notion of an object is epistemic: the idealist contention that objects are not merely *relative to* a conceptual scheme but also *mental constructions*. Even in this extremely constructivist framework there is an intuitive difference between what one might call *purely subjective* states like pains and *seemingly objective* states like visual or tactile perceptual experiences. If one holds

that objects are mental constructions, one has to offer an account of the intuitive difference between these two sorts of states. After all, pains *do not seem* to refer to items in the world, at least in the same way as visual and tactile experiences do. The situation becomes even more delicate in the presence of linguistic devices such as proper names, which *do seem* to have referential functions that go beyond our mental life. This clearly shows that the problem of reference is not only a problem for the metaphysical realist. It is a problem for everybody, including the idealist who grants that there is a distinction between seemingly subjective and seemingly objective mental states. Whereas the metaphysical realist may be puzzled by the gap between minds and a realm of epistemic-free objects, the idealist should be puzzled by the difference between seemingly subjective and seemingly objective mental states. To be sure, the idealist cannot formulate the problem of reference as the problem of relating mental states to mind-independent items in the world. Yet, she will face the complementary problem of *internal differentiation*: How do conceptual schemes or languages generate the difference between seemingly subjective and seemingly objective mental states? What cognitive abilities underwrite that contrast? The problem of internal differentiation is the anti-realist counterpart of the more familiar problem of reference.

8. The Autonomy of the Intentional

Some modest philosophers might concede that we can study intentionality by building models available from our intentional perspective. They might also grant that even the anti-realist faces a problem that is very similar to the problem of reference. Nevertheless, they might reply that psychological concepts are not well suited to provide scientific explanations of reference or intentionality. And, when such explanations are provided, they always leave out a crucial aspect of reference or intentionality. This line of thought is implicit in McDowell's writings on this topic:

An account given from outside is an account that denies itself the only descriptions under which we know that linguistic actions make rational sense, and we have been given no reason to suppose we can still see the activity of a speaker as hanging together rationally if we are required to describe it in other terms. (McDowell 1997: 113; see also Wettstein 2004)

The intuition behind this remark is that explanations at the sub-personal level fall short of providing a *full* understanding of reference, which is a personal-level notion. After all, it is *people* who refer to Aristotle by means of 'Aristotle,' not their parts or organs. So, cognitive accounts of reference leave out a central dimension of reference. This dimension includes the social and normative aspects of language use (see also Wettstein 2004: 108-9).

Consider an analogy. Someone asks: 'How does this machine work?' An engineer could try to find an answer by opening the machine, identifying the different parts, and seeing how they are related to each other. This would enable her to identify the function of the parts and the principles by which their interplay enables the machine to perform a complex task. The modest philosopher might insist, however, that this identification of smaller parts leaves something out. In the case of persons, the individual parts that compose them are not endowed with intentionality, for intentionality is the resultant of all the parts working in concert within a broader social and historical context. When you look for the smaller parts that make up the machine, you are not talking about intentionality anymore; the concepts of intentionality are designed to understand minds as situated in the world but not the mechanisms underlying them:

[F]olk-psychological concepts can express a kind of understanding of a person that seems to have little or no relation to predictive power. [...] If the understanding that common-sense psychology yields is *sui generis*, there is no reason to regard it as a primitive version of the understanding promised by a theory of inner mechanisms. The two sorts of understanding need not compete for room to occupy. (McDowell 1995: 413; see also McDowell 1994)²³

These remarks do not offer a clear argument against the semantic psychologist. Yet, they clearly delineate a possible view that, if true, would undermine the program of delivering a psychological account of reference. We can formulate the challenge as the conjunction of two claims:

(A) Autonomy:

(A1) Metaphysical: The personal level is independent from sub-personal levels.

²³ For a critical discussion of this sort of view, see Engel (1996), Bermúdez (2005), and Burge (2005).

(A2) Epistemological: Our understanding of ourselves as minds is confined to the personal level.

(I) Irreducibility:

The mind is irreducible to any sub-personal mechanism. In other words, the mind is not like an organ or a part of the person. Minds *are* persons.

If these claims were true, it would not be possible to understand the sub-personal level by employing or adapting aspects of the vocabulary of intentionality. Moreover, it would not be possible to shed light on personal-level notions like reference by engaging in sub-personal psychological theorizing. This approach enables us to see the debate between modest and full-blooded theorists in a new light. Psychological modesty is not merely a 'defeatist' attitude; it relies on a very specific view on the relation between personal and sub-personal levels.²⁴ Modesty hinges on the substantial claim that we cannot frame any full-blooded account of semantic competence in *sub-personal* terms.²⁵

I propose to conclude this paper by providing a consideration against the modest outlook. According to modesty, in order to elucidate language, intentionality, etc., we must remain at the personal level. If we move to a lower level, we cannot find anything that may be recognized as genuinely intentional. When we theorize on lower levels, we are just changing the topic. Brain processes are too distant from personal experience to tell us anything interesting about philosophical muddles.

I take this challenge as a serious one. We cannot predict whether we will be able to provide adequate *psychological* explanations of reference. Despite this difficulty, there is at least one reason that militates in favor of a psychological account of reference.²⁶ This reason exploits the notion of *intuitive understanding*. Imagine that neuroscience has progressed so much, that it enables us to correlate types of brain activation with uses of types of words like 'not,' 'but,' 'Aristotle,' and so on. Whenever a person uttered 'Aristotle,' we would observe an activation pattern *A*; whenever she uttered 'Plato,' we would observe an activation pattern *P*, etc. Similarly, neuroscience would manage to

²⁴ *Pace* Dummett (1987) and Smith (2006: 951).

²⁵ For a well-documented discussion of the autonomy claim, see Bermúdez (2005: 3.1-3.2).

²⁶ The following considerations are inspired by Cussins (1987, 1992).

provide accurate descriptions of the use of word-types in the context of sentences. Thus, the sentence-type: 'Aristotle is a Great Philosopher' would correspond to the activation pattern *AGP*. Even in this happy scenario, where a perfect correlation would obtain between types of neural events and types of linguistic behavior, we would not think that *this* correlation provided everything we needed to *understand* the referential abilities underlying our uses of 'Aristotle,' 'Plato,' etc. Even if we were ready to accept the existence of such type-type correlations, they would remain *too distant* from *our* personal-level understanding of our use of proper names.

These considerations can be accepted by both modest and full-blooded theorists because they display an explanatory insufficiency of type-type correlations. Now, their reasons for being unsatisfied would be different. The semantic anti-psychologist would reject psychophysical type-identity because she wants to stress the autonomy of the personal over the sub-personal level. The semantic psychologist, by contrast, would probably exploit the prior intuition on the epistemic insufficiency of such perfect type-type correlations to introduce intermediary explanatory levels to account for referential abilities. Even after having established detailed correlations between types of linguistic behavior and types of brain activity, an *explanatory gap* would remain.

I think this consideration provides a *slight* advantage to semantic psychologists over semantic anti-psychologists. If the autonomy picture is right, there is nothing we can do to bridge the gap between our commonsense understanding of reference and the sort of understanding that could be provided by neuroscience. If semantic psychologists are right, however, there is something we can do to bridge the gap: develop *intermediary explanations* of reference couched in a vocabulary enabling us to get *an intuitive* understanding thereof. This is precisely what cognitive science and AI do when they idealize over some details of brain activity, and elaborate a theoretical apparatus that borrows some concepts from commonsense psychology. This theoretical apparatus is sufficiently similar to commonsense to offer an intuitive understanding of a number of phenomena but also sufficiently different therefrom to advance our understanding of our mental life beyond the confines of our commonsense view.

Notice that the argument is based on epistemological considerations. The limits of some forms of reductionism and pluralism are not merely ontological. They are epistemological as well (Cussins 1992; Kim 2010). Even if psychological phenomena turned out to be physics in the long run, we would still mind the gap. We would want to understand why things that look so different turned out to be identical. This is, I think, the force of the prior argument:

It does not beg the question against the modest theorist who insists on the autonomy of our personal-level view of the world. It shows how uncomfortable that view is for *understanding* our own place in a world that is—I take it—fundamentally physical.

Given the epistemological orientation of this argument, however, some theorists might declare it unstable. Consider a well-known episode in the history of astronomy. The geocentric theory provided *the* 'intuitive' understanding of the motions of the stars. According to that doctrine, the Earth was a stationary body at the center of the universe, and the celestial bodies moved around it. It seems natural to think that this theory provided a more intuitive understanding than the subsequent heliocentric theory. After all, the geocentric theory could be easily mapped onto our everyday experience of the sky, while the heliocentric view contradicted that ordinary experience. Still, the intuitive view was rejected in favor of a more counterintuitive explanation. At the end of the day, people had to learn to think in a new way that contradicted their most entrenched intuitions. Why should semantic competence and intentionality be different? Why persist in accounting for reference by multiplying levels of explanation?

I have no definite response to these questions. For the time being, I would justify the relevance of sub-personal explanations by stressing a potential difference between the geocentric case and the hypothetical case of a perfect type-type correlation between types of brain activity and types of uses of words. When one engages in sub-personal theorizing, one is tacitly assuming that the gap between the higher level of language use and intentionality and the lower level of neurological mechanisms is so deep that intermediary explanatory levels are necessary. What we do is look at our personal-level concepts and try to exploit them as models to bridge the gap. But there is nothing analogous to this explanatory strategy in the geocentric example. *The hypothesis that the stars turn around the earth is not a step toward a better understanding of the heliocentric hypothesis but an obstacle thereto.* If we introduce sub-personal explanations of reference, it is because we mind the gap between explanations in neural terms and the semantic vocabulary we use to characterize language and intentionality. Bridging the gap requires the establishment of conceptual bridges, not a sharp personal/neural dichotomy.

9. Concluding Remarks

The contrast between full-blooded and modest theories of meaning articulates an opposition between an optimistic and a pessimistic attitude toward the explanation of reference. Since the relevant notion of explanation can be understood in different ways, there are many different ways of drawing the line. In this paper, I focused on a strand of the full-blooded/modest divide that has played a central role in Engel's work: whether one could provide a psychological account of referential abilities. I defended this program by displaying the insufficiency of truth-conditional semantics, and responding to some influential arguments against what I called 'semantic psychologism.'

The discussion of these arguments showed that, far from being a defeatist view, modesty hinges on substantial claims about the nature of mind and its relation to the world. My response to these arguments purported to show that, even if one grants some of these substantial claims, the problem of reference would not disappear. And our last argument led us to a more fundamental issue: the contrast between a view of the mind as an autonomous realm and a more interactive picture that seeks to spell out the relations that various levels bear to each other. Since the latter picture is not committed to reductionism, it bears some similarities to Engel's idea of world 2½: semantic psychologism may be seen as the study of how psychology and semantics are related to each other. This sort of inquiry requires a revision of the usual view of the philosophy of language as a branch of logic, and a rejection of anti-psychologism as our default view of normativity.

10. References

- Almog, J. (2005) Is A Unified Description of Language-and-Thought Possible? *The Journal of Philosophy*, 102 (10): 493-531
- Barwise, J. & J. Perry (1983) *Situations and Attitudes*, Cambridge (MA.), MIT Press / Bradford Books
- Berkeley, G. (1710) *A Treatise Concerning the Principles of Human Knowledge*, in his *Philosophical Writings*, edited by Desmond M. Clarke, Cambridge, Cambridge University Press, 2008
- Bermúdez, J. L. (2005) *Philosophy of Psychology: A Contemporary Introduction*, New York, Routledge

- Block, N. (1995) The Mind as the Software of the Brain, in E. E. Smith & D. N. Osherson, eds., *An Invitation to Cognitive Science*, 2nd edition, vol. 3: Thinking, Cambridge (MA.), MIT Press: 377-425
- Burge, T. (1977) Belief *De Re*, *The Journal of Philosophy*, 74 (6): 338-62. Reprinted with a postscript in his *Foundations of Mind*, Oxford, Clarendon Press, 2007: 44-81
- Burge, T. (1990) Frege on Sense and Linguistic Meaning, in David Bell & Neil Cooper, eds., *The Analytic Tradition: Meaning, Thought, and Knowledge*, Oxford, Blackwell: 30-60. Reprinted in his *Truth, Thought, Reason: Essays on Frege*, Oxford, Clarendon Press, 2005: 242-69
- Burge, T. (2005) Disjunctivism and Perceptual Psychology, *Philosophical Topics*, 33 (1): 1-78
- Clark, A. (2005) Intrinsic Content, Active Memory, and the Extended Mind, *Analysis*, 65 (285): 1-11
- Cussins, A. (1987) Varieties of Psychologism, *Synthese*, 70 (1): 123-54
- Cussins, A. (1992) The Limitations of Pluralism, in D. Charles & K. Lennon, eds., *Reduction, Explanation, and Realism*, Oxford, Clarendon Press: 179-223
- Davidson, D. (1984) *Inquiries into Truth and Interpretation*, Oxford, Clarendon Press
- Davies, M. (1987) Tacit Knowledge and Semantic Theory: Can a Five Per Cent Difference Matter? *Mind*, 96 (384): 441-62
- Devitt, M. (1981) *Designation*, New York, Columbia University Press
- Donnellan, K. S. (1974) Speaking of Nothing, *The Philosophical Review*, 83 (1): 3-31
- Dretske, F. (1981) *Knowledge and the Flow of Information*, Cambridge (MA.), MIT Press
- Dummett, M. (1973) *Frege: Philosophy of Language*, 2nd edition, Cambridge (MA.), Harvard University Press
- Dummett, M. (1975) What Is a Theory of Meaning? In S. Guttenplan, ed., *Mind and Language*, Oxford, Oxford University Press. Reprinted with an Appendix in his *The Seas of Language*, Oxford, Clarendon Press: 1-33
- Dummett, M. (1976) What Is a Theory of Meaning? (II) In G. Evans & J. McDowell, eds., *Truth and Meaning: Essays in Semantics*, Oxford / New York, Oxford University Press. Reprinted in his *The Seas of Language*, Oxford, Clarendon Press: 34-93

- Dummett, M. (1978) What Do I Know When I Know a Language? First published in his *The Seas of Language*, Oxford, Clarendon Press: 94-105
- Dummett, M. (1987) Reply to John McDowell, in B. Taylor, ed., *Michael Dummett: Contributions to Philosophy*, Dordrecht, Martinus Nijhoff: 253-68
- Dummett, M. (1991) *The Logical Basis of Metaphysics*, London, Duckworth
- Engel, P. (1989) *La norme du vrai : Philosophie de la logique*, Paris, Gallimard
- Engel, P. (1994) *Davidson et la philosophie du langage*, Paris, PUF
- Engel, P. (1996) *Philosophie et psychologie*, Paris, Gallimard
- Engel, P. (2001) The False Modesty of the Identity Theory of Truth, *International Journal of Philosophical Studies*, 9 (4): 441-58
- Engel, P. (2006) Logic, Reasoning and the Logical Constants, *Croatian Journal of Philosophy*, 6 (17): 219-35
- Engel, P. (2007) Review of Ernest Lepore, Kirk Ludwig, *Donald Davidson's Truth-Theoretic Semantics*, *Notre Dame Philosophical Reviews*, 8
- Engel, P. (Ms.) Une théorie de la signification peut-elle être autre que modeste ?
- Evans, G. (1981a) Understanding Demonstratives, in H. Parret & J. Bouveresse, eds., *Meaning and Understanding*, Berlin, New York, Walter de Gruyter: 280-303. Reprinted in his *Collected Papers*, Oxford, Clarendon Press: 291-321
- Evans, G. (1981b) Semantic Theory and Tacit Knowledge, in S. Holtzman & C. Leich, eds., *Wittgenstein: To Follow a Rule*, London, Routledge & Kegan Paul. Reprinted with a new section in his *Collected Papers*, Oxford, Clarendon Press: 322-42
- Evans, G. (1982) *The Varieties of Reference*, edited by John McDowell, Oxford, Clarendon Press
- Field, H. (1972) Tarski's Theory of Truth, *The Journal of Philosophy*, 69: 347-75
- Field, H. (1978) Mental Representation, *Erkenntnis*, 13: 9-61. Reprinted with a postscript in N. Block, ed., *Readings in Philosophy of Psychology*, vol. 2, Cambridge (MA.), Harvard University Press, 1981: 78-114
- Frege, G. (1891) Funktion und Begriff, *Vortrag, gehalten in der Sitzung vom 9. Januar 1891 der Jenaischen Gesellschaft für Medizin und Naturwissenschaft*, Jena, H. Pohle, II, 31 pp. Reprinted in G. Frege, *Funktion – Begriff – Bedeutung*, edited by M. Textor, Göttingen, Vandenhoeck & Ruprecht, 2002: 1-22

- Frege, G. (1893) *Grundgesetze der Arithmetik, begriffsschriftlich abgeleitet*, vol. 1, Jena, Pohle. Reprinted: Hildesheim, Olms, 1962
- Frege, G. (1918-1919) Der Gedanke – eine logische Untersuchung, *Beiträge zur Philosophie des deutschen Idealismus*, 1: 58-77. Reprinted in G. Frege, *Logische Untersuchungen*, edited by G. Patzig, Göttingen, Vandenhoeck & Ruprecht, 2003: 35-62
- Heck, R. (2002) Do Demonstratives Have Senses? *Philosophers' Imprint*, 2 (2): 1-33
- Higginbotham, J. (1998) Conceptual Competence, *Philosophical Issues*, vol. 9: Concepts: 149-162
- Kant, I. (1781/1787) *Kritik der reinen Vernunft*, Hamburg, Felix Meiner, 1998
- Kaplan, D. (1989) Afterthoughts, in J. Almog, J. Perry & H. Wettstein, eds., *Themes from Kaplan*, New York, Oxford University Press: 565-614
- Kim, J. (2010) *Essays in the Metaphysics of Mind*, Oxford, Oxford University Press
- Kripke, S. A. (1979) A Puzzle About Belief, in A. Margalit, ed., *Meaning and Use*, Dordrecht, Reidel: 239-83. Reprinted in P. Ludlow, ed., *Readings in the Philosophy of Language*, Cambridge (MA.) / London, MIT Press / Bradford Books: 875-920
- Kripke, S. A. (1980) *Naming and Necessity*, Cambridge (MA.), Harvard University Press
- Kripke, S. A. (2008) Frege's Theory of Sense and Reference: Some Exegetical Notes, *Theoria*, 74: 181-218
- Lewis, D. (1969) *Convention: A Philosophical Study*, Cambridge (MA.), Harvard University Press
- Lewis, D. (1980) Index, Context, and Content, in S. Kanger & S. Öhman, eds., *Philosophy and Grammar*, Dordrecht, Reidel: 79-100. Reprinted in his *Papers in Philosophical Logic*, Cambridge, Cambridge University Press, 1998: 21-44
- McDowell, J. (1977) On the Sense and Reference of a Proper Name, *Mind*, 86: 159-85. Reprinted in his *Meaning, Knowledge, and Reality*, Cambridge (MA.) / London, Harvard University Press: 171-98
- McDowell, J. (1987) In Defense of Modesty, in B. Taylor, ed., *Michael Dummett: Contributions to Philosophy*, Dordrecht, Martinus Nijhoff: 59-80. Reprinted in his *Meaning, Knowledge, and Reality*, Cambridge (MA.) / London, Harvard University Press: 87-107

- McDowell, J. (1992) Putnam on Mind and Meaning, *Philosophical Topics*, 20: 35-48. Reprinted in his *Meaning, Knowledge, and Reality*, Cambridge (MA.) / London, Harvard University Press: 275-91
- McDowell, J. (1994) The Content of Perceptual Experience, *The Philosophical Quarterly*, 44: 190-205. Reprinted in his *Mind, Value, and Reality*, Cambridge (MA.) / London, Harvard University Press: 341-58
- McDowell, J. (1995) Knowledge and the Internal, *Philosophy and Phenomenological Research*, 55: 877-93. Reprinted in his *Meaning, Knowledge, and Reality*, Cambridge (MA.) / London, Harvard University Press: 395-413
- McDowell, J. (1996) *Mind and World*, 2nd edition, with a New Introduction, Cambridge (MA.), Harvard University Press
- McDowell, J. (1997) Another Plea for Modesty, in R. G. Heck, ed., *Language, Thought, and Logic: Essays in Honour of Michael Dummett*, Oxford / New York, Oxford University Press: 105-29. Reprinted in his *Meaning, Knowledge, and Reality*, Cambridge (MA.) / London, Harvard University Press: 108-31
- McKinsey, M. (2009) Thought by Description, *Philosophy and Phenomenological Research*, 78 (1): 83-102
- Marr, D. (1982) *Vision*, San Francisco (CA.), W. H. Freeman
- Millikan, R. G. (1984) *Language, Thought, and Other Biological Categories*, Cambridge (MA.) / London, MIT Press / Bradford Books
- Millikan, R. G. (2005) *Language: A Biological Model*, Oxford, Oxford University Press
- Predelli, S. (2005) *Contexts: Meaning, Truth, and the Use of Language*, Oxford, Clarendon Press
- Pylyshyn, Z. W. (2007) *Things and Places: How the Mind Connects with the World*, Cambridge (MA.) / London, MIT Press / Bradford Books
- Putnam, H. (1988) *Representation and Reality*, Cambridge (MA.) / London, MIT Press / Bradford Books
- Putnam, H. (1990) *Realism With a Human Face*, Cambridge (MA.), Harvard University Press
- Rorty, R. (1979) *Philosophy and the Mirror of Nature*, Princeton / Oxford, Princeton University Press
- Sainsbury, R. M. (2005) *Reference Without Referents*, Oxford / New York, Oxford University Press

- Salmon, N. (1986) *Frege's Puzzle*, Cambridge (MA.) / London, MIT Press / Bradford Books
- Searle, J. R. (1983) *Intentionality*, Cambridge, Cambridge University Press
- Sedivy, S. (2004) Minds: Contents Without Vehicles, *Philosophical Psychology*, 17 (2): 149-81
- Sellars, W. (1956) Empiricism and the Philosophy of Mind, *Minnesota Studies in Philosophy of Science*, vol. 1, edited by H. Feigl & M. Scriven, Minneapolis, University of Minnesota Press. Reprinted as *Empiricism and the Philosophy of Mind*, with an Introduction by R. Rorty and a Study Guide by R. Brandom, Cambridge (MA) / London, Harvard University Press
- Smith, B. C. (2006) What I Know When I Know A Language, in E. LePore & B. C. Smith, eds., *The Oxford Handbook of Philosophy of Language*, Oxford, Oxford University Press: 941-82
- Wettstein, H. (2004) *The Magic Prism: An Essay in the Philosophy of Language*, Oxford / New York, Oxford University Press