

Semantic Externalism and Psychological Externalism

Åsa Wikforss
Department of Philosophy
Stockholm University
asa.wikforss@philosophy.su.se

Abstract

Externalism is widely endorsed within contemporary philosophy of mind and language. Despite this, it is far from clear how the externalist thesis should be construed and, indeed, why we should accept it. In this entry I distinguish and examine three central types of externalism: what I call *foundational externalism*, *externalist semantics* and *psychological externalism*. I suggest that the most plausible version of externalism is not in fact a very radical thesis and does not have any terribly interesting implications for philosophy of mind, whereas the more radical and interesting versions of externalism are quite difficult to support.

1. Introduction

Externalism has created a great deal of excitement within philosophy of mind and language. It is widely perceived to have overthrown traditional theories not only of meaning and mental content, but also of the nature of psychological states. Indeed, externalism has been so successful that the primary focus of today's debate is not so much on whether externalism is right or wrong, but rather on what its implications are. For instance, there is an on-going debate concerning the compatibility of externalism and self-knowledge, and more generally, a debate concerning the epistemological implications of externalism. (For recent contributions see Brown 2004 and Nuccetelli 2003). Here, however, I shall suggest that we focus on the basics: What is externalism? And why should it be accepted?

In what follows I distinguish and examine three central types of externalism which are often conflated: what I call *foundational externalism*, *externalist semantics* and

psychological externalism. The two first theses are semantic theses, whereas the third is a thesis concerning the individuation of types of psychological states. My main concern is with kind terms (such as 'water', 'arthritis' and 'sofa') although I shall touch upon the issue of singular terms as well. I will also concentrate on foundational externalism, since this is the type of externalism most commonly endorsed, and since spelling it out in some detail provides the background against which externalist semantics and psychological externalism can be seen. I shall suggest that the most plausible version of externalism is not in fact a very radical thesis and does not have any terribly interesting implications for philosophy of mind, whereas the more radical and interesting versions of externalism are quite difficult to support.

2. Two types of semantic externalism

The label 'externalism' is most commonly used to refer to some kind of thesis concerning meaning. Various formulations of the thesis can be found in the literature: There is Hilary Putnam's catchy 'meanings ain't in the head', as well as less memorable ones, such as the claim that the meaning of some types of terms *depends* on features of the external environment, or that the meaning of these terms is *determined* by external features. What hides behind these phrases?

As stressed by Robert Stalnaker, we have to distinguish between two types of semantic theories: foundational semantics and descriptive semantics (1997: 535-538). A foundational semantic theory tells us something about the *determination* of meaning: For instance, in virtue of what does the word 'water' have the meaning that it does? What are the facts that determine this meaning and how is the function from these facts to the meaning to be understood? Different theories appeal to different sets of facts: One theory might refer to facts about the speaker's psychological states, another to facts about S's linguistic usage or facts about causal connections. These facts constitute the *determination basis* such that the

term *T* having meaning *M* depends on them.¹ Descriptive semantics, by contrast, tells us something about the *semantic values* of the terms of a given language. The notion of 'semantic value', here, is used liberally, as to include whatever it is that provides an interpretation of the expression and contributes to the proposition expressed. For instance, it might be held that 'water' is directly referential and lacks all descriptive content, or that 'water' expresses a composite, descriptive concept (*the wet, thirst-quenching, transparent liquid that flows in rivers and taps...*) or, perhaps, that it expresses two distinct intensions along the lines of two-dimensionalism.

Externalism in its most familiar form is a thesis within foundational semantics. It states that the external environment plays a role in the determination of meaning (and, according to some, mental content) and is most famously defended by Hilary Putnam (1975) and Tyler Burge (1979, 1982, 1986). Accordingly, I shall call this type of externalism *foundational externalism*. Sometimes, however, 'externalism' is used to refer to a particular thesis within descriptive semantics: i.e. the thesis that certain types of terms have an object-dependent semantic value. The thesis of object-dependence is defended by, among others, Gareth Evans and John McDowell (Evans 1982; McDowell 1977, 1986). I shall label it *externalist semantics* in order to distinguish it from foundational externalism. Let us begin by examining foundational externalism in some detail.

3. Foundational externalism

3.1 The role of the physical environment

According to Putnam traditional theory of meaning rests on two unchallenged assumptions:

- (i) Knowing the meaning of a term *T* is just a matter of being in a certain psychological state.

(ii) The meaning (intension) of *T* determines its extension in the sense that sameness of intension entails sameness of extension.

It follows from (i), Putnam suggests, that the psychological state of the speaker *determines* the meaning of *T* and hence, in accordance with (ii), that sameness in psychological state entails sameness in extension (1975: 219, 222; see also Putnam 1970: 139, 140). For instance, according to the traditional descriptivist account of general terms, the meaning of a term such as 'lemon' is determined by the set of descriptions that speakers associate with it (yellow color, tart taste, a certain kind of peel, etc.). These descriptions provide individually necessary and jointly sufficient conditions for something to be a lemon. If two speakers associate the same set of descriptions with 'lemon', therefore, the term as used by them must have the same meaning and extension.

Putnam's contention is that what characterizes natural kind terms is precisely that the descriptions associated by ordinary speakers fail to determine meaning since they fail to provide necessary and sufficient conditions for the application of the term. In order to illustrate that the associated descriptions are not sufficient, Putnam hypothesizes Oscar on Earth and his twin on Twin-Earth ('Toscar') that are said to be in the same psychological state. Oscar and Toscar have associated all the same descriptions or stereotypes with their respective term 'water' (colorless, transparent, tasteless, thirst-quenching, etc). And yet, Putnam argues, their terms have a different extension as a result of the fact that they are situated in environments that differ in one essential respect: on Earth the liquid called 'water' has the chemical composition H₂O, whereas on Twin Earth it has an entirely different chemical composition, abbreviated XYZ. Putnam suggests that this leaves us one of two options: We can either reject (i), and give up the idea that meaning is determined by psychological state, or reject (ii) and grant that 'water' has the same meaning on the two planets despite the difference in extension.

Putnam recommends against giving up (ii). This, he argues, would be the correct move in the case of *absolutely* indexical terms, but natural kind terms are not absolutely indexical even though indexical elements play a role when these terms are introduced into the language. For instance, Putnam suggests, we may point to a glass of water and say 'this liquid is water' (1975: 225, 231). However, he argues, unlike an indexical such as 'I', a change in the extension of 'water' would entail a change in its meaning (1975: 245-255). Consequently, the moral of the thought experiment is that we should reject (i): What the thought experiment shows is precisely that the psychological states of speakers do not determine meaning. 'Water' as used by Oscar has a different meaning than 'water' as used by Toscar, despite the fact that the two individuals are psychologically identical.

Putnam, therefore, is plausibly construed as defending a version of foundational externalism. In its most general form, foundational externalism can be characterized as the thesis that the set of all facts that determine the meaning of *T* include external facts; that is, features of the external environment are included in the determination basis. Applied to natural kind terms foundational externalism may therefore be formulated as follows:

(FE) For all natural kind terms *T*, and all meanings *M*, the totality of facts that determine that *T* expresses *M* include external facts.²

Foundational internalism, by contrast, is the thesis that the determination basis includes only internal facts. The distinction between externalism and internalism may then be illustrated as follows (where the arrow indicates the determination relation):

Foundational externalism

'Water' means water



External (and internal) facts

Foundational internalism

'Water' means water



Internal facts

Externalism is often characterized as the thesis that meaning (and content) fails to supervene on internal facts. Thesis (FE) coheres with this characterization, since the suggestion that external facts are included in the determination basis implies that meaning does not supervene on internal facts. However, since supervenience may fail for a variety of reasons, it is not very fruitful to characterize externalism in terms of *failure* of supervenience.³ Rather, foundational externalism should be understood as a *positive* supervenience thesis, i.e. as the thesis that meaning supervenes on a set of external and internal facts.

A central question concerns how the distinction between external and internal facts is to be drawn. As has been pointed out in the literature, the distinction should not be drawn in terms of what is 'inside' and 'outside' the skin, since there may be natural kind terms that refer to physical conditions of the body ('meningitis') and it seems plausible that an externalist Twin Earth thought experiment could be run on these too (Farkas 2003: 191). I believe that the relevant distinction between internal and external facts should be seen against the background of the rejection of descriptivism. At the heart of externalism lies the idea that although the descriptions ordinary speakers associate with *T* may serve a role in *fixing* the reference and extension of a term *T* (picking out objects in the actual world), they do not provide necessary and sufficient conditions and hence do not *determine* the meaning of *T*: this, instead, is done by the underlying physical nature of the paradigms or objects fixed upon. The 'object' may be a puddle of water, for instance, or a gold ring or a tiger, and the

'underlying' nature is standardly assumed to be the microphysical structure of the object in question. In the case of water, again, it is the underlying *chemical composition* that is suggested to play this meaning determining role, with the implication that if 'water' picks out H₂O in this world, it picks out H₂O in all possible worlds. This characterization of externalism allows for the possibility of giving an externalist account also of terms referring to 'objects' inside the skin of the agent: What determines the meaning of these terms is not the descriptions associated with them, but the underlying nature of the 'object' fixed upon.

Once foundational externalism is extended to mental content, of course, the picture becomes more complex. Putnam, again, characterizes his position by saying that Oscar and Toscar are psychologically the same (have all the same intentional states, etc.) despite the fact that 'water' as uttered by Oscar has a different meaning than 'water' as uttered by Toscar. Putnam therefore attempts to combine foundational externalism with respect to *meaning*, with foundational internalism with respect to *mental content*.⁴ (For other such attempts see Fodor 1980; Loar 1985; McGinn 1989.) According to the *content* externalist, however, the externalist determination of meaning carries over to content in such a way that if the meaning of 'water' is determined externally, so is the corresponding concept expressed. This view is driven by the conviction that mental content is truth-conditional and hence the external determination of truth-conditions will carry over to the level of mental content. Indeed, according to Tyler Burge the meaning of 'water' is not distinct from the concept expressed by the term (Burge 1993: 312). Consequently, on Burge's view, Oscar and Toscar have different 'water' concepts and hence are not psychologically the same - while Oscar thinks that water is wet, Toscar thinks that twater is wet, etc. (Burge 1982: 101).⁵

Burge concludes from this that Putnam's attack on the traditional picture is misguided. The content externalist, he argues, can retain both of the traditional assumptions (i) and (ii) listed by Putnam. That is, the content externalist can hold both that knowing the

meaning of a term is a matter of being in a certain psychological state and that meaning determines extension (Burge 1993: 319- 320). Oscar knows the meaning of 'water' in the sense that he uses 'water' to express beliefs, and these beliefs contain a concept that determines the referent or extension of the term. Hence, Burge argues, psychological states do determine meaning and meaning does determine extension.

However, this is to conflate two distinct notions of determination.⁶ On the content externalist view the psychological state 'determines' meaning in the sense that sameness in psychological state entails sameness in meaning and extension. This just means that there is a function from psychological facts to meanings and is in that sense a purely formal notion of determination. It does not tell us in virtue of what 'water' has the meaning it does in the first place, which is the relevant, foundational notion of determination here. In this latter sense, clearly, psychological states do *not* determine meaning on Burge's view. After all, according to Burge the concept expressed, the concept that figures in Oscar's beliefs, is determined jointly with the meaning of 'water', and hence the answer to the foundational question (what determines the meaning of 'water'?) cannot refer to these very psychological states. Instead, on the content externalist view, the types of facts that are included in the determination basis for concept *C* must be specified independently of the psychological states containing the concept.

Moreover, like Putnam, Burge is committed to the idea that the descriptions associated with a term (concept) by ordinary speakers fail to provide necessary and sufficient conditions for the application of the term (concept). While, therefore, he wishes to endorse the thesis that knowledge of meaning is a psychological state, the resulting position diverges from the traditional one. Traditionally, knowing the meaning of a term entailed knowing the necessary and sufficient conditions for the application of the term. On Burge's view, by contrast, knowledge of meaning does not entail knowledge of necessary and sufficient

conditions since such conditions are determined by factors not known to the individual. The meanings of many terms and the identities of many concepts, Burge writes for instance, "are what they are even though what the individual knows about the meaning or concept may be insufficient to determine it uniquely. Their identities are fixed by environmental factors that are not entirely captured in the explicatory or even discriminatory abilities of the individual..." (1993: 318).

3.2 The role of the social environment

A further complication concerns the suggestion that not only the *physical* environment plays a role in determining meaning but also the *social* environment. This idea is most famously defended by Burge in his 'arthritis' thought experiment (1979).⁷

Burge postulates an individual who thinks that 'arthritis' applies to all types of rheumatoid diseases and utters 'I have arthritis in my thigh'. However, in the individual's community the term applies only to rheumatoid diseases of the joints by 'standard dictionary definition'. Since, Burge says, the individual is committed to her community practice, in the sense that she stands corrected by the experts, the term as used by her must nevertheless be said to have its standard extension and express the standard concept. Hence, the individual expresses a (conceptually) false belief by her utterance. In a counterfactual community where the term 'arthritis' is given a wider definition and applies to rheumatoid diseases of the joints as well as the ligaments, however, the individual would have another concept (the concept of tharthritis) and hence express a true belief when uttering 'I have arthritis in my thigh'. The upshot is that features of the individual's social environment play a part in the determination of meaning and mental content (1979: 79).

Although often treated on a par with physical externalism, it should be clear that social externalism, or as Burge likes to call it 'anti-individualism', is a rather different thesis.⁸

Like physical externalism, social externalism is based on the observation that the descriptive information that ordinary speakers have associated with their terms fail to provide necessary and sufficient conditions for the application of the term. Unlike physical externalism, however, the idea is not that the underlying features of the *objects* picked out serve to determine the meaning of *T*, but that the practice of the *experts* determine the meaning. Hence, social externalism does *not* stand in opposition to the traditional idea that meaning is determined by facts about individual speakers - it only stands in opposition to the idea that meaning is determined by facts about ordinary, non-expert speakers. This has the consequence that social externalism will not yield the same results as physical externalism. For instance, social externalism does not support Putnam's claim that in 1750 (before the development of modern chemistry) 'water' on Earth expressed a different meaning than 'water' on Twin Earth.⁹ In 1750 the experts on Earth and the experts on Twin Earth would have associated all the same descriptions with 'water', and hence the term 'water' in English would have had the same meaning as the term 'water' in Twin-English.

Indeed, what is distinctive about social externalism is not so much that it represents a break with traditional accounts of the determination of meaning (and content), as that it represents a break with traditional accounts of linguistic (and conceptual) *competence*. Social externalism relies on the assumption that meanings are 'transmitted' within the linguistic community, even when the individual has a rather incomplete grasp of the meaning (concept) in question (Burge 1979: 83). On Burge's view, the individual's commitment to the community practice entails that when she uses a word in a non-standard way the word should nevertheless be interpreted as having its standard meaning. Consequently, the individual should be described as having made a linguistic (conceptual) error, rather than as using her words with a non-standard meaning (concept). Burge concedes that there are cases of radical misunderstandings, where charity requires that the individual is reinterpreted rather than

ascribed a linguistic error (he gives the example of the person who claims to have 'orangutans' for breakfast, thinking 'orangutan' applies to a fruit drink (1979: 90)). However, he suggests that there are many cases where the linguistic (and conceptual) error is not sufficiently radical for reinterpretation to be required (as when the patient utters 'I have arthritis in my thigh'). This class of cases provides the foundation for Burge's social externalism: "The thought experiment depends only on there being some cases in which a person's incomplete understanding does not force reinterpretation of his expressions in describing his mental contents. Such cases appear to be legion." (1979:92)

Much of the plausibility of social externalism hinges on whether this claim can be defended. One difficulty concerns Burge's rather selective use of the principle of charity. If rationality considerations require reinterpretation in the case of *some* linguistic deviances, some departures from the community usage, why not in the case of all such deviances? In all such cases, after all, the failure to reinterpret means introducing a problematic gap between the individual's conceptions and the concept in question which leads to well-known difficulties accounting for her cognitive perspective (see Bilgrami 1992, Loar 1985). Moreover, in all such cases the individual will typically stand corrected and change her ways. Of course, it may well be that it is implausible to reinterpret the individual who utters 'I have arthritis in my thigh', in particular since she otherwise uses the term so competently. But this could just be taken to show that the error in question is not plausibly construed as a conceptual error. That is, the reason we do not have the intuition that any rationality constraints are violated in this case is precisely that the individual cannot be described as having an incomplete grasp of the concept in question. However, as Burge himself stresses, the assumption that the individual has an incomplete grasp of the concept is essential if the thought experiment is to go through.¹⁰

Social externalism is therefore best construed as a thesis concerning the relation between the community practice and individual competence. Unlike physical externalism, it is compatible with the traditional idea that the meaning of *T* (the concept expressed) is determined by facts about the collective use of *T*. Hence, if 'external facts' are construed as suggested above, as facts about the objects picked out by the relevant term, social externalism does not qualify as a form of foundational externalism. In this sense, Burge's preferred label for his position, 'anti-individualism', is more apt than that of 'social externalism'.

3.3 The status of (FE)

Let us therefore focus on foundational externalism of the sort that appeals to the role of the speaker's *physical* environment in the determination of meaning, along the lines of (FE).

Three questions need to be addressed: (i) What type of dependency on the external world is entailed by (FE)? (ii) What is the scope of foundational externalism? (iii) Why should we accept (FE)?

First, it should be clear that (FE) is a rather weak thesis. That the *totality* of facts that determine meaning (that 'water' means water, say) include external facts, does not mean that every set of facts that determine this meaning must include external facts - several different sets of facts within the determination basis could serve to determine the same meaning (content). In general, as stressed by Peter Pagin, the function from the meaning determining facts to meaning (content) need not be *one-one*, but can be *many-one* such that several different sets of facts within the basis determine the same meaning (1997: 23-24; 2006). Thus, thesis (FE) is compatible with the following foundational account:

'Water' means water



External & internal facts

Internal facts

Although **(FE)** suffices to rule out foundational internalism, it does therefore not support the claim, commonly made, that the meaning of natural kind terms is *relational* in the sense that meaning water by 'water' requires having interacted with instances of water (see for instance Davies 2003: 25). To support this claim a stronger form of foundational externalism is required, according to which *every* set of facts that determine *M* includes a given type of external fact *E* (such as the presence of water):

(FED) For all natural kind terms *T*, and all meanings *M*, every set of facts that determine that *T* means *M* includes external fact *E*.

However, contrary to received opinion, the Twin Earth thought experiment does not in itself support **(FED)**.¹¹ What Putnam's thought experiment illustrates (if we accept it) is that less is required in the way of descriptive information than traditionally assumed in order to use a kind term with its standard meaning - as long as one is placed in the right kind of environment. This tells us something about the *sufficient* conditions for meaning water by 'water' (for possessing the concept of water), but it does not tell us anything about the necessary conditions. It is fully compatible with what Putnam says, that one may mean water by 'water' in virtue of some other facts. For instance, the thought experiment does not rule out the possibility that one does so in virtue of being an expert and having a rich and complex pattern of holding sentences true ('Water is H₂O', 'Under normal pressure water boils at 100C', etc.), even though one has never in fact interacted with actual water (Burge 1982 grants this).

Second, the scope of foundational externalism: As stated above, **(FE)** is a thesis concerning natural kind terms. This raises two questions: What is a natural kind term? And can the thesis be extended to other kind terms, such as terms for artifacts? As to the first question, there has been much discussion recently. Initially it was taken for granted that whether a term is a natural kind term depends merely on the speaker's semantic intentions. This motivated the idea that these terms were to be given a separate semantic treatment. However, the question arose what to say of those cases where a purported natural kind term fails to pick out an underlying kind with a unified micro-structural nature. A famous example is 'jade', which was discovered to pick out two distinct minerals, jadeite and nephrite (see for instance LaPorte 1996). It seems intuitively implausible to count such cases as cases of reference failure. The standard reaction to this difficulty has been to suggest that whether or not a term is a natural kind term depends on whether it in fact succeeds in picking out a natural kind (see for instance Brown 2004, Gallois 1996, McLaughlin & Tye 1998, Nuccetelli 2003). It follows that the assumption that a term falls under **(FE)** is falsifiable by scientific discoveries. This is not an entirely unproblematic move however. In particular, it raises the question why we should not say that the presence of XYZ on Twin Earth (which Putnam assumes to be in the same world as planet Earth) constitutes a falsification of the assumption that 'water' is a natural kind term. Of course, unlike nephrite, twin water cannot be found on planet Earth. But why should that matter? If the assumption that 'water' names a natural kind is fallible, then it seems arbitrary to limit possible falsifications to planet Earth (see Wikforss 2005 for a discussion of this difficulty).

Related to this is the question of whether foundational externalism can be extended to apply to non-natural kind terms. Although it has not received much discussion, Burge has suggested that a version of **(FE)** can be defended in the case of all 'empirically applicable terms': a large class of terms applying to 'everyday, empirically discernible

objects', including 'mud', 'stone', 'tree', 'bread', 'knife', 'sofa' (1989, 181; 1986a, 699; 2007, 24). The meaning of these terms, according to Burge, is fixed by the use of examples ('That's a sofa' etc) and the individual's explicational abilities therefore do not exhaust their meaning: Even in the case of an expert, he argues, there will typically be things that are true of the object picked out that the individual is unaware of and that serves to determine the meaning of the term and the content of her thoughts. To show this Burge uses an argument that might be called 'the argument from cognitive distance'. The suggestion is that anti-individualism is a precondition of the very notion of objectivity, the idea that there is a mind-independent world that we can be right and wrong about in our judgments. For instance, Burge argues, it is essential that our perceptual judgments may be mistaken and this requires that perceptual content is determined non-individualistically, by objective features in the individual's environment rather than by her dispositions to discriminate among such features (1986b; 1986c; 2007). The distinction between superficial appearances and underlying natures, central to the thought experiments concerning natural kind terms, Burge writes, "is just a special case of our lack of omniscience with respect to any objective empirical subject matter, even the superficial features" (2007:23).

How plausible is this extension of (FE)? It is clear that the mere fact that an individual has an incomplete *knowledge* of the properties of instances of a kind (failing to know everything there is to be known about sofas, for instance) does not suffice to support the externalist conclusions. To show that the meaning of the term is determined externally two additional assumptions have to be made: First, that what I fail to know is so central to the kind that nothing could be an instance of the kind without having this property. Second, that despite my ignorance (and error) I am to be ascribed thoughts containing this concept. The trouble is finding good examples where both of these conditions are plausibly fulfilled. There are many things I do not know about sofas (how they are made, for instance), but, arguably,

that does not show that I have an incomplete grasp of the meaning of 'sofa', nor does it suffice to support a Twin Earth style argument. And although one may imagine an individual who is radically confused about the objects called 'sofa' (thinking they are not pieces of furniture, for instance), the question is whether such an individual really could be said to use 'sofa' with its standard meaning.¹² In this respect natural kind terms are at least *prima facie* different: In the case of natural kinds the distinction between superficial and underlying features coincides with the distinction between accidental and essential features (according to the standard view) and it does not appear problematic to ascribe a natural kind concept to someone who fails to know the underlying essence of the kind.

Moreover, the argument from cognitive distance is not compelling. Granted, objectivity requires that a distinction can be drawn between what is right and what seems right to the agent, i.e. that there is a *possibility* of error (which, of course, is distinct from demanding actual fallibility). However, this distinction does not, in turn, presuppose that there is a distinction between what an individual takes a word to mean (her conceptions) and what the word in fact means (the expressed concept). Incomplete knowledge of the empirical world simply does not presuppose an incomplete grasp of one's own concepts, as Burge would have it (see especially Burge 2007: 23). There are theories of content that have difficulties accounting for error, for instance versions of the causal theory and certain types of unrestricted holistic theories. But there is no principled reason why internalist theories should be worse off in this respect than externalist ones. (See Boghossian 1989, 534-540: and Wikforss 2001: 207-212, for a discussion of the problem of error.)

This takes us to the third question, which is why thesis **(FE)** should be endorsed in the first place - even when limited to the case of natural kind terms. As noted above, Twin Earth style thought experiments are based on the observation that the descriptive information typically possessed by ordinary speakers is too sketchy and incomplete to provide sufficient

conditions for something to fall under the term.¹³ However, there are several responses available to this observation other than foundational externalism.

A common internalist strategy is to appeal to the idea that the relevant psychological states contain an indexical element (Donnellan 1993; Farkas 2006; Searle 1983: 204-205). Putnam, as noted above, grants that indexical elements play a role in fixing the reference of natural kind terms but denies that natural kind terms function like indexicals. The indexical proposal disputes this. Drawing on Kaplan's distinction between character and content, it is suggested that natural kind terms have a constant meaning captured by a semantic rule ('character') that, together with the external environment, determines extension and truth-conditions ('content') (Kaplan 1989). For instance, associated with 'water' is a rule that says that the term applies to the wet, thirst-quenching, transparent, etc. liquid *around here*. The indexical definition is 'in the head', and hence what is in the head determines meaning which, together with context, determines extension (Searle 1983: 205). The meaning of 'water' as used by Oscar is determined by his psychological state and is identical to the meaning of 'water' as used by Toscar, despite the fact that the two terms differ in extension - much like the meaning of 'here' in Oscar's mouth is the same as the meaning of 'here' in Toscar's mouth, despite the difference in reference.¹⁴

Whether this proposal succeeds depends on whether natural kind terms really can be construed along the lines of indexicals. Is there, in this case, something corresponding to Kaplan's notion of character? (See Stalnaker 1989 for a critical discussion.) An obvious difficulty is that unlike indexicals, natural kind terms do not change extension from one context of speech to another (Burge 1982: 103-105; 2007, 10-11). Or, rather, if it is granted that 'water' changes extension with speech context it would follow that the Earthlings who visit Twin Earth would be saying something true when uttering 'There is water in the well' - contrary to the assumptions of the thought experiment. The question is whether the internalist

can block this implication, without undermining the suggested parallel with indexicals. One proposal is that in the case of natural kind terms what matters is not the context of *use*, but the context of *acquisition* (Farkas 2006). However, giving up the idea that it is the context of use that determines the extension of 'water' (together with its meaning) is precisely to give up the idea that these terms can be compared to indexicals. We are thus left without an account of how to understand the function from context to extension.

Even if the indexical account fails, however, other responses are available. First, it should be noted that the insufficiency of the associated descriptive information is exaggerated. Granted, the descriptions typical speakers have associated with words such as 'gold', 'water' and 'tiger' is such that it would also fit twin-gold, twin-water and twin-tigers (in the case of 'water', of course, this is dubious since most speakers today do have the belief that water is H₂O). However, luckily, there are no true twin-substances and twin-animals in the actual world.¹⁵ Hence, the descriptive information suffices to secure sameness in extension in the actual world, despite big differences among speakers in the richness of the associated descriptions. The expert who has a fully developed theory about tigers and the non-expert who only associates stereotypical descriptions with the word ('tigers are striped', 'tigers are carnivorous', etc) still manage to pick out the same group of animals (more or less). The world, as it were, cooperates to secure a basic level of agreement, essential to communication, also while the perspectives of individual speakers on this world vary. Of course, this does not extend across all possible worlds, and when the perspectives radically diverge one should expect disagreement with respect to some counterfactual scenarios, in particular, far removed scenarios where the laws of nature are different. But this, after all, is precisely what is illustrated by the debate on externalism: Despite the central role played by intuitions in the externalist arguments, there is little agreement on what we are to say about the various counterfactual scenarios presented, such as Twin Earth. Some people have the intuition that

twin-water is a type of water, others do not, and there are scenarios discussed in the literature that people have very different reactions to, such as Dry Earth, a completely dry planet where the inhabitants are merely under the illusion that there is a wet, thirst-quenching, etc. liquid in their environment (see Boghossian 1997 and McLaughlin & Tye 1998). One possibility, then, is that this disagreement reflects conceptual differences that do not surface in ordinary linguistic interactions.

Moreover, there is no reason why the descriptive properties would have to be limited to superficial properties.¹⁶ This limitation is an essential assumption underlying Putnam's claim that associated descriptions are insufficient. However, while it is clear that ordinary speakers do not typically have beliefs about the microstructural features of the substances and liquids in their environment, they do have the belief that some kinds are natural kinds in the sense that they have a unified, underlying nature. Indeed, the assumption that ordinary speakers have this kind of belief plays a central role in the externalist argument for why our term 'water' does not apply to twin-water. The internalist, similarly, can appeal to this belief if she wishes to secure the intuition that XYZ is not a type of water: 'Water' does not apply to XYZ since it is a natural kind term and these pick out unified underlying kinds, not a motley. This allows the internalist to accommodate at least some of the externalist intuitions without having to fall back on the idea that natural kind terms function like indexicals. Again, once it is granted that the assumption that 'water' picks out a natural kind can fail, it is not obvious that twin-water does not qualify as a kind of water. However, this difficulty is one that the externalist faces too, given the central role played by the assumption that 'water' names a natural kind in the externalist argument.

Finally, it is possible to question the very starting point for foundational externalism - the search for facts that determine necessary and sufficient conditions. Underlying Putnam's reasoning is the idea that since the associated descriptions fail to provide

necessary and sufficient conditions for the application of natural kind terms, such conditions have to be provided some other way, by features of the external environment. To support the idea that the external environment should provide such conditions, externalists have appealed to Kripke's claim that theoretical generalizations such as 'Water is H₂O' express *identities* and hence are necessary if true (Kripke 1980; Putnam 1975, 231). If, therefore, 'water' picks out H₂O in the actual world, being H₂O constitutes both a necessary and a sufficient condition for something to fall under the term. However, as recent discussions have made clear there are many problems with this suggestion deriving from the fact that natural kind terms are not plausibly construed as singular terms (in their most basic use) but rather as predicates (Soames 2002).¹⁷ Hence, the idea that theoretical generalizations of this sort are necessary if true, cannot be defended simply by appealing to the doctrine of the necessity of identity. Rather, it seems clear that controversial, essentialist assumptions about natural kinds have to be made. (For a criticism of these assumptions see for instance Dupré 1981, Häggqvist 2005, Needham 2000).

Someone who is skeptical of the modal claims made on behalf of the externalist may therefore react rather differently to Putnam's observation about associated descriptions, and simply say that what it shows is just that we should be skeptical of the search for necessary and sufficient conditions - whether provided internally or externally - and accept that there will always be indeterminacies. Whether 'water' applies to twin-water, on this view, is not something that is determined by our use of 'water' but something that calls for a decision (see LaPorte 1996 and 2004). It is worth noting that Putnam himself appears to have moved in this direction, for instance in his paper 'Is Water Necessarily H₂O?' (1990). Putnam discusses the question of what we are to say about the application of 'water' when it comes to hypothetical worlds that obey different laws, for instance a world in which H₂O behaves nothing like in our world:

Is it clear that we would call a (hypothetical) substance with quite different behavior *water* in these circumstances? I now think that the question, 'What is the necessary and sufficient condition for being water *in all possible worlds*?' makes no sense at all. And this means that I now reject 'metaphysical necessity'. (1990: 70)

4. Externalist semantics

4.1 Object-dependence

Like foundational externalism, externalist semantics has its roots in the opposition to descriptivism. Drawing on Kripke's arguments against descriptivism, it is suggested that some terms lack all descriptive content and that, therefore, the semantic value of the term is dependent on the existence of the external object referred to (Kripke 1980). Unlike foundational externalism, as stressed above, externalist semantics is a theory concerning the *semantic value* of certain terms, not a theory within foundational semantics.

We need to distinguish between two versions of externalist semantics. First, there is the Millian claim that the only semantic contribution made by a singular term, such as a proper name, is its referent. That is, the semantic value of the term is exhausted by the object referred to:

(ES) If *T* is a singular term, then the semantic value of *T* is the external object referred to by *T*.

Of course, this is a very minimal notion of 'externalism', given that the very idea underlying the Millian account of proper names is that meaning is nothing over and above reference. However, applied to the level of thought, it does support the thesis that certain types of thoughts (those containing a singular element) have an object-dependent content and this plausibly qualifies as an interesting form of externalism: Thinking the thought in question presupposes the existence of the external object.

Second, there is the thesis of object-dependent *senses*, championed by among others Gareth Evans and John McDowell (Evans 1982, McDowell 1977 and 1986). This is an attempt to reconcile externalist and Fregean elements.¹⁸ The sense is object-dependent since it cannot be understood in terms of any general descriptions but depends on the external object that the thought is about. If the object in question is absent there is no thought expressed. At the same time these object-dependent senses are Fregean, since they are individuated by the individual's attitudes: If the subject takes different attitudes towards thought T1 than towards thought T2 (believing one while disbelieving the other) then the thoughts have a different content - even if the thoughts are about the same object (Evans 1982: 18; McDowell 1986: 142). Hence, while the sense of a singular term, a singular concept, is distinct from the external object (that is the point of speaking about senses), it is dependent on it:

(ES*) If *T* is a singular term, then the semantic value of *T* is an object-dependent sense.

Although clearly different, externalist semantics and foundational externalism are often not kept apart but indiscriminately labeled 'externalism' (Burge 2007:11 objects to this habit). This may be a result of the tendency to characterize externalism in terms of failure of supervenience: Foundational externalism, again, entails that meaning (and content) fails to supervene on internal facts, but so does externalist semantics - simply imagine a counterfactual situation where the external object is absent or replaced by a duplicate, and it follows that the term has a different semantic value despite the fact that the speakers are internally identical. (For instance, Oscar in world 1 utters 'Bush won the election' and Oscar in world 2 utters the same words, only in his world there is no Bush but merely an evil demon producing impressions as of Bush.) This is one reason why the characterization of externalism in terms of failure of supervenience is unfortunate.

Externalist semantics faces the familiar difficulty of what to say about cases of reference failure. This difficulty is all the more serious when externalist semantics is applied to the level of thought. The idea of object-dependent singular thoughts implies that a subject may be under the impression that she thinks such a thought when she does not. This appears to pose a threat to self-knowledge, a consequence McDowell happily endorses because of its radical 'anti-Cartesian implications' (McDowell 1986: 146). However, even if one wishes to get rid of the Cartesian conception of self-knowledge, there are reasons to worry about the implications of accepting illusions of thoughts - in particular if the Fregean requirement of accounting for cognitive perspective is taken seriously. From the point of view of the individual, after all, it is as if there was a thought available, one that she reasons with and acts on (Oscar in world 2 votes against Bush in the elections, takes part in rallies, etc). How can this be explained if one endorses (either version of) externalist semantics?

A popular response to this difficulty that has emerged recently is the suggestion that if the object is absent then the semantics of the thought expressed is affected. There is a thought there, only it is not a singular thought but a general one (Brown 2004; Ludlow 2003). Whether a term such as 'Hesperus' expresses an object-dependent concept or not, therefore, depends on whether the term succeeds in referring. The difficulties facing externalist semantics concerning empty thoughts are thus avoided by falling back on descriptivism: Instead of having an illusion of thought the individual is simply mistaken in thinking that she has a singular thought. The semantics, as it were, becomes a posteriori.

It is debatable whether this proposal, 'a posteriori semantics', provides a satisfactory solution to the problem of empty thoughts. It seems plausible to hold that whether a term has a descriptivist semantics or an object-dependent one is determined by our semantic intentions. According to the proposal under consideration, however, our intentions cannot play this role since even if we intend 'Hesperus' to have an object-dependent semantics, it may

have a descriptivist one. Moreover, a posteriori semantics introduces a potential gap between how the individual reasons with the thought in question and the actual logical and conceptual connections of the thought. As Ludlow points out, if the semantics of the thought is a posteriori so is its logical form. As a consequence, he suggests, the very notion of what counts as a logical inference "will come unhinged" (2003: 411). If so, it is a serious question whether a posteriori semantics helps with the problem it was designed to solve, that of explaining the individual's cognitive perspective. (See Häggqvist & Wikforss 2007 for a critical discussion.)

4.2 Externalist semantics and natural kind terms

Theses (ES) and (ES*) concern singular terms. The question is whether a version of externalist semantics can be applied to the case of kind terms. The standard assumption is that this can be done, and that there are thoughts that are 'kind-dependent' (McLaughlin & Tye 1998: 292). However, it is far from clear that this is correct. There are of course those who defend the Millian position in the case of natural kind terms. These terms, it is argued, lack all descriptive content and function like singular terms: they are directly referential (LaPorte 1998, Marti 2004). However, notice, the object referred to is then taken to be the *kind*, an abstract object, and is thus not to be found in the external environment of the individual.¹⁹ To turn the Millian position into an externalist thesis, additional arguments are required. One proposal is that the semantic value of natural kind terms is constituted by *instances* of the kind (Ben-Yami 1991, 177). But it is very difficult to see how instances of a kind, individual tigers or puddles of water, could constitute the semantic value of a *kind* term. Perhaps, rather, it is the mereological sum of all tigers, etc. that constitute the semantic value of the term? However, this has the counterintuitive implication that kind terms such as 'water' and 'tiger' could not have the same meaning in a world where the instances of the kind are slightly different (a few gallons less, a few animals more).

There are therefore reasons to believe that the thesis of object-dependence cannot be adapted to the case of kind terms. If this is correct, it follows that in the case of these terms, foundational externalism is the only type of externalism available. Moreover, it undercuts another common argument in support of the claim that the meaning of natural kind terms is *relational* in the sense mentioned above, i.e. the claim that meaning *water* by 'water' (possessing the concept of water) requires having interacted with instances of water. If externalist semantics does not apply to kind terms then natural kind terms do not have a semantic value that depends on there being instances of the kind in the individual's environment.

This leaves us with the question what semantics these terms *should* be given - a very large issue, obviously, that cannot be properly discussed here. It should be pointed out, however, that one may accept foundational externalism, thesis **(FE)**, without thereby being committed to the Millian wholesale rejection of descriptivism. Even if it is granted that associated descriptions do not determine meaning independently of the environment it does not follow that the terms in question lack all descriptive content. Showing the falsity of descriptivism as a theory concerning the semantic value of natural kind terms requires further arguments.²⁰

Several proposals are available in the literature of how to combine descriptivism with foundational externalism. One is the suggestion mentioned above, that the relevant descriptions contain an indexical element.²¹ A related proposal, currently gaining in popularity, is the so-called neo-descriptivism defended by two-dimensionalists, according to which natural kind terms express two distinct intensions, one of which is to be understood descriptively and captures the speaker's cognitive perspective. For instance, 'water' is to be understood in terms of a rigidified description along the lines of 'the watery stuff of our acquaintance' (Chalmers 1996, Jackson 1998). This implies that an utterance, such as 'Water

is wet', express two distinct propositions, one of which is determined externally, the other internally. Hence, two-dimensionalism combines **(FE)** with descriptivism only at the price of splitting the content expressed into two.

However, it is also possible to combine foundational externalism with a form of descriptivism without splitting the content into two: By accepting that the externalist determination carries over to the descriptive content. For instance, it might be held (in a vaguely Davidsonian vein) that the meaning of Oscar's term 'water' is determined by Oscar's pattern of holding sentences true, in a context ('Water is wet', 'There is water in the glass', 'Water is a natural kind', 'That is water', etc. etc.). Thus, Oscar's use of the term, together with the context, serves to determine that his term has a certain descriptive content that, in turn, determines a certain extension. Indeed, this may be one way of developing Putnam's suggestion that stereotypical descriptions such as wet, thirst-quenching etc. capture the cognitive or conceptual aspect of meaning, even though they do not suffice to determine extension (1975: 249-251, 269). The important difference is that it is recognized that external factors may play a role in the determination of the content of the descriptions as well.²²

Although foundational externalism is typically coupled with a Millian account of natural kind terms, therefore, it should be clear that it is quite possible to endorse foundational externalism without being committed to the claim that these terms lack all descriptive content. Foundational externalism is a thesis within foundational semantics, whereas the Millian position is a position within descriptive semantics, and while they draw on related arguments it is important to see that the theses are relatively independent of one another.

5. Psychological externalism

Foundational externalism and externalist semantics, again, are both species of *semantic* externalism. Sometimes, however, 'externalism' is used to refer not to a semantic thesis, but to a thesis concerning the individuation of *types of psychological states*; more precisely, the thesis that certain types of psychological states are *relational* in character, in the sense that being in such a state requires standing in a relation to a certain feature of the external environment. In order to distinguish this species of externalism from semantic externalism, let us call it *psychological externalism*:

(PE) For all psychological states of type S, the individual is in a state of this type only if she stands in an appropriate relation to external fact E.

Psychological externalism is very widely endorsed. Indeed, externalism is often simply characterized as a thesis about psychological states, along the lines of **(PE)**. For instance, in the *Stanford Encyclopedia of Philosophy*, John Lau characterizes content externalism as follows: 'Externalism with regard to mental content says that in order to have certain types of intentional states (e.g. beliefs), it is necessary to be related to the environment in the right way' (Lau 2003:1). Sanford Goldberg, similarly, characterizes externalism as the thesis that 'some propositional attitudes depend for their individuation on features of the thinker's (social and/or physical) environment' (Goldberg 2005:135). And in his recent book, Burge characterizes anti-individualism as a thesis about the nature of psychological kinds: Relations with the wider environment, Burge writes, 'are constitutively necessary for the states and events to be the specific kinds of states or events that they are' (2007:3).

How is semantic externalism related to psychological externalism? It is uncontroversial that there is a small class of psychological states that are relational or 'wide' in the sense of **(PE)**; i.e. factive states, such as 'seeing that *p*' or 'knowing that *p*'. Putnam also proposes that jealousy is such a state, since *x is jealous of y* entails that *y* exists (1975, 220).

Semantic externalism is perceived by many to promise a radical expansion of this class: It is taken to show that regular de dicto beliefs, such as the belief that water is wet, are relational in just the sense that factive states are.²³

However, as should be clear from above, the semantic support for psychological externalism is much more scant than is ordinarily recognized. Thus, standard foundational externalism, thesis **(FE)**, fails to support the thesis that there are certain *types* of content that are determined externally (even if **(FE)** is assumed to apply to mental content). In the case of *Oscar*, the content of 'Water is wet' is determined by facts in his external environment, but this does not show that thinking this thought requires the presence of these external facts, since it does not exclude the possibility that the same content is determined internalistically (in the case of an expert, say). Consequently, **(FE)** does not support the claim that psychological states involving the concept water are relational in the sense of **(PE)**.

Externalist semantics, thesis **(ES)**, does of course support the claim that certain types of intentional states are relational - this is part and parcel of the notion of object-dependent contents. If one endorses externalist semantics it therefore follows that psychological externalism holds for certain types of psychological states, i.e. those containing the required singular element (*believing that Bush will soon retire or hoping that that apple is not rotten*). However, as noted above, externalist semantics cannot be adapted to the case of kind terms. Hence, it cannot be employed to show that psychological externalism holds in the case of psychological states containing kind concepts.²⁴ Neither foundational externalism, nor externalist semantics, underwrites the commonly made claim that psychological externalism holds for psychological states containing kind concepts (such as *believing that water is wet or hoping to find gold*).

Burge, in fact, grants that his version of externalism does not entail that certain types of contents are relational. The relevant constitutive relations, he says, need not be

relations to the thing represented: "One can think about aluminum as aluminum even though one's cognitive system never interacts with aluminum. For instance, a Martian scientist could theorize about the structure of aluminum and imagine its macro-properties." (2007:3). Burge also rejects object-dependent semantics and grants that one may have concepts that apply to nothing, such as phlogiston or Santa Claus (2007: 4; 1982: 97-98). Thus, Burge both rejects externalist semantics and accepts the weaker construal of foundational externalism, thesis **(FE)** instead of thesis **(FED)**. At the same time, he claims that relations with the wider environment are constitutively necessary for thinking certain kinds of thought. Burge's idea is that there must, in general, be some causal contact between an individual's thoughts and the objects that her thoughts are about (he speaks of 'collateral causal connections'). This may well be plausible as an empirical thesis about the individual's cognitive development: as a matter of fact, acquiring a language, and developing one's cognitive capacities, requires complex relations with one's physical and social environment. However, Burge insists that his point 'is not merely developmental' but *constitutive* (2007:2). It is far from clear how to understand this claim, much less defend it. Moreover, even if there is some philosophical argument to the effect that, by and large, an individual must interact with her environment to have contentful thoughts, it does not follow that any particular type of content is such that thinking it requires standing in a relation to something in the individual's 'wider environment'.

Externalism about kind terms therefore does not have the radical implications for philosophy of mind that it is usually taken to have – even if it is assumed to extend to mental content. The Twin Earth argument, if accepted, supports a thesis within foundational semantics, thesis **(FE)**. This, again, suffices to accomplish what Putnam sets out to accomplish, i.e. reject the idea that the meaning of kind terms is always determined 'internally'. However, it does not support his grander claims, such as the famous dictum 'meanings just ain't in the head', nor does it support psychological externalism.²⁵ What follows

is merely that in the case of *some* terms (and we may not know which prior to scientific investigations) external factors *sometimes* (when the speaker is not very knowledgeable) play a role in the determination of meaning and content. Although not uninteresting, this thesis hardly constitutes the sea change in philosophy of mind and language that externalism has been touted as.²⁶

References

- Ben-Yami, H. 2001. "The Semantics of Kind Terms", *Philosophical Studies*, 102: 155-184.
- Bilgrami, A. 1992. *Belief and Meaning*, Oxford: Blackwell
- Boghossian, P. 1989. "The Rule-Following Considerations", *Mind* 98.
----1997. "What the Externalist Can Know A Priori." *Proceedings of the Aristotelian Society*, vol. XCVII.
- Brown, J. 2004. *Anti-Individualism and Knowledge*. Cambridge, Mass: MIT Press.
- Brueckner, A. "What An Anti-Individualist Knows A Priori", *Analysis* 52: 111-118.
- Burge, T. 1979. "Individualism and the Mental", *Midwest Studies in Philosophy* 4: 73-121.
---- 1982. "Other Bodies", *Thought and Object. Essays on Intentionality*, ed.A. Woodfield, Oxford: Clarendon Press.
----1986a. "Intellectual Norms and the Foundations of Mind", *Journal of Philosophy* 83: 697-720.
----1986b. "Cartesian Error and the Objectivity of Perception", in *Subject, Thought, and Context*, eds. P. Pettit & J. McDowell, Oxford: Clarendon Press.
----1989. "Wherein is Language Social?", *Reflections on Chomsky*, ed. A. George, Oxford: Blackwell, 175-191.
----1993. "Concepts, Definitions, and meaning", *Metaphilosophy* 24: 309-325.
----2007. "Introduction", *Foundations of Mind*, Oxford: Oxford University Press, 1-31.
- Chalmers, D. 1996. *The Conscious Mind*. Oxford: Oxford University Press.
- Davidson, D. 1987. "Knowing One's Own Mind", *Proceedings of the American Philosophical Association*: 441-458.
- Davies, M. 1998. "Externalism, Architecturalism, and Epistemic Warrant", in Wright, Smith & Macdonald. (eds.) 1998: 321-362.
---- 2000. "The Problem of Armchair Knowledge", in Nuccetelli 2003: 23-56.

- Donnellan, K. 1993. "There Is a Word for That Kind of Thing: An Investigation of Two Thought Experiments", in *Philosophical Perspectives, 7. Language and Logic*, ed. J.E. Tomberlin, Atascadero, Ca: Ridgeview.
- Dupré, J. 1981. "Natural Kinds and Biological Taxa", *Philosophical Review*, XCI, 66-90.
- Evans, G. 1982. *The Varieties of Reference*. Oxford: Clarendon Press.
- Farkas, K. 2003. "What is Externalism?", *Philosophical Studies*, 112: 187-208.
- 2006. "Semantic Internalism and Externalism", in *The Oxford Handbook of Philosophy of Language*, E. LePore & B. Smith eds.
- Fodor, J. 1980. "Methodological Solipsism Considered as a Research Strategy in the Cognitive Sciences", *Behavioral and Brain Sciences*, 3: 63-73.
- Gallois, A. 1996. *The World Without, The Mind Within. An Essay on First-Person Authority*, Cambridge: Cambridge University Press.
- Glüer, K. (forthcoming) "Critical Notice: Donald Davidson's Collected Essays", *Dialectica*.
- Goldberg, S. 2005. "The Dialectical Context of Boghossian's Memory Argument", *Canadian Journal of Philosophy* 35: 135-148.
- Hurley, S. (forthcoming). "Varieties of Externalism", in press in R. Menory, ed. *The Extended Mind*, Ashgate Press.
- Häggqvist, S. 2005. "Kinds, Projectibility and Explanation", *Croatian Journal of Philosophy*, 13: 71-88.
- Häggqvist, S & Wikforss Å. "Externalism and a posteriori semantics", forthcoming in *Erkenntnis* (online publication July 7, 2007).
- Jackson, F. 1998. *From Metaphysics to Ethics*, Oxford: Clarendon Press.
- Kaplan, D. 1989. "Demonstratives". In *Themes from Kaplan*, eds. J. Almog, J. Perry, and H. Wettstein. Oxford: Oxford University Press.
- Kripke, S. 1980. *Naming and Necessity*. Oxford: Blackwell.
- LaPorte, J. 1996. "Chemical Kind Term Reference and the Discovery of Essence", *Nous*, 30: 112-132.
- 1998. "Rigidity and Kind", *Philosophical Studies* 97: 293-316.
- 2004. *Natural Kinds and Conceptual Change*, Cambridge, MA: Cambridge University Press.
- Loar, B. 1985. "Social Content and Psychological Content", in *Contents of Thought*, eds. H. Grimm & D. Merrill, University of Arizona Press.

- Lau, J. 2003. "Externalism About Mental Content", *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/content-externalism>
- Lewis, D. 1994. "Reduction of Mind", in *A Companion to the Philosophy of Mind*, ed. S. Guttenplan, Oxford: Blackwell, 412-431.
- Ludlow, P. 2003. "Externalism, Logical Form and Linguistic Intentions", in *Epistemology of Language*, Oxford: Oxford University Press, 399-414.
- Marti, G. 2004. "Rigidity and General Terms", *Proceedings of the Aristotelian Society* 104: 129-146.
- McDowell, J. 1977. "On the Sense and Reference of a Proper Name", *Mind* 86: 159-185.
 ----1986. "Singular Thought and the Extent of Inner Space", in *Subject, Thought, and Context*, Pettit, P. and McDowell, J. (eds). Oxford: Clarendon Press: 137-168.
- McGinn, C. 1989. *Mental Content*. Oxford: Blackwell.
- McLaughlin B. and Tye M. 1998. "Externalism, Twin Earth, and Self-Knowledge", in Wright, Smith & Macdonald (eds.) 1998: 285-320.
- McKinsey, M. 1987. "Apriorism in the Philosophy of Language", *Philosophical Studies* 52: 1-32.
 ----1991. "Anti-Individualism and Privileged Access". *Analysis* 54: 9-16.
 ----2002. "Forms of Externalism and Privileged Access", *Philosophical Perspectives* 16: 199-224.
- Needham, P. 2000. "What is Water?", *Analysis*, 60, 13-21.
- Nuccetelli, S. (ed.) 2003. *Semantic Externalism and Self-Knowledge*, Cambridge, Mass: MIT Press.
- Pagin, P. 1997. "Is Compositionality Compatible with Holism?", *Mind & Language*, 12: 11-33.
 ----2006. "Meaning Holism", *Handbook of Philosophy of Language*. Oxford: Oxford University Press: 213-32.
- Pessin, A. and Goldberg S. eds. 1996. *The Twin Earth Chronicles*, New York: M.E. Sharpe.
- Putnam. H 1970. "Is Semantics Possible?", reprinted in *Mind, Language and Reality, Philosophical Papers*, vol. 2, Cambridge, MA: Cambridge University Press, 215-271.
 ----1975. "The Meaning of 'Meaning'", in *Minnesota Studies in the Philosophy of Science*, vol. 2, ed. K. Gunderson, Minneapolis: University of Minnesota Press.
 ----1990. "Is Water Necessarily H₂O?", *Realism with a Human Face*, Cambridge, MA: Harvard University Press, 54-79.
 ----1996. "Introduction", in Pessin A. & Goldberg S. 1996: xv-xxii.
- Rowlands, M. 2003. *Externalism. Putting Mind and World Back Together Again*, Chesham: Acumen.

- Searle, J. 1983. *Intentionality*. Cambridge: Cambridge University Press.
- Segal, G 2000. *A Slim Book about Narrow Content*, Cambridge, Mass: MIT Press.
- Soames, S. 2002. *Beyond Rigidity: The Unfinished Semantic Agenda of Naming and Necessity*, Oxford: Oxford University Press.
- Stalnaker, R. 1989. "On What's in the Head", *Philosophical Perspectives 3: Philosophy of Mind and Action Theory*.
- 1997. "Reference and Necessity", in *A Companion to Philosophy of Language*, B. Hale & C. Wright eds. Oxford: Blackwell: 534-554.
- Wikforss, Å. 2001. "Semantic Normativity", *Philosophical Studies*, 102: 203-226.
- 2004. "Externalism and Incomplete Understanding", *The Philosophical Quarterly*, 54: 287-294.
- 2005. "Naming Natural Kinds", *Synthese*, 145: 65-87.
- C. Wright, B. C. Smith, and C. Macdonald eds. 1998. *Knowing Our Own Minds*, Oxford: Clarendon Press.

¹ As Peter Pagin [1997: 13] has argued, the notion of *determination* of meaning is not an epistemic notion but, rather, a metaphysical one.

² This is not to say that the set does not also include internal facts. For instance, many externalists require some minimal level of competence on part of the speaker such that a speaker cannot mean *water* by 'water' unless there is a reasonable set of associated descriptions in place (see Burge 1979, 90-91).

³ See Burge (2007, 153) who rejects the characterization of externalism, or anti-individualism, in terms of failure supervenience (2007, p. 153). For a good discussion of externalism and the supervenience thesis see also Hurley (forthcoming).

⁴ It should be noted that Putnam later changes his view, and endorses foundational externalism with respect to content as well (1988, 1996).

⁵ This does not yet exclude the possibility of saying that Oscar and Toscar associate all the same descriptions with their respective term: Wet, transparent, thirst quenching etc. However, if the content externalist wishes to extend the externalism to these other concepts as well, the associated descriptions will of course be different.

⁶ In general, we must separate the notion of 'determination' that belongs to foundational semantics (concerning the facts that determine meaning), from the notion that belongs to the level of semantic content (concerning how the intension determines the extension).

⁷ Putnam, too, appeals to the role of the social environment when he famously suggests that there is 'division of linguistic labor' (1975: 227). However, unlike Burge, Putnam limits the role of the social environment to a determination of extension, and does not extend it to concepts and mental content.

⁸ For a nice discussion of these differences see Donnellan 1993. See also McKinsey 2002.

⁹ This is sometimes disputed on the grounds that social and physical factors *jointly* determine meaning. It is not clear what to make of this idea of a 'joint determination', however, since the social and physical factors manifestly deliver different results. If, instead, the idea merely is

that the descriptions of the experts help *fix* the extension, the resulting theory is not a species of social externalism (since the fixing descriptions do not *determine* meaning).

¹⁰ Burge 1986a puts forth a thought experiment that he suggests does not rely on the assumption of incomplete understanding. However, it should be noted that the thought experiment does not support social externalism but a version of physical externalism. Moreover, it is far from clear that the experiment does not, after all, presuppose that the individual has an incomplete understanding of the relevant concepts (see for instance Wikforss 2004).

¹¹ It is very commonly assumed that the thought experiment does support (**FED**). See for instance Davies 1998: 327-328; McGinn 1989: 35-36; Nuccetelli 2003: 2-7; Rowlands 2003:5.

¹² This, it should be clear, is reminiscent of the difficulty afflicting social externalism, noted above, concerning whether there are cases where it could plausibly be said that the individual makes a genuine linguistic error and yet should be ascribed the standard concept.

¹³ This observation also plays a role in Kripke's criticisms of descriptivism. Scott Soames labels it 'the semantic argument' (2002, 266).

¹⁴ Notice, however, that although the indexical construal is compatible with foundational internalism with respect to *meaning* it is not compatible with foundational internalism with respect to thought content - at least not on the assumption that thought content is truth-conditional. The full truth-conditional content, after all, is determined by the meaning in conjunction with the external context.

¹⁵ What characterizes true twin-substances (water and twater) is that there are no differences on the macro-physical level. By contrast, actual 'twin-substances' (such as jadeite and nephrite, gold and fool's gold) always involve differences on the macro-physical level. As Putnam himself later acknowledged: 'differences in microstructure invariably (in the actual world) result in differences in lawful behavior' (1990, 69).

¹⁶ As David Lewis points out, this assumption seems to be a remnant of old-style 1950's versions of descriptivism (1994: 424). See also Glüer, forthcoming, and Segal 2000:124.

¹⁷ This has spurred a debate concerning whether there is an interesting notion of rigidity that applies in the case of natural kind terms. See for instance Marti 2004.

¹⁸ For this reason the position is sometimes labeled 'Fregean externalism' (see for instance Brown 2004, pp. 20-21).

¹⁹ Clearly, if by 'external factors' one includes abstract objects, the externalist thesis threatens to become utterly trivial (Burge runs this risk in 2007: 154). For instance, it would follow that a foundational internalist who construes meanings as abstract objects qualifies as an externalist.

²⁰ Soames suggests that in addition to the semantic argument Kripke employs two further arguments against descriptivism: his famous modal argument and what Soames calls the epistemic argument (2002, 19).

²¹ Although, again, the indexical proposal is not in fact compatible with thesis (**FE**) since the idea is that the external determination does not extend to meaning. However, it may extend to thought content (on the assumption that thought content is truth-conditional) and if so the indexical account does combine descriptivism about meaning with foundational externalism about thought content.

²² Burge, too, appears to combine foundational externalism with a form of descriptivism. For instance, Burge is happy to speak of definitions, although he stresses that these are not determined by our linguistic practices (individual or social), but rather by the nature of the external objects themselves (1986; 1993; 2007: 9-12).

²³ This idea also underlines the heated discussion concerning the compatibility of externalism and self-knowledge, more precisely, the debate following Michael McKinsey's reductio argument against the compatibilist position (McKinsey 1991; 2002).

²⁴ McKinsey is well aware of these difficulties. He suggests that we construe natural kind-terms as relational by appealing to the idea that their meaning is given by a semantic rule that contains a genuine term, such as 'that group' (1987).

²⁵ As Davidson points out, Putnam's dictum is not a happy formulation of the externalist thesis in any case (not even of the stronger foundational thesis, thesis **FED**, mentioned in section 3.3 above). That sunburn presupposes the existence of the sun does not imply that sunburn is not a condition of the skin (1987: 451).

²⁶ Thanks to Katalin Farkas, Sören Häggqvist, Michael McKinsey and Peter Pagin for very helpful discussions of central issues of the paper. Special thanks to Kathrin Glüer for invaluable comments on earlier drafts.