SEMANTIC EXTERNALISM:
From Wittgenstein
to Putnam

by

Peter T. Bahr

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts

at

Dalhousie University
Halifax, Nova Scotia
August 2001

SEMANTIC EXTERNALISM:
From Wittgenstein
to Putnam

by

Peter T. Bahr

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts

at

Dalhousie University
Halifax, Nova Scotia
August 2001

*To my parents, without whose help and support none of this
would have been possible.*

Contents

Abstract

It is not widely acknowledge that Putnam's initial foray into semantic externalism is Wittgensteinian. Some philosophers (Anthony Rudd, for one) argue that, as he has become more Wittgensteinian in recent years, Putnam has effectively abandoned his own early externalism. In this thesis I suggest that this perception is mistaken.

I argue, contrarily, that Putnam's early externalism, circa "The Meaning of 'Meaning'", is a continuation of the account of language offered by Wittgenstein in the *Philosophical Investigations*. Specifically, I argue that: (1) Wittgenstein and Putnam share a negative thesis, whereby meaning is not determined by psychological facts about speakers; and (2), they share a positive thesis—according to which, the environment itself has a role to play in determining meaning—as the 'use' in Wittgenstein's 'meaning is use' is identified in terms of a speaker's interactions with the social and physical environment in which the speaker is situated. That is, Putnam's semantic externalism just is Wittgenstein's view that use determines meaning.

Acknowledgements

I owe a special thanks to my supervisor Mike Hymers, and to my readers Duncan MacIntosh and Bob Martin. Thanks also to J.S. Bahr, Jennifer Epp, and Rockney Jacobsen for editing, reading early drafts, and helpful discussions.

Introduction

The present Anglo-American philosophical landscape is a diverse and interesting one, but with close scrutiny one notices some related trends. Three examples come readily to mind: philosophy of mind is slowly taking priority over philosophy of language (i.e., discussions about meaning are *now* about intentionality); cognitive science programmes flourish both in the United States and the United Kingdom; and, finally, in the relevant literature, the works of Donald Davidson and Hilary Putnam are regularly subjected to rigorous critical scrutiny. This suggests, I think, that internalism—the idea that meaning is dependent upon psychological states of speakers—is a received theory, of sorts, in the modern philosophical climate.

But not everyone is seduced by this view. Davidson, Putnam, Ludwig Wittgenstein, and, perhaps, W.V.O Quine have all mounted serious attacks against the received view. Quine, for example, thinks that if meaning is something mental, then it is nothing more than a mysterious and superfluous entity, while for Davidson internalism involves an epistemic priority of the mental (over language) that he believes is better avoided. My concern is the manner in which both Putnam and Wittgenstein, in particular, urge us towards anti-internalism. They both reject internalism and replace it with an account of language in which meaning is dependent upon speakers' interactions with their social and physical environment.[1] Call this view semantic externalism.

---

[1] Davidson, too, advocates such view, but for present purposes I am only concerned with the similarity between Putnam and Wittgenstein's views.

1

It is not my task to take a side in this debate (although I side with the externalist), or to persuade anyone of the strengths of one view as opposed to the other. My goal is a far humbler one. In this thesis, I intend to demonstrate that the work of Wittgenstein and Putnam represent a continuous strand of anti-internalist thought. That is, I demonstrate that a significant degree of continuity exists between their externalisms. Anthony Rudd argues that these two views cannot be integrated and that they are distinct species of externalism. I suggest he is wrong.

In Chapter One, I survey a number of internalist positions in order to situate subsequent discussions. Although, at first glance, the discussion may seem somewhat scattered, it will serve to draw out the different ways in which internalism can be formulated. I shall focus, in particular, on the internalisms of John Locke and Bertrand Russell. The discussion of these two philosophers allows me to introduce two types of internalism: first, one in which psychological states *are* the referents of our terms—that is, words immediately refer to some mental object; and second, one in which meaning may not be identical to a psychological state, but nonetheless the psychological state of a speaker *determines* meaning. (The importance of separating these out becomes apparent later as both Wittgenstein and Putnam address these accounts separately.) I conclude this chapter with a brief summary of how these views are integrated into modern philosophy and the various guises in which we may find modern internalism

In the second chapter I discuss the externalist nature of Wittgenstein's thought. It is not my goal in this chapter to provide a novel interpretation of Wittgenstein's

work. Rather, I examine three aspects of his philosophy that I take to be overtly externalist and which I believe are relevantly similar to points made by Putnam. To that end, I do the following: first, I discuss Wittgenstein's remarks on naming as they consist of a rejection of the idea that something mental can determinately fix meaning or reference; second, I analyze his private diarist example and argue that it may be read as a refutation of the internalist idea that a mental object enters into the meaning of a term; and third I suggest that Wittgenstein's doctrine of 'meaning is use' is an externalist one because meaning is determined by the use a speaker makes of a term, where usage is identified in terms of interactions with the speaker's social and physical environment.

In Chapter Three I arrive at the heart of my argument. Here I present an analysis of certain features of Putnam's philosophy that demonstrates the continuity to be found between his works and the ideas of Wittgenstein, which were presented in the previous chapter. First, I argue that Putnam's 'Twin Earth' thought-experiment is a modern variant of the private language argument, and of the private diarist example in particular. Second, I point to similarities between Putnam's discussion of intentionality and Wittgenstein's conclusions about naming. Third, I suggest that Putnam's 'semantic externalism' can be identified with, or seen as an instantiation of, the Wittgensteinian 'use determines meaning' doctrine. And finally, I respond to a number of objections, raised by Rudd, which claim that the two accounts are not similar species of externalism.

1

Internalism

My overall aim in this thesis is to demonstrate that a continuity exists between the externalisms of Wittgenstein and Putnam. That is, I intend to show that their respective positions represent a continuous anti-internalist strand of thought. Before discussing this continuity, however, it is important to understand the view that they oppose. As Putnam claims:

> To explain what is wrong with the way philosophers and cognitive scientists have generally approached questions about meaning, it will be necessary to examine a number of different ways in which the standard approach has manifested itself.[2]

That is, to fully appreciate the reaction one must understand what is being reacted against. Accordingly, in this chapter I undertake an analysis of a variety of positions that fall broadly under the rubric of internalism. The various theories I survey do not exhibit any overt continuity, but instead serve as a sort of scenic overview of the internalist landscape. Yet, despite this variety, the theories to be discussed share the idea that meaning is, in some important way or other, determined by something internal to the speaker. To begin, we turn our attention to the initial origins of internalism and the works of Aristotle.

The Origins of Internalism

Aristotle is an important figure in the history of the philosophy of language,

---

[2] Hilary Putnam, *Representation and Reality* (Cambridge, Mass: MIT Press, 1988), p. 4.

for it is with the writing of his *De interpretatione*[3] that one encounters the beginnings of internalist (or, alternatively, 'traditional') theories of meaning. He writes that "spoken words are the symbols of mental experience and…the mental experiences, which these directly symbolize, are the same for all."[4] Thus, with a simple pen-stroke, Aristotle introduces a particular picture of language that has come to dominate philosophy for over two millennia—a conception of meaning that has become so entrenched that it is virtually dogma in many areas of philosophy. The key postulate of the traditional theory (and that which makes it internalist) is simple: for any word, sign, term, etc. a mental representation is associated with it in the mind of a speaker or hearer, and this mental representation determines what the word, sign, or term means.

Now, to say that this particular picture of language has become dogma is not to say that it has remained static. This traditional 'mental representation' has undergone a number of changes. It has been construed variously as experience, idea, picture, sense-datum and concept. But what has remained unchanged is the idea that some 'mental thing' (no matter how it is construed) is an essential aspect of the meaning or reference of a term. That is, a term could not refer to what it does without some associated mental representation.

It is this very tenet—that meaning depends on something internal to the speaker—that Wittgenstein's and Putnam's respective externalisms seek to counter. It is important to note, however, that although I take Aristotle to be the originator of

---

[3] Aristotle, *De Interpretatione* in *Aristotle*, ed. W.D. Ross (New York: Scribner, 1938).

[4] Aristotle, *De Interpretatione*, p. 7.

internalist thought, it is not obvious that Aristotle's account is itself internalist. This is a matter of debate and various commentators have differing views. Some take Aristotle to be explicitly internalist (e.g., the Putnam of *Representation and Reality*), while others hold that his conception of the mental differs from our own to such an extent that we cannot 'read back' internalism onto his position (e.g., the Putnam of *Words and Life*).[5] But the core idea is still there—that words symbolize something mental. Regardless of Aristotle's thinking, it is this idea, borrowed from Aristotle, that later philosophers appropriated and developed into the internalism which survives today. However, since this is a debate better avoided (for present purposes at least), let us begin our foray into internalism proper with the introduction of internalism into modern philosophy—the 'idea theory' of John Locke.

John Locke: the Idea Theory

In *An Essay Concerning Human Understanding*,[6] Locke's ultimate concern is to discover the limits or scope of our epistemological practices—that is, the limits of human knowledge. Yet to say that this is his sole concern would be to underestimate the breadth of Locke's thinking. Locke's concern *is* epistemological, but one also finds the recognition that our epistemological practices cannot be understood outside of our linguistic practices. He writes:

---

[5] For the views of these 'two Hilary Putnams' see: Putnam, *Representation*, pp. 19-20; and Hilary Putnam, "How Old is the Mind?" *Words and Life*, ed. James Conant (Cambridge, Mass: Harvard UP, 1994), pp. 3-5.

[6] John Locke, *An Essay Concerning Human Understanding*, ed. Raymond Wilburn (London: Dent, 1948).

> It is impossible to speak clearly and distinctly of our knowledge, which all consists in propositions, without considering, first, the nature, use, and signification of Language.[7]

In other words, one cannot understand knowledge *simpliciter*, without first understanding the vehicle through which that knowledge gets expressed—namely, language.

For Locke, the proper function of language is the communication or expression of one's ideas. In his account, all language is understood to be descriptive, constative, declarative, etc. The *telos* of language, then, is solely to make known one's ideas or mental life[8]—it is to report what is in one's head to another.

Since the primary function of language is to communicate one's ideas it seems only natural, for Locke, that the referents of words just are ideas. He states, "the use, then, of words, is to be sensible marks of ideas; and the ideas they stand for are their proper and immediate signification."[9] The picture that Locke is advocating, then, is quite simple: the referent of any word is an idea. When one utters the word 'cat', for example, what one immediately refers to with that word is a cat-idea, and not a cat. But the Lockean picture is not yet complete.

Lockean ideas are understood to be private objects—they are, as he puts it, 'hidden in our own breasts'. In referring one does not signify some publicly held idea (e.g., something akin to a Platonic Form or a Fregean *Sinn*), rather one immediately

---

[7] Locke, II, xxxiii, 19.

[8] Locke also claims that a function of language is to record one's thoughts (i.e., it assists in remembering), but I do not take this to be an 'end' for Locke, because it is a prerequisite on reporting one's ideas that one remembers those ideas. See III, i, 1 through III, ii, 2.

[9] Locke, III, ii, 1.

refers only to one's *own* idea. So. the Lockean view is not merely that a word refers to an idea. but that "words, in their primary or immediate signification, stand for nothing but *the ideas in the mind of him that uses them*."[10] Moreover. Locke takes this to be necessarily so.

Suppose that words could stand for something other than one's own ideas. Locke takes this to be a contradictory state of affairs—to say 'a word refers to something other than one's idea' is circumlocution for 'words. both. do and do not refer to one's ideas'.[11] To see why he takes this to be contradictory one must be cognizant of the fact that, for Locke, individuals are always and only aware of their own ideas. He writes:

> Ideas we receive from sensation and reflection are the boundaries of
> our thoughts; beyond which the mind, whatever efforts it would make.
> is not able to advance one jot; nor can it make any discoveries, when it
> would pry into the nature and hidden causes of those ideas.[12]

Perceptions, experiences, thought, etc. are always mediated through ideas—that is, all awareness is of one's own ideas and of nothing more. So (given this notion that awareness is mediated through ideas) to say that a word refers to something other than one's own idea is to assert the following: first, a word refers to one's own idea only if it does not refer to something else;[13] second, that some particular word does, in fact, refer to something other than one's idea; and third, that this same word refers to one's

---

[10] Locke, III, ii, 2.

[11] Locke, III, ii, 2.

[12] Locke. II, xxiii, 29.

[13] Thanks to Michael Hymers for pointing out this implicit premise that Locke needs for his argument to succeed.

own idea (as that is all one is ever cognizant of anyway). Simply, a word does and does not refer to something other than one's idea. Thus, contrary to the assumption above, a word can only refer to one's own ideas.

The theory of language we find emerging in Locke, then, is this: a word immediately signifies or refers to an idea of a particular speaker, and it is the proper function of a word to express its associated idea to an audience. So, for example, a speaker's token utterance of 'The cat is on the mat' immediately refers only to the speaker's cat-idea and mat-idea, and a particular relation that the speaker takes to hold between these two ideas; and, moreover, the purpose of this utterance is to express this (more complex) idea to some audience.

Although this particular picture is nice, simple and succinct, a problem arises, and the nature of it should be fairly obvious: we have no reason whatsoever to expect that successful communication can take place. In fact, if anything, we should expect the contrary. For any speaker's utterance, if it is to refer to anything at all, it can only refer to one's own idea. Likewise, for the audience—the words 'The cat is one the mat' can immediately refer only to their own ideas. The problem is this: if ideas are hidden and private, as Locke takes them to be (i.e., they are 'hidden in our own breasts'), then there is no reason to believe that people ever understand one another. What the audience understands with the speaker's utterance are their own ideas, which are signified by the utterance, and *not* the idea that the speaker desired to express. So, there is no guarantee that we are communicating, in spite of the fact that communicating is supposed, by Locke, to be the proper function of language.

In recognition of this defect. Locke claims that we must make some further suppositions in order to understand how communication is possible. First. we must assume that one's own ideas match those of others. People must suppose. he claims. that "their words...[are] marks of the ideas in the minds also of other men."[14] The speaker's cat-idea and the audience's cat-idea must match or else, as demonstrated above. communication is not taking place. If this 'matching assumption'[15] is not made, then speaker and audience are at cross-purposes, with hidden meanings, and never expressing or understanding the desired ideas.

Yet surely this matching assumption is unfounded. It seems that the only reason one ought to suppose ideas match in the manner discussed, is if one is John Locke, wishing to ensure that communication is possible within his theory. However, Locke believes that we *can* safely suppose that ideas match, given a second supposition—the 'representational assumption'. This supposition is that one's ideas, and derivatively one's words, stand for or represent "the reality of things."[16] Words and ideas, according to Locke, represent "whether they be only *constant effects*, or else *exact resemblances* of something in the things themselves."[17] An idea represents some thing when it is caused by and/or resembles that same thing. In the most basic cases, one's idea represents some particular thing because it resembles whatever thing

---

[14] Locke, III, ii, 4.

[15] I borrow the terms 'matching assumption' and the subsequent 'representational assumption' from Gregory McCulloch, *The Mind and Its World* (London: Routledge, 1995), pp. 33-4.

[16] Locke, III, ii, 5.

[17] Locke, II, xxx, 2.

the idea is about—the idea 'pictures' the thing. A speaker's cat-idea, for example, represents a cat because the idea looks like a cat. In more complex cases, an idea represents by virtue of being caused by some particular thing. So, a speaker's idea of red would be caused by an object reflecting wavelengths of light between 650 and 700 nanometres, which enter the eye, stimulating the photoreceptors in the retina, etc.

But what has representation to do with the matching assumption? If ideas are caused by or resemble some objective mind-independent thing, then the referents of our terms (e.g., our ideas) are not purely subjective entities—they have an objective component. So we *can* suppose that the ideas of speaker and audience match. If a speaker's idea resembles some particular thing, then so too does a hearer's idea. More correctly, if a speaker's idea resembles some object and a hearer's idea resembles that same object, then the two ideas resemble each other—that is, they match.[18] They match because they both represent some thing by resembling it. One's cat-idea is the same as another's cat-idea because they both represent, and resemble, cats. So, successful communication can occur, in Locke's account, because the idea produced in an audience (i.e., what the audience understands), by a speaker's utterance, matches the idea that the speaker desired to express.

With these two suppositions in place Locke believes that we have an accurate account of reference and language in general. The Lockean account of language, then, may be summarized as follows: first, it is the proper function of language to express a speaker's X-idea to some audience (i.e., to communicate some idea to an

---

[18] Although resemblance *simpliciter* is not transitive, Lockean resemblance is transitive precisely because it involves an objective form of picturing.

audience); second, the referent of a speaker's utterance 'X' just is the speaker's X-idea; third, an idea, which is the immediate object of awareness, represents some thing X, by virtue of being caused by or resembling X; fourth, since an X-idea represents X (via causality or resemblance), a speaker's and audience's X-ideas match; and, therefore, successful communication occurs as a speaker's utterance of 'X' produces in the audience an association with the audience's X-idea, which matches that of the speaker. Consider, for example, a speaker's utterance of 'the cat is on the mat'. This sentence is uttered in order to express the speaker's idea that the cat is on the mat, and what the utterance immediately refers to just is that idea. Moreover, this idea represents the 'way things are' because the speaker's idea resembles an actual cat on a mat (think here, perhaps, of the idea as a mental picture of a cat on a mat). The result is that this utterance causes the idea of a cat on a mat in the mind of the hearer. Thus, language succeeds in its communicative function because the hearer's idea matches that of the speaker, as it, too, resembles a cat on a mat—that is, the speaker's idea gets expressed to the hearer.

Although my concern in this chapter is mainly exegetical (i.e., to historically situate the writings of Putnam and Wittgenstein), a few final comments are perhaps in order to show why later philosophers rejected the Lockean account. First, recall that, in Locke's account, successful communication occurs only because we suppose that people's ideas match; and this, in turn, is justified only because we suppose that words and ideas represent objects. That is, according to Locke, the matching requirement is satisfied because ideas resemble or are caused by 'the reality of things'. But are these

assumptions at all tenable? Consider the case of resemblance: since ideas bear a resemblance relation to those things that they are ideas of, one's idea will also resemble that same idea in another. But this position, which was meant to secure the matching of ideas, is not one that is available to Locke. To appreciate why, recall that Locke holds a representational theory of perception whereby experiences are always mediated through ideas. The problem is that if all one is ever aware of is an idea, then one can never 'get past' the idea to some object to compare the idea against. That is, one cannot know if an idea resembles some non-idea thing because all one can ever be aware of is an idea—the comparison would only be betwixt idea and idea.[19] But if a speaker and audience cannot know that an idea resembles the thing it represents, then they have no guarantee that their ideas will match during any linguistic transaction.

It would seem that, at best, all these interlocutors can know is that some particular thing *causes* their idea. Yet, even taking this latter tack, Locke's matching assumption cannot be satisfied. Analyzing representation causally is problematic because causality is paradigmatically an external relation: as David Hume argued, one can have knowledge of an event without having knowledge of its cause.[20] It follows, therefore, that a speaker's and audience's respective ideas, during any given linguistic interaction, may be caused by different things. In other words, the interlocutors can have knowledge of the idea, without having concomitant knowledge of the cause.

---

[19] As Berkeley somewhat trivially quipped, "an idea can be like nothing but an idea": See George Berkeley, *A Treatise Concerning the Principles of Human Knowledge*, in *The Works of George Berkeley Bishop of Cloyne*, ed. T.E. Jessup, vol. 2 (London: Nelson, 1949), I, 8.

[20] David Hume, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, (Oxford: Clarendon, 1949), I, II, xiv. For further discussion on this point see Michael Hymers, *Philosophy and its Epistemic Neuroses* (Boulder, Co.: Westview, 2000), pp. 76-7.

Again, our speaker and hearer cannot know that matching occurs. Thus, either way Locke construes representation (i.e., as a causal relation or a resemblance relation), it fails to satisfy the matching requirement. Thus his account is left wanting, burdened by the difficulty that initially motivated the two assumptions—that is, there is no reason to suppose that successful communication can take place.

The more fundamental problem with the Lockean account of language, however, is that the doctrine is, in some sense, self-refuting—it is a linguistic non-starter of sorts. Locke tells us that language was created so that people may express their ideas:

> The comfort and advantage of society not being to be had without communication of thoughts, it was necessary that man should find out some external sensible signs, whereof those invisible ideas, which his thoughts are made up of, might be made known to others.[21]

So, in order for speakers to know what their words mean, they must first know the content of their own minds. That is, the mental is prior to meaning. But notice that if this is correct, then meaning cannot be what is used to know the contents of one's own mind (i.e., one needs to know what one is thinking, but in order to do so, one needs to know the meaning of one's thoughts). Yet without meaning, how can speakers possibly 'understand' their own ideas? If meaning cannot be used to know one's ideas, then one cannot properly be said to understand the idea either.[22] Similarly, if

---

[21] Locke, III, ii, 1.

[22] This follows because 'meaning' and 'understanding' are correlative concepts—to understand just is to have knowledge of meaning, or something of the sort. So, understanding is a mental state determined by the meaning of that being understood. In this case, since there is no recourse to meaning, there is no proper object of understanding.

an idea cannot be meaningfully described, even to oneself, then there is nothing comprehensible to express (or that one might wish to express). So, whatever might be in one's mind, it cannot be an expressible idea. Ideas, in Locke's view, begin to look radically inexpressible—akin to a 'something I know not what'. Somewhat paradoxically, then, language was created for the expression of ideas that speakers cannot (prior to language, at least) even express to themselves. But what need is there for language if a speaker has nothing to express? In short, if Locke is correct, it is not clear how language could ever 'get off the ground' as it were.

It should be apparent, from the preceding discussions, that Locke's view does not offer an adequate account of our linguistic practices. Accordingly, many of Locke's views on language were rejected in the years subsequent to the 1690 publication of his *Essay*. Yet, at the same time, much of the Lockean picture remains intact, and has survived three centuries of scrutiny to become widely accepted as received theory. Two notions, in particular, survive: first, the basic idea that language consists primarily of names referring to objects; and second, the view that the referent of a term just is some mental object. Both ideas recur in the work of Bertrand Russell.

Russell: Proper Names and Definite Descriptions

Russell, in response to seminal works on language and logic by Gottlob Frege, undertakes an analysis of proper names—that is, names that refer to some one particular thing (whether person, place or object). And, although Russell does not follow Locke in proposing a general theory of language, his account, of this particular subset of language, is similar to Locke's in that it is an internalist one. As one might

expect. then. for Russell. proper names refer (or denote) by virtue of a mental representation in the mind of a speaker.

In "On Denoting",[23] Russell lists three puzzles that he believes a theory of reference ought to be capable of solving: first, it must allow for the non-truth functionality of substitution into intensional, or referentially opaque, contexts (e.g., belief contexts, propositional attitude contexts, modal contexts, etc.); second, it must give a satisfactory account of how bearerless names are meaningful, despite lacking a referent; and third, it must allow for the denial of some particular thing's 'being', without paradoxically asserting its existence.[24] While each of these is interesting on its own terms, for present there is only one which we need consider as a framework in which to discuss Russell's account of proper names. Consider, then, non-truth functional contexts.

The occurrence of a name, word, etc. (i.e., the context) is said to be extensional or referentially transparent if the substitution for that name of a co-referring term is truth functional or truth preserving. A name's occurrence is intensional or referentially opaque, contrarily, if the substitution of a co-referential term is non-truth functional. Consider some examples to aid in fixing the terminology.

(1) It is true that Scott was the author of Waverly.

---

[23] Bertrand Russell, "On Denoting," *Mind* 14 (1905): 479-93; reprinted in Bertrand Russell, *Logic and Knowledge*, ed. Robert Charles Marsh (London: Allen & Unwin, 1956). References are to the latter.

[24] Russell, "Denoting", pp. 47-8.

Consider, also, that the names 'Sir Walter' and 'Scott', from (1), refer to the same person—they are co-referential terms. That is:

(2) Scott is Sir Walter.

It follows (by 'substitutivity of identity') that:

(3) It is true that Sir Walter is the author of Waverly.

If Scott is Sir Walter, then whatever is true of one is true of the other—if Scott is, in fact, the author of Waverly, then so too is Sir Walter. So, in (1), the context '_____ was the author of Waverly' is said to be extensional or referentially transparent because the substitution of a name for another with the same reference occurs without a change in the truth-value of the embedding sentence.

Now, compare the above argument with the following:

(4) King George IV believed that Scott was the author of Waverly.
(5) Scott is Sir Walter.
(6) Therefore, King George IV believed that Sir Walter was the author of Waverly.

Despite the superficial similarities between the two arguments, they are different in one important way: the first argument is valid and the second is not, despite the grammatical similarity between the two. The conclusion of the latter could, in fact, turn out to be false, even though (4) and (5) may be true. So, the context 'King George IV believed that _____' in (4) is intensional or referentially opaque because the substitution of coextensive terms into this context is non-truth functional—the substitution does not preserve the truth of the original sentence. Preliminaries aside, let us now turn to Russell's attempt to solve the problem of intensional contexts and his theory of proper names.

To make explicit why intensional contexts are thought to be problematic, consider a logical analysis of (4) through (6). Suppose that: B is a 'King George IV believed that' operator, such that for any proposition P, $B(P)$ yields 'King George IV believed that P'; Wx is a predicate, such that 'x is the author of Waverly'; a denotes Scott; and, b denotes Sir Walter. Suppose, also, that George IV does, in fact, believe that Scott is the author of Waverly, but, at the same time, does not believe that Sir Walter is the author of Waverly (i.e., he lacks the knowledge that Scott is Sir Walter). Using standard notation, (4) and (5) may be analyzed as:

(7) $B (Wa)$
(8) $a=b$

It follows, by Leibniz's Law or the Sub Id inference rule that:

(9) $B (Wb)$

One is, then, left with two related problems. First, (9), which just is (6) in logical notation, is derived from the true (7) and (8) (i.e., (9) ought to validly follow), despite the fact that it is false. That is, (7) is true and (9) is false, even though Scott is the author of Waverly if and only if Sir Walter is the author of Waverly (i.e., Wa if and only if Wb). And second, there is a certain respect in which (9) may be understood as true. King George does, in fact, believe that the individual denoted by 'Sir Walter' is the author of Waverly, although he may never assent to such a claim.[25] He believes, in other words, that *that person* wrote Waverly, no matter what that person may be

---

[25] What he lacks is not knowledge that 'that particular individual' is the author of Waverly. He does know this because he knows that Scott, who just is 'that particular individual', is the author of Waverly. All that he lacks is the knowledge that 'that particular individual' is called Sir Walter. In other words, read *de re* (9) is true, although read *de dicto* it is false.

called. So. the puzzle for Russell is to give an account of reference that offers an adequate solution to both of these problems.

Russell's handling of referentially opaque contexts involves denying that (7) through (9) are the correct analyses of (4) through (6). These problems arise. for Russell, only if one analyzes terms such as 'Scott' and 'Sir Walter' as proper names, rather than as definite descriptions—that is, if one mistakes proper names for definite descriptions. Russell's account of proper names, then, is essentially reductionist in nature as it involves the "reduction of all propositions in which denoting phrases occur to forms in which no such phrases occur."[26] That is, so-called proper names (i.e., denoting phrases) are to be reduced to definite descriptions.

Yet what justifies such a reduction? This reduction is warranted, thinks Russell, because of the different roles played by names and descriptions in a language. A genuine name, for Russell, is defined as follows:

> [It] is a simple symbol, directly designating an individual which is its meaning. and having this meaning in its own right, independently of the meanings of all other words.[27]

A genuine proper name, then, does no more, and no less, than denote an individual or a particular thing. But, urges Russell, if one looks at the usage of what we take to be proper names, one finds that they do not function in this manner. While considering the status of existential claims, for example, Russell writes:

And so, when we ask whether Homer existed, we are using the word

---

[26] Russell, "Denoting", p. 45.

[27] Bertrand Russell, *Introduction to Mathematical Philosophy* (London: Allen & Unwin, 1970), p. 174.

"Homer" as an abbreviated description: we may replace it by (say) "the author of the Iliad and the Odyssey." The same considerations apply to almost all uses of what look like proper names.[28]

The point here is that, if a term like 'Homer' were to function as a genuine name, one could not meaningfully inquire into Homer's existence—precisely because a name refers to an existent individual. That is, if the term functions as a name, it, in effect, tacitly asserts the existence of its referent. The only way in which existential claims make sense, thinks Russell, is if so-called proper names are actually descriptions. So, for Russell, to inquire into an individual's existence is not to ask 'Does *that* individual exist?' because, of course, the individual exists—that was just confirmed *in* asking the question (i.e., by pointing to *that*). Instead, such a query purports to describe some individual (e.g., the use of 'Homer' describes the individual who wrote the Iliad and the Odyssey), and then inquires as to whether any individual satisfies that description, or something of the sort. Russell concludes, in the case of existential claims at least, that so-called proper names do not function as genuine names.

Russell generalizes this point to extend beyond the merely existential use of proper names. He states:

We may even go so far as to say that, in all such knowledge as can be expressed in words—with the exception of "this" and "that" and a few other words of which the meaning varies on different occasions—no names in the strict sense occur, but what seem like names are really descriptions.[29]

So, the role of a so-called proper name is not to denote an individual, but to describe

---

[28] Russell, *Mathematical Philosophy*, p. 179.

[29] Russell, *Mathematical Philosophy*, p. 178.

one. That is, when one uses a proper name "the individual, instead of being named [i.e., denoted], is being described as the person having that name."[30] So, for Russell, proper names are merely stand-ins or short forms for descriptions; and what is meant by a name just is some set of uniquely describing characteristics. Our well worn 'Scott', for example, does not pick out *that* particular individual; rather it describes an individual as 'called Scott', 'born in Edinburgh', 'author of Ivanhoe', etc. The name describes the unique individual, if any, that is Scott.

The idea that proper names are merely descriptions in disguise gives Russell the means to demonstrate how (7) through (9), above, misrepresent the logical structure of (4) through (6). Recall that the problem with intensional contexts was that:

(7) $B\,(Wa)$
(8) $a=b$
(9) Therefore, $B\,(Wb)$

appears (syntactically) to be a valid argument, despite the fact that (9) is false. Russell's contention is that this is problematic only if the constituents 'Scott' and 'Sir Walter' are analyzed as proper names. But, if the names 'Scott' and 'Sir Walter' are taken as abbreviated descriptions, then the argument yields a different analysis. Suppose, with Russell, that 'Scott' is an abbreviation for the individual called Scott (Sx), who was born in Edinburgh (Ex) and is author of Ivanhoe (Ix). What George IV believes, then, in believing that Scott wrote Waverly (Wx), is not (7), but:

---

[30] Russell, *Mathematical Philosophy*, p. 174.

(7′) $B\ [(\exists x)((Sx\ \&\ Ex\ \&\ Ix\ \&\ Wx)\ \&\ (\forall y)(Wy \rightarrow x=y))]$[31]

which roughly states that King George believes there is an individual called Scott, born in Edinburgh, and author of Ivanhoe, and that same individual is the writer of Waverly, if anyone is. Similarly the identity claim in (8) is construed, not as an identification of individuals, but as the claim that there is one unique individual that is called, both, 'Scott' and 'Sir Walter' (SWx):

(8′) $(\exists x)[Sx\ \&\ SWx\ \&\ (\forall y)(Sy \rightarrow x=y)\ \&\ (\forall z)(SWz \rightarrow x=z)]$

That is, there is an individual called 'Scott' and 'Sir Walter' and anything that is called 'Scott', and anything that is called 'Sir Walter' just is that same individual. Russell's solution, here, is twofold: first, there is no constituent 'Scott' (i.e., the a in the previous formulation) in (7′) for which one can substitute 'Sir Walter' (i.e., from $a=b$) to yield the false (9); and second, the identity claim in (8′) is not an identity claim *simpliciter*, but a claim about a unique individual that has two particular properties (i.e., being called by two names) and, as such, does not admit of substitution into (7′).[32] Thus, (9), or in Russellian terms:

(9′) $B\ [(\exists x)((SWx\ \&\ Ex\ \&\ Ix\ \&\ Wx)\ \&\ (\forall y)(Wy \rightarrow x=y))]$

may, in fact, be false, but this is no longer of concern because it cannot be derived

---

[31] This is not precisely how Russell would formulate such a statement. In order to ensure the uniqueness of the individual called Scott, we would need to add: for all y, if it is S and E and I, then it is x. The correct formulation, then, is: there is one and only one individual called Scott, born in Edinburgh, author of Ivanhoe, and that unique individual wrote Waverly (e.g., $B\ [(\exists x)((Sx\ \&\ Ex\ \&\ Ix\ \&\ (\forall y)((Sy\ \&\ Ey\ \&\ Iy) \rightarrow x=y))\ \&\ (\forall z)(Wz \rightarrow x=z))]$): See Russell, *Mathematical Philosophy*, pp. 175-80. However, (7′) suffices for the sake of simplicity.

[32] Russell, "Denoting", pp. 51-2. For an analysis, similar to the one I have given, of Russell's solution to the problem of intensional contexts, see Alexander Miller, *Philosophy of Language* (Montreal: McGill-Queen's UP, 1998), pp. 67-9.

from the true (7') and (8').

Lest this be thought a long digression, we must consider the internalist nature of Russell's account and its relevance to this discussion. As discussed at length. Russell takes proper names to be abbreviated descriptions or descriptions in disguise. The meaning of 'Scott', for example, may be given by the following:

(10) $(\exists x)[(Sx \ \& \ Ex \ \& \ Ix \ \& \ \ldots x) \ \& \ (\forall y)((Sy \ \& \ Ey \ \& \ Iy) \rightarrow x=y)]$

In other words, 'Scott' means that there is one unique individual thing that is called Scott, is born in Edinburgh, and authored Ivanhoe. Russell is clear, however, that what is important, for meaning, is the description held in the mind of a speaker. Consider the following:

> The thought in the mind of a person using a proper name correctly can generally only be expressed explicitly if we replace the proper name by a description. Moreover, the description required to express the thought will vary for different people, or for the same person at different times.[33]

And:

> Thus, for example, when we make a statement about Julius Caesar…we have in mind some *description* of Julius Caesar… Thus our statement does not mean quite what it seems to mean, but means something involving, instead of Julius Caesar, some description of him which is composed wholly of particulars and universals with which we are acquainted.[34]

There are a number of things in these passages that need sorting out to fully appreciate Russell's internalism. First, there is the now familiar idea that the meaning of a proper name is given by some description. That is, the meaning of 'Scott', or 'Julius

---

[33] Bertrand Russell, *The Problems of Philosophy* (London: Oxford UP, 1959), p 54.

[34] Russell, *Problems*, p. 59.

Caesar' for that matter, can be given by something similar to (10) above. Second, the meaning of a proper name depends, in some sense, on the definite description held in the mind of a speaker. It depends, in other words, on the mental or psychological state of a speaker. So, the description that a particular speaker associates with 'Scott' may differ from the description that a hearer associates with 'Scott'. And finally, the meaning of the description is ultimately reducible to, and its meaning determined by, the constituents with which a speaker is acquainted. These latter two points are the entrance to the robust internalism of Russell's account.

For Russell, 'Scott' is to be analyzed as 'the man who was called Scott', 'the author of Ivanhoe', 'the man who was born in Edinburgh', or something of the sort— that is, the proper name is to be analyzed as a description. And the meaning of 'Scott' is determined by the particular definite description. Moreover, different speakers may associate different descriptions with the same name. Russell's overall picture, although somewhat simplified, is that the description in the mind of a speaker is to be analyzed into its simplest meaningful parts (e.g., genuine proper names, propositional functions, and logical connectives) and the meaning of these 'simples' determines the meanings of the description, and, derivatively, of the proper name.[35] A speaker, for example, may use 'Scott' as an abbreviation for 'the man called Scott' and 'the author of Ivanhoe', while a hearer may associate the name with 'the man called Scott' and 'the author of Waverly', such that their respective beliefs about Scott are given by the following (where B is now a 'speaker believes that'-operator):

---

[35] Russell, "Denoting", pp. 42-4.

(11) B $[(\exists x)(Sx \ \& \ Ix \ \& \ (\forall y)((Sy \ \& \ Iy) \rightarrow x=y))]$
(12) B $[(\exists x)(Sx \ \& \ Wx \ \& \ (\forall y)((Sy \ \& \ Wy) \rightarrow x=y))]$

Now, since meaning is determined compositionally—that is, by its parts—what 'Scott' means differs for each of these interlocutors. Thus, Russell's account is, broadly speaking, internalist because the meaning of a proper name, or at the very least, what a speaker understands by that name, depends upon the psychological state (i.e., the description held in mind) of a speaker.

Although Russell's account, construed thusly, is evidently internalist, it becomes even more so once one considers that, for Russell, the meaning of a description is determined by things with which *we* are acquainted. That is:

> We must attach *some* meaning to the words we use, if we are to speak significantly and not utter mere noise; and the meaning we attach to our words must be something with which we are acquainted.[36]

As discussed above, a Russellian analysis of a proposition containing proper names will result in a logically correct proposition that contains only simples—genuine proper names, universals, propositional functions, logical connectives, and the like. Returning to the example of 'Scott', suppose that its final reduction yields 'the man who was called Scott and is the author of Waverly'. This proposition is comprised of a number of simples, such as the universals 'man' and 'author', a uniqueness function in 'the', the connective 'and', and the genuine proper names 'Scott' and 'Waverly'.[37]

---

[36] Russell, *Problems*, p. 58.

[37] Suppose for the sake of a simpler analysis that 'Waverly' is functioning as a proper name indicating *that* book—in actuality, it, too, would need further analysis until there remained only logical simples. Note, that 'Scott' is not similarly problematic as it indicates, not Scott, but *that* particular sound or *phone*.

Russell contends that if these simples are to be meaningful at all. then the meaning must be ‘something with which we are acquainted’.

But what is it with which we are acquainted? For Russell. as for Locke before him. all we are ever immediately acquainted with are mental experiences—or sense-data.[38] The object of a speaker’s acquaintance with ‘Waverly’ is not Waverly (i.e.. *that* book), but a sense-datum of (or caused by) that book. Similarly, in ‘Scott’, one is acquainted with the sense-datum of the sound produced by an utterance of S-c-o-t-t. Russell’s account becomes overtly internalist once one recognizes that the meaning of a simple *just is* the ‘something with which we are acquainted’. That is, the meaning of a simple is the sense-datum of a speaker. [39] So, it is not only that proper names differ in meaning because of the differing descriptions in the minds of speakers. But, also, the meanings of proper names differ because the simples, which the proper names reduce to, refer to the sense-data of an individual speaker.[40]

Repacking this motley of related ideas allows Russell to *explain* referential opacity. The problem, it will be recalled, is that one cannot validly infer from ‘King George IV believed that Scott was the author of Waverly’ and ‘Scott is Sir Walter’ that ‘King George IV believed that Sir Walter is the author of Waverly’. The

---

[38] Russell, *Problems*, pp. 7-12, 54-5.

[39] If it is not obvious that the meaning of the simples is identical to the things with which we are acquainted, which just happen to be sense-data, then recall that, for Russell, the meaning of a simple is the thing it designates: see Russell. *Mathematical Philosophy*, p. 174.

[40] This led Russell to ultimately conclude, “a logically perfect language, if it could be constructed, would not only be intolerably prolix, but, as regards its vocabulary, would be very largely private to one speaker. That is to say, all the names that it would use would be private to that speaker and could not enter into the language of another speaker.”: see Bertrand Russell, *The Philosophy of Logical Atomism*, in *The Collected Papers of Bertrand Russell*, vol. 8 (London: Allen & Unwin, 1986), p. 176.

explanation is that, since the meanings of proper names depend on the psychological state of a speaker (i.e., both the description held in mind, and the individual's sense-data), what George IV believes by 'Scott is the author of Waverly' differs from the belief that 'Sir Walter is the author of Waverly'. Moreover, one cannot derive the latter belief from the former because, as Russell puts it, "we can make no judgements concerning this man [Scott] which involve knowledge about him beyond what the description gives."[41] That is, unless King George's description of Scott contains 'the man who is also called Sir Walter', or something of the sort, it is impossible for the former belief to entail the latter.

Like Locke's theory before his, Russell's account of proper names is susceptible to a number of criticisms—chief amongst them, the critiques of Wittgenstein and Putnam to follow. However, even philosophers, such as John Searle, who follow Russell in offering a descriptivist treatment of proper names, have pointed to problems with Russell's theory. In "Proper Names",[42] Searle argues that the Russellian account ought to be rejected as it leads to several counter-intuitive results. To illustrate the problems that arise with Russell's descriptivism, recall that, for Russell, a proper name is equivalent in meaning to some unique set of descriptions. In the example above, 'Scott' is equivalent to the set of characteristics (e.g., the man called Scott, born in Edinburgh, author of Ivanhoe, etc.) that uniquely describe Scott. So, for an object x to satisfy the description abbreviated by 'Scott' just

---

[41] Russell, *Problems*, p. 56.

[42] John Searle, "Proper Names," *Mind* 67 (1958): 166-73; reprinted in *The Philosophy of Language*, ed. A.P. Martinich, 3rd ed. (New York: Oxford UP, 1996). References are to the latter.

is for *x* to have all of the characteristics described of that object. That is, in equating

the proper name to a set of definite descriptions, we are, in effect, giving the necessary

and sufficient conditions to determine whether an object is to be called by that proper

name.

To demonstrate one of the problems with Russell's account, Searle supposes

that our history books turn out to be incorrect. It is not the case that Scott was born in

Edinburgh; rather, he was born in Glasgow. In fact, we could imagine this happening

for any one of the properties or characteristics in the definite description. This is

problematic for Russell because he seems forced to conclude that either the meaning

of 'Scott' has changed or else Scott did not exist;[43] and neither of these alternatives

seems satisfactory. To see why, reconsider the logical notation of (10). In this

scenario, Searle is saying that the characteristic 'being born in Edinburgh' (the *Ex* of

(10)) is false of Scott. Note also, that if any one of the characteristics contained in

(10) (e.g., *Sx* & *Ex* & *Ix*) is false, then the sentence itself is false. But, if *Ex* is false,

this amounts to saying that 'it is not the case that there exists one and only one object

called Scott, born in Edinburgh, etc.' That is, Scott does not exist. The only

alternative is to say that the meaning of 'Scott' has changed. This requires a simple

rewriting of the definition of 'Scott' omitting the problematic characteristic *Ex*. But

this, also, amounts to a change in the meanings of our words each and every time

some new historical discovery is made. The problem, then, is that both of these

responses, which seem 'forced' by Russell's theory, are somehow too strong. In

---

[43] Searle, "Proper Names", p. 250. Searle's account uses the name 'Aristotle'; for continuity, however, I have chosen to reformulate his objections using Russell's 'Scott' example.

giving necessary and sufficient conditions, Russell's descriptivism invites these counter intuitive results whenever a characteristic that is part of the definite description fails to obtain.

The second source of trouble with Russell's theory is the idea that a proper name is equivalent in meaning to a unique definite description. Since proper names are simply abbreviations for definite descriptions, and, as such, are semantically equivalent to the descriptions, the proper name can always be replaced with the definite description. To best appreciate this problem, consider the usage of proper names in a language. Generally speaking, a speaker uses a name in a language to pick out an object, such that 'this name refers to that thing'. The upshot of this usage is that the name can be used to refer to an object; there is no need to find the object, pick it up, point to it, etc. on each occasion that a speaker wishes to refer to it. However, Searle notes that if Russell is correct and a proper name is just a definite description in disguise, then this function could be performed equally well by the description. That is, a linguistic community could agree to stop using the name 'Scott' and instead refer to him by the description 'the author of Ivanhoe'. The problem is that it is then a necessary truth that the man referred to (i.e., *that* man) is the author of Ivanhoe; but, in fact, it is a contingent matter whether or not Scott wrote anything at all.[44] The fundamental problem, here, is that the Russellian account of proper names, in some manner, conflates the necessary with the contingent. Russell's theory results in, what

---

[44] Searle, "Proper Names", pp. 252-3.

Searle calls, a "hoary metaphysical system."[45]

There is a final critique of Russell's solution to the problem of referential opacity that is directly related to his internalism. Russell's theory of definite descriptions is an attempt to give a solution to the problem that is, in a sense, 'better behaved' than Frege's. Frege's solution was to say that in referentially opaque contexts the referent of a proper name was its sense or meaning. So, insofar as 'Scott' and 'Sir Walter' may have different meanings (i.e., insofar as what is grasped by the speaker may differ), then it is possible for George IV to believe that Scott is the author of Waverly without believing that Sir Walter is. Russell, however, thinks it arbitrary to say that sometimes the referent of a proper name is an object and other times a meaning. So, Russell wants to propose, alternatively, an extensional solution to the problem—that is, a solution in which the meaning or truth of propositions containing proper names are determined only by the referents of those names (i.e., by the objects denoted). But here Russell fails.

Alexander Miller points to the following problem with Russell's solution. Suppose that 'Scott' and 'Sir Walter' are abbreviations for the following functions:

(13) $(\exists x)[(Sx \,\&\, Ix \,\&\ldots x) \,\&\, (\forall y)((Sy \,\&\, Iy) \rightarrow x=y)]$
(14) $(\exists x)[(SWx \,\&\, Wx \,\&\ldots x) \,\&\, (\forall y)((SWy \,\&\, Wy) \rightarrow x=y)]$

Now, the extensions of these functions are sets of ordered pairs, matching predicates with truth-values, that result when predicates are substituted for the ellipses. So, if 'born in Edinburgh' (Ex), 'author of Romeo and Juliet' (Rx), and 'married Charlotte Carpenter' (Mx) are substituted for the ellipsis in (13), the following set is generated:

---

[45] Searle, "Proper Names", p. 251.

$\{(E, T), (R, F), (M, T)\}$. The problem, Miller says, is that, since everything that is true or false of Scott is also true or false of Sir Walter, the sets of ordered pairs (i.e., the extension) for (13) and (14) are identical. In other words, these are not two different functions, but the same function with two different names.[46] That is, if Russell's analysis is truly extensional, then, if King George believes that Scott is the author of Waverly, he also believes that Sir Walter is the author of Waverly. Something has gone awry.

What has gone wrong here is that Russell's solution gives the appearance of working only because of his intensionalist claim, that King George's belief that there is an x who wrote Waverly and is called Scott differs from the belief that there is an x who wrote Waverly and is called Sir Walter.[47] That is, what is doing the work in Russell's solution is not his extensional analysis, but his internalism—if his solution works at all, it is only because different speakers have different sense-data and different definite descriptions in mind when using the terms 'Scott' and 'Sir Walter'. Moreover, insofar as these things determine reference, they function in exactly the same way as Frege's sense[48] So, Russell's solution is not an extensional one, different in kind from that of Frege. His solution does no more than remove Frege's sense from the so-called 'third realm' and locates them in the realm of the mental. By his own standards, Russell fails at his task.

---

[46] Miller, pp. 70-2.

[47] We previously encountered this in Russell's *explanation* of referential opacity.

[48] Thanks to Michael Hymers for discussion on the above points.

But Russell's tacit intensionalism, driven by his internalism, complicates matters further. Recall that Russell's solution to referential opacity was to analyze the premises 'King George IV believes that Scott is the author of Waverly' and 'Scott is Sir Walter' thusly:

(7′) $B [(\exists x)((Sx \;\&\; Ex \;\&\; Ix \;\&\; Wx) \;\&\; (\forall y)(Wy{\rightarrow}x{=}y))]$
(8′) $(\exists x)[Sx \;\&\; SWx \;\&\; (\forall y)(Sy{\rightarrow}x{=}y) \;\&\; (\forall z)(SWz{\rightarrow}x{=}z)]$

such that (8′) does not admit of substitution into (7′) to yield the false 'King George IV believes that Sir Walter is the author of Waverly'. If Russell's analysis is correct, however, consider what follows. Let T be a truth operator, such that for any proposition P, $T(P)$ yields 'It is true that P'. Now, according to Russellian analysis 'It is true that Scott is the author of Waverly' and 'Scott is Sir Walter' ought to be analyzed as follows:

(15) $T [(\exists x)((Sx \;\&\; Ex \;\&\; Ix \;\&\; Wx) \;\&\; (\forall y)(Wy{\rightarrow}x{=}y))]$
(16) $(\exists x)[Sx \;\&\; SWx \;\&\; (\forall y)(Sy{\rightarrow}x{=}y) \;\&\; (\forall z)(SWz{\rightarrow}x{=}z)]$

The problem, for Russell, is that if (8′) does not substitute into (7′), then neither will (16) substitute into (15); 'It is true that Sir Walter is the author of Waverly', which ought to validly follow, is not derivable. That is, if Russell's solution to referential opacity works, then referentially transparent contexts are not truth-preserving either. Russell's solution, in effect, merely flips the problem on its head—there is no longer a problem with referential opacity, but there is now one with referential transparency. If Russell is to reclaim the truth-preserving quality of transparent contexts, it seems he must analyze proper names differently for different contexts. But this is the same arbitrariness for which he admonishes Frege. That is, as Russell's internalism

functions similarly to Fregean sense. Russell seems forced into a position that requires the same type of ad hoc manoeuvre.

Having given a thorough explication of both Locke's and Russell's internalisms and some brief discussion about the problematic aspects of their respective accounts, a few final comments are in order regarding the modern day status of the traditional theory.

Internalism Modernized

As mentioned above, this chapter serves one purpose: to provide a survey of the internalist landscape so that we may better appreciate the criticisms of Wittgenstein and Putnam that follow. To that end, it may appear that there is no overt continuous thread binding the preceding discussions together. Locke's theory, on the one hand, involves an inquiry into the nature of general terms and communication in particular, while Russell's account deals only with a particular class of names and is, in effect, an attempt to solve a number of problems he inherited from Frege. But what they do share is their internalism—the idea that something mental plays an essential role in meaning or reference.

The choice of Locke and Russell as stalking horses, of sorts, was also particularly deliberate. With the work of these two authors we are presented with two species of internalism. Locke's theory introduces an internalism whereby the objects of reference just are the psychological states of a speaker—in this case, ideas. While with Russell's account, although names also refer to mental objects, the more important idea is the introduction of an internalism in which the psychological state of

a speaker fixes or determines the reference or meaning of a term.

In the course of the preceding discussions. I enumerated a number of potential problems with the Lockean and Russellian accounts. Yet. despite these problems. the internalist aspects of their theories are still present in modern day philosophy—in particular, the latter idea that something mental determines meaning. Modern internalist theories abound in a number of familiar guises. According to the pragmatic programme of H.P. Grice (and adherents such as Searle and Stephen Schiffer) meaning is ultimately reducible to a speaker's intention to induce a belief in an audience.[49] In the Gricean programme, meaning is to be explained in terms of psychological concepts such as beliefs and intentions—that is, meaning depends upon a speaker's *mental* act of intending and the *mental* content of one's thoughts. Noam Chomsky's work in linguistics is similarly internalist. According to Chomsky, everyone is born with an innate stock of concepts and language acquisition amounts to the assigning of words to those concepts.[50] (Chomsky takes this so seriously, in fact, that he supposes children are born complete with complex concepts. such as 'carburetor' and 'bureaucrat'). A final example of modern internalism is found in Jerry Fodor's 'language of thought hypothesis' and the cognitive science programme

---

[49] See, for example, H.P. Grice, "Meaning," *Philosophical Review* 66 (July 1957): 377-88; John Searle. *Speech-acts* (Cambridge: Cambridge UP, 1969); and Stephen Schiffer, *Meaning* (Oxford: Clarendon Press, 1972). Although, Schiffer now expresses reservations that the Gricean programme can be satisfactorily carried out: see Stephen Schiffer, *The Remnants of Meaning* (Cambridge, Mass: MIT Press, 1987).

[50] Noam Chomsky, "Language and Problems of Knowledge," *The Philosophy of Language*, ed. A.P. Martinich, 4th ed. (New York: Oxford UP, 2001), pp. 595-6.

(at least of the non-connectionist variety).[51] For Fodor. thought and representation require a medium. and this medium is 'mentalese'. The idea. albeit oversimplified. is that the direct objects of our beliefs. desires. and. most importantly for present purposes. understanding. are mental representations.[52] The reason. for example. that two terms (perhaps in two languages) mean the same things is that they both encode the same bits of mentalese. Meaning, then, is parasitic upon mental representations.

The point of these examples is simply to demonstrate that internalism survived and flourishes today. In fact, it flourishes so much that both exponents and detractors alike take it to be received theory.[53] And, lest one think I overestimate the ubiquity of the internalist position, consider that, in much philosophical discourse, philosophy of mind is now seen as central to. and inseparable from. philosophy of language, and that questions about meaning are often, now, treated as questions about intentionality.

In subsequent chapters, this received theory will be challenged. And, moreover, I shall argue that the criticisms of it by Wittgenstein and Putnam represent a continuous strand of thought—a strand that denies the very possibility of internalist accounts and sees meaning determined externally, rather than internally. To the beginnings of this, I now turn.

---

[51] See, for example, almost anything by Fodor, especially, Jerry Fodor, *The Language of Thought* (New York: Crowell, 1975); and Jerry Fodor, *Psychosemantics* (Cambridge, Mass: MIT Press, 1987).

[52] The importance of understanding is, again, that what one understands *is* a meaning.

[53] For examples of this idea, that internalism is received theory, in the internalist and externalist camps respectively, see Michael McKinsey, "The Internal Basis of Meaning," *Pacific Philosophical Quarterly* 72 (1991), p. 143; and Putnam, *Representation*, p. 19.

2

Wittgenstein

In the previous chapter, I summarized a number of positions that are internalist in nature. While some of these positions differ in a number of important ways, essential to them all was the idea that facts about meaning are determined by, or sometimes reducible to, facts about the psychological states of speakers. Near the end of that discussion I suggested that internalism is received theory and dominates much of philosophy. That being said, where internalism is presently the dominant view, it was once the only view; and it was not until the first half of the twentieth century that philosophers thought about challenging the received view. Wittgenstein was amongst the philosophers who began this challenge.

The purpose of the present chapter is threefold: first, I discuss two of the later Wittgenstein's criticisms of internalism—particularly, his naming considerations and the private language argument; second, I examine the view of language that Wittgenstein offers in place of internalist accounts—that is, the doctrine that meaning is determined by use; and third, I shall structure the discussion in such a way as to make later comparison with the work of Putnam possible.

The Early Wittgenstein

Wittgenstein's *Tractatus Logico-Philosophicus*[54] is a continuation of the

---

[54] Ludwig Wittgenstein, *Tractatus Logico-Philosophicus*, trans. D.F. Pears and B.F. McGuinness (London: Routledge, 1974).

works in language and logic undertaken by Frege and Russell. And, although not overtly internalist in nature, Wittgenstein's theory of language, in the *Tractatus*, is very similar to that espoused by Russell.[55] Language, for the early Wittgenstein, is pictorial—statements are meaningful as they picture possible states of affairs.

"A picture", for Wittgenstein, "is a model of reality."[56] In a picture, the objects of the picture correspond to the things that are being pictured. That is, "in a picture the elements of the picture are the representatives of objects."[57] Moreover, the objects of the picture are related to one another in the very same way as those things which they represent. In Wittgenstein's words, "the fact that the elements of a picture are related to one another in a determinate way represents that things are related to one another in the same way."[58] Language, for the early Wittgenstein, is like a picture as it represents states of affairs in precisely the same way as do pictures—it depicts them.

As a picture is a model of reality, so too is a proposition or a statement. Language consists in an assemblage of depictions. Wittgenstein writes:

> In order to understand the essential nature of a proposition, we should consider the hieroglyphic script, that depicts the facts that it describes. And alphabetic script developed out of it without losing what was essential to depiction.[59]

---

[55] I qualify the possibility of the *Tractatus* position being internalist because although Wittgenstein never states that terms refer to sense-data or that meanings involve pictures in the mind, the most natural way to make sense of his position would be to construe it as an internalist position.

[56] Wittgenstein, *Tractatus*, § 2.12.

[57] Wittgenstein, *Tractatus*, § 2.131.

[58] Wittgenstein, *Tractatus*, § 2.15.

[59] Wittgenstein, *Tractatus*, § 4.016.

and, similarly:

> One name stands for one thing, another for another thing, and they are combined with one another. In this way the whole group—like a tableau vivant—presents a state of affairs.[60]

The idea is this: a statement says what some possible state of affairs is by taking names (which stand in for the objects in the state of affairs) and using those names to show the form of the state of affairs—that is, the names are arranged, just as they are in the state of affairs. Consider, for example, 'the cat is on the mat'. On the early Wittgensteinian view, 'the cat' and 'the mat' stand in for the objects (i.e., the cat and the mat) in the state of affairs, while the use of 'on' shows the relation between the two objects just as a painting depicting a cat on a mat would. So, for Wittgenstein, "a proposition constructs a world with the help of a logical scaffolding, so that one can actually see from the proposition how everything stands."[61]

Pause to consider a brief rejoinder to Wittgenstein and this picture theory of language. The objection is this: some words and propositions are far too complex to be thought of as any sort of pictorial representation. Here Wittgenstein falls back on Russellian atomism and its notion of logical analysis. That is, *every* proposition will ultimately be analyzable into simples (i.e., names that stand for, and have as their meaning, the objects they denote) and the logical relations that hold between them. So, like Russell before him, Wittgenstein thinks, "it is obvious that the analysis of propositions must bring us to elementary propositions which consist of names in

---

[60] Wittgenstein, *Tractatus*, § 4.0311.

[61] Wittgenstein, *Tractatus*, § 4.023.

immediate combination."[62] So, one need not worry that propositions do not appear to be pictorial representations because these are meaningful only insofar as they are analyzable into elementary propositions—and these *do* picture possible states of affairs.

For the early Wittgenstein, problems in philosophy (whether they be ethical, metaphysical, etc.) only arise because the "logic of language is misunderstood."[63] Having completed Russell's work, with his picture theory demonstrating all that can be meaningfully said, Wittgenstein claimed:

> The truth of the thoughts that are here communicated seems to me
> unassailable and definitive. I therefore believe myself to have found,
> on all essential points, the final solution of the problems.[64]

Believing himself to have solved the problems of philosophy, Wittgenstein left the academy to become a schoolteacher in rural Austria. However, as is often the case, all was not what it seemed, and in 1929, after many years, Wittgenstein returned to philosophy in order to dismantle the Russellian (and in a large part the internalist) programme—a programme that one Ludwig Wittgenstein had brought to its culmination in the *Tractatus*.

Remarks on Naming

In the *Philosophical Investigations*,[65] Wittgenstein begins his attack on

---

[62] Wittgenstein, *Tractatus*, § 4.221.

[63] Wittgenstein, *Tractatus*, p. 3.

[64] Wittgenstein, *Tractatus*, p. 4.

[65] Ludwig Wittgenstein, *Philosophical Investigations*, trans. G. E. M. Anscombe, 3rd ed. (Englewood Cliffs, NJ: Prentice Hall, 1958).

internalist theories with an attempt to demonstrate that naming cannot be a primary or fundamental aspect of meaning. Recall that, for Locke, language consists in the application of names to ideas, which, in turn, represent objects in the world. Russell argues that the meanings of terms just are the objects that are named. And in modern variants of internalism, some mental act of intending or naming functions to determine meaning. One of the seemingly plausible aspects of internalism, then, is the role that naming plays in our linguistic practices. That is, language in its most basic function seems to involve the naming of objects—and part of the *prima facie* plausibility of internalism derives from the securing of this function. But it is this 'seeming plausibility' that Wittgenstein first attacks.

Wittgenstein concedes that naming is an integral part of our linguistic practices, and that many words name or denote objects. What he denies is that this offers any elucidation of the concept of meaning. Internalists, according to Wittgenstein, think that "we name things and then we can talk about them: can refer to them in talk."[66] This, for Wittgenstein, is an incomplete picture. He considers a speaker defining the number two, to an audience, by pointing to two nuts and uttering 'This is two'. This seems similar to the way we do, in fact, teach people new words, yet Wittgenstein thinks this can tell us nothing about meaning.[67] The problem is that the hearer may take 'two' to refer to nuts, or this group of nuts, or the arrangement

---

[66] Wittgenstein, *Investigations*, § 27.

[67] Wittgenstein, *Investigations*, § 28. Wittgenstein concedes that we often do teach people through ostensive definition, but as we shall see in what follows, this is only after we have some fairly sophisticated linguistic abilities.

between the nuts, etc. That is, the hearer may misunderstand what exactly is being pointed to. Someone may object, of course, that the speaker ought to define 'two' thusly: 'This number is two'. The use of 'number' is supposed to ensure that the hearer knows that the quantity of nuts is what the speaker is trying to define. But this is of no assistance because all of the previous considerations will now apply to the term 'number'. One seems either caught in a vicious circle or destined to offer definitions *ad infinitum*.

Yet matters are even worse than they appear. It is not only that the hearer can misunderstand or not determine what object is being defined (i.e., pointed to), but also the very act of pointing is indeterminate. As Wittgenstein points out, "where is it said which way I am to follow it; whether in the direction of the finger or (e.g.) in the opposite one?"[68] That is, nothing can determine, in advance, how one ought to take the act of pointing itself. Without a sophisticated enough background of practices (i.e., without already knowing how pointing works), pointing could be construed as going in the direction of the finger, in the direction of the shoulder, or even as a command to walk in that direction.[69] Naming, ostensively defining, pointing, etc. seem fraught with indeterminacy, and if there is a tight correlation between naming and meaning (as the internalist suggests), then meaning itself seems in peril.

Before drawing out the implications of these comments and presenting a

---

[68] Wittgenstein, *Investigations*, § 85.

[69] Now as a matter of fact, we do take pointing in very similar ways (i.e., we follow in the direction of the finger). But we must appeal here to shared natural or learned reactions, and the point is then that this is all background or stage-setting and cannot be used to explain our linguistic practices.

coherent criticism, pause to consider an (early perhaps) internalist reply. Internalists may grant that these behaviours and activities seem indeterminate, but they, in fact, are not because some mental act of naming underlies them. So, the story goes, a speaker means this, rather than that, because some psychological state or mental act of intending or correlation with some mental language *determines* that this, rather than that, is what the speaker means. In short, something mental determines meaning. So, the reason that naming, pointing, etc. are not indeterminate, as envisioned by Wittgenstein, is because they are grounded or determinately fixed by something mental. But this, thinks Wittgenstein, can be of no help to the internalist. First, all of the problems with naming will re-emerge at the level of the mental. For internalists, a speaker's understanding (of meaning) consists in knowing what names apply to which things—yet, as Wittgenstein demonstrates, this is problematic. But the internalist now wants to secure this by way of some mental language, or something of the sort. Wittgenstein characterizes this position as follows:

> [It] is as if the child came into some strange country and did not understand the language of that country; that is, as if it already had a language, only not this one. Or again: as if the child could already think, only not yet speak. And "think" would here mean something like "talk to itself".[70]

So, the internalist posits an internal language that fixes or determines our 'outer' linguistic practices in order to solve the naming problems, discussed above. That is, internalists explain how it is that our linguistic practices are meaningful by supposing that there is some mental language that is meaningful. But notice that the internalist

---

[70] Wittgenstein, *Investigations*, § 32.

reply is equally problematic because, now, an explanation is needed at the level of the mental. In other words, if naming, pointing, etc. could not ground our linguistic abilities, neither can they ground our mental abilities. The internalist now owes an explanation of how one understands oneself (i.e., in one's mental language). Is there a 'meta-mental language' that determines meanings in the mental language? If so, the internalist is off on a regress.

The second problem with this internalist reply is that, for Wittgenstein, as a matter of fact, a mental accompaniment cannot suffice to determine meaning. To get at the heart of the problem, consider Wittgenstein's primitive builders' language that consists only of names.[71] In this language, a speaker's utterance of 'Slab!', for example, is meant to get a hearer to bring the speaker a slab. Now, Wittgenstein says the following:

> If you shout "Slab!" you really mean: "Bring me a slab".—But how do you do this: how do you *mean that* while you *say* "Slab!"? Do you say the unshortened sentence to yourself? And why should I translate the call "Slab!" into a different expression in order to say what someone means by it? And if they mean the same thing—why should I not say: "When he says 'Slab!' he means 'Slab!'"? Again, if you can mean "Bring me the slab", why should you not be able to mean "Slab!"?— But when I call "Slab!", then what I want is, *that he should bring me a slab*!—Certainly, but does 'wanting this' consist in thinking in some form or other a different sentence from the one you utter?[72]

The preceding discussion made it seem as if naming, pointing, etc. left meaning, in a certain sense, indeterminate. The internalist's solution is to posit something mental

---

[71] Wittgenstein considers this language in various sections: see especially, Wittgenstein, *Investigations*, §§ 2, 6, and 8.

[72] Wittgenstein, *Investigations*, § 19.

(e.g., a mental act of intending or an association with a psychological state, etc.) to help determinately fix what a speaker means. Suppose, with the internalist, then, that a mental 'Slab' accompanies the builder's 'Slab!' and that, likewise, an English speaker's 'Bring me a slab' is accompanied by a mental 'Bring me a slab'. Wittgenstein's concern is this: nothing mental can be what fixes meaning because these two utterances mean the same thing (i.e., that the hearer should bring the speaker a slab) despite the differing mental accompaniments. That is, no matter what the mental accompaniment (i.e., no matter what psychological state the speaker is in at the time of the utterance) the utterance still means bring that slab. Why?—simply, because the utterances are used to get a hearer to bring the speaker a slab. This leads Wittgenstein to query, "Doesn't the fact that the sentences have the same sense [meaning] consist in their having the same use?"[73] The 'something mental' is not what leads to similarity of meaning; rather it is the fact that the utterances are used for the same purposes that yield similarity of meaning. So, the internalist cannot attempt to secure meaning by under-girding it with something mental, because that mental thing is not what meaning consists in. In short, this possible internalist rejoinder either removes all of the naming and pointing problems by pushing them one level back (i.e., into the mental), or it posits some superfluous fact which is of no help in determining meaning.

All of these considerations lead Wittgenstein to ultimately conclude that naming cannot suffice to explain our linguistic abilities. The internalist sought to

---

[73] Wittgenstein, *Investigations*, § 20.

45

explain our linguistic abilities in terms of naming—that is, understanding a language just is to know what names apply to which objects. But, for Wittgenstein, to say that a "word in a language signifies something" is, so far, to say "nothing whatever."[74] Wittgenstein may be construed, then, as making two related points in the preceding discussions: first, it does not suffice for understanding that one have knowledge of the names of objects (in which case, neither does it suffice for meaning); and second, naming (whether a mental intending or not) cannot be invoked to explain our linguistic practices and abilities.

Consider the former. For Wittgenstein, knowing what names stand for which objects is, at best, an incomplete account of understanding. Recall that naming and pointing left meaning, in certain respects, indeterminate. This, thinks Wittgenstein, is because one can know what names apply to which objects without understanding the name—that is, without knowing the meaning of the name. Imagine, for example, a group of people learning the builders' language discussed above. In this particular language, recall, utterances are used in order to get someone to bring the speaker the object that was named. 'Slab!', for example, is uttered so that the hearer will bring the speaker a slab. Suppose, now, that the individuals learning this language are taught by ostensive definition—they are shown a number of different slabs and 'Slab!' is uttered, until a time when the learners are capable of identifying slabs as slabs in accordance with the correct utterance. Wittgenstein's point is that, even if pointing and naming are determinate, this still does not constitute understanding of the word

---

[74] Wittgenstein, *Investigations*, § 13.

'Slab!' This does not suffice for understanding because there is a piece of knowledge that the learners still lack. That is, the learners do no yet know that when 'Slab!' is uttered they are meant to bring the speaker a slab. It is "only together with a particular training" (e.g., one in which the learners are taught that the utterance is a command to bring the object) that the learner has a complete understanding, and "with different training the same ostensive teaching of these words would have effected a quite different understanding."[75] So, for Wittgenstein, even with knowledge of the names of objects, the learners do not yet know what the words mean. Wittgenstein considers an analogous example:

> When one shews someone the king in chess and says: "This is the king", this does not tell him the use of the piece—unless he already knows the rules of the game up to this last point: the shape of the king.[76]

The point of this example is the same as the last: one can know the names of every piece on the chessboard, but without knowing how the pieces move, that they even are pieces in a game, etc. one actually knows very little. The same, thinks Wittgenstein, is true of language—"the shape of the chessman corresponds here to the sound or shape of a word."[77]

This leads directly to Wittgenstein's latter point—that naming cannot *explain* our linguistic abilities. The idea that naming does not suffice for understanding

---

[75] Wittgenstein, *Investigations*, § 6.

[76] Wittgenstein, *Investigations*, § 31.

[77] Wittgenstein, *Investigations*, § 31.

demonstrates that "naming is so far not a move in the language game."[78] But one ought not to conclude that Wittgenstein is sceptical about meaning, nor about naming or ostensive definition. On the contrary, Wittgenstein *does* believe that naming, pointing, etc. have a role to play in language. He states, clearly, "the ostensive definition explains the use—the meaning—of the word when the overall role of the word in the language is clear."[79] To appreciate how this is to be understood, consider another chess example:

> I am explaining chess to someone; and I begin by pointing to a chessman and saying: "This is the king; it can move like this,… and so on."—In this case we shall say: the words "This is the king" (or "This is called the 'king'") are a definition only if the learner already 'knows what a piece in a game is'. That is, if he has already played other games, or has watched other people playing 'and understood'—*and similar things*…We may say: only someone who already knows how to do something with it can ask significantly of a name.[80]

Wittgenstein's contention is not that naming plays no role in our linguistic practices, but that it can function correctly only after a significant number of other practices, techniques and training are in place. That is, naming can proceed successfully only after one is already "master of a language."[81]

What has this to do with internalism? This speaks directly to those internalists who claim that meaning is determined by some kind of mental naming (e.g., *this* name means *that* because that is how the speaker intended it). But naming, for

---

[78] Wittgenstein, *Investigations*, § 49.

[79] Wittgenstein, *Investigations*, § 30.

[80] Wittgenstein, *Investigations*, § 31.

[81] Wittgenstein, *Investigations*, § 33.

Wittgenstein. presupposes that one has some fairly sophisticated linguistic abilities. So. one cannot explain our linguistic practices and abilities in terms of naming or intending because they are part and parcel of those very same abilities and practices. That is. on pain of circularity. naming cannot be offered as an *explanation* for our linguistic abilities, as those abilities are presupposed by the act of naming (i.e.. one can name only after one is master of a language). The internalist, then, cannot use naming to explain anything precisely because it is part of what is in need of explanation.

A brief summary: Internalists (e.g.. Locke or Russell) tend to argue that the meaning of a term just is the object that the term denotes. Understanding a language, then, just is the knowledge of terms and their referents—to know the meaning of a term is nothing more than to know what object it denotes. But Wittgenstein demonstrates that this is not an adequate account of our linguistic practices. The acts of naming, pointing, and ostensively defining are fraught with indeterminacy; and even if they were not, to know what name applies to which object yields only a partial or incomplete understanding of a term. That is, even knowing a name and its referent, one may not yet be capable of using a term correctly. And. as mentioned earlier, still other internalists (e.g., Grice or Searle) claim that meaning is determined intentionally, perhaps, although this is somewhat to oversimplify, through some mental act of naming. But this, thinks Wittgenstein, explains nothing. For Wittgenstein. the ability to name or to intend to refer is an aspect *of* our linguistic ability, and as such is part of what is in need of explanation. And, as the ability to

name is part of our linguistic practices, the attempt to explain those practices in terms of naming explains nothing. So, naming, which is an integral aspect of the internalist view, is ill equipped, no matter what its guise, to aid in explaining the concept of meaning. However, to my mind, at least, Wittgenstein offers a more devastating critique of the internalist position in the form of the 'private language argument'.

The Private Diarist and Private Languages

Wittgensteinian interpretation is at the best of times a controversial matter. Present Wittgensteinian literature consists largely of commentators arguing for and defending their preferred interpretations against the interpretations of others. The so-called 'private language argument' is no exception. Some commentators believe that the private language argument consists of the entire first section of the *Investigations*, while others see it starting at remark 243 and running through remarks in the early 300's.[82] What seems agreed upon, however, is that the private language argument consists of a number of remarks that are meant to wean us away from the intuitively plausible idea that sensation terms (e.g., 'pain', 'fear', etc.) are in some important way private—that is, they have private meanings. I think, however, that, no matter which interpretation is correct, the cornerstone of this argument is the example of the private diarist.

---

[82] An example of the former may be found in Hilary Putnam, "Sense, Nonsense, and the Senses" *Journal of Philosophy* 91 (September 1994): 445-517; reprinted in his *The Threefold Chord: Mind, Body, and World* (New York: Columbia UP, 1999). References are to the latter: see p 25. For an example of the latter interpretation see Jonathan Dancy, "Foundationalism and Other Minds," *Introduction to Contemporary Epistemology* (Oxford: Blackwell, 1985), p. 76. Saul Kripke, on the other hand, claims that the standard view is that the private language argument begins at remark 243, but he thinks that it is in the sections preceding 243: see Saul Kripke, *Wittgenstein on Rules and Private Language: An Elementary Exposition* (Cambridge, Mass: Harvard UP, 1982), pp 2-5.

Before undertaking an examination of the private diarist example, one needs to be clear about what Wittgenstein means by 'a private language'. The private language is not simply a language that one person does, in fact, speak. Rather, it is a language that, in principle, only one person can speak. Wittgenstein describes it thusly:

> The individual words of this language are to refer to what can only be known to the person speaking; to his immediate private sensations. So, another person cannot understand the language.[83]

So, the private language is thought to include the following: first, the words of the language refer to objects (i.e., private sensations) that are available, accessible, etc. only to the speaker who has them; second, the speaker can know that the words refer to these objects—that is, the speaker can understand the language; and third, no audience can know what the meanings of the words in this language are because no one else has access to the sensations to which the words refer.[84] It is this notion of a private language that the private language argument (and the private diarist example in particular) attempts to counter. Having attended to this preliminary detail, we may return to exegetical matters.

Since there is debate about the appropriate interpretation of the private diarist example, it is best to first consider the example as Wittgenstein presents it:

> Let us imagine the following case. I want to keep a diary about the recurrence of a certain sensation. To this end I associate it with the sign "S" and write the sign in the calendar for every day on which I have this sensation.—I will remark first of all that a definition of the sign cannot be formulated.—But still I can give myself a kind of

---

[83] Wittgenstein, *Investigations*, § 243.

[84] One ought to be reminded here of both Locke and Russell's picture of language—especially as it relates to ideas and sense-data.

ostensive definition.—How? Can I point to the sensation? Not in the ordinary sense. But I speak, or write the sign down, and at the same time I concentrate my attention on the sensation—and so, as it were, point to it inwardly.—But what is this ceremony for? For that is all it seems to be! A definition surely serves to establish the meaning of a sign.—Well, this is done precisely by the concentrating of my attention; for in this way I impress on myself the connexion between the sign and the sensation.—But "I impress it on myself" can only mean: this process brings it about that I remember the connexion *right* in the future. But in the present case I have no criterion of correctness. One would like to say: whatever is going to seem right to me is right. And that only means that here we can't talk about 'right'.[85]

Although the argumentation is rather elegant and subtle, the example itself is relatively straightforward. The private diarist wants to record the occurrences of a sensation, *S*. To that end, the diarist decides to call the sensation '*S*' and will write an '*S*' on the calendar for every occurrence of the sensation. Now, since the sensation is essentially private, the diarist can ostensively define the term only through a kind of inward pointing—an inner ceremony whereby the diarist says, "I hereby name this sensation '*S*'", or something of the sort. The diarist must mentally 'impress the connexion' between word and sensation, and perhaps in this way give the word meaning. Moreover, for '*S*' to have meaning and be understandable to the diarist, it is required that the connection between *S* and '*S*' (i.e., between sensation and word) be remembered correctly—that is, that the diarist uses '*S*' correctly in the future. Wittgenstein's point is that, in the case of the private diarist, there is no criterion of correctness (or wrongness, for that matter). So, '*S*' *cannot* be meaningful.

Since the private diarist example relies on remarks both preceding and following it, I think it best, at this point, to sort out the various pieces of the puzzle

---

[85] Wittgenstein, *Investigations*, § 258.

and present it in something resembling an argument.[86] Before beginning this reconstruction, however, it is important to note that Wittgenstein already has an argument against such a private language. Recall, in the discussion on naming, that Wittgenstein concluded that the meaning of a term was not simply its referent. This was because speakers can know what names apply to which things, yet still not understand the names because they did not know how to go on to use the word (e.g., the builders' language and the chess example). Similar considerations could apply here. Contrary to the definition of a private language, speakers do not understand the words of a private language precisely because it does not suffice for understanding that speakers know the referents of their terms. So, as a matter of fact, speakers cannot understand such a private language. But Wittgenstein is after bigger game in this example. The private diarist example is meant to demonstrate, not only that such a private language does not occur, as a matter of fact, but also that such a private language is logically impossible.

From earlier discussion, we know that simply attaching the name '*S*' to the sensation is not enough to make '*S*' meaningful. For the diarist to know that '*S*' refers to *this* sensation is not enough to constitute understanding of '*S*'. Now, immediately following the diarist example Wittgenstein asks, "Are the rules of the private language *impressions* of rules?"[87] This suggests a connection between meaning and rules. But what is this connection? For Wittgenstein, the diarist's term '*S*' can be meaningful

---

[86] For an analysis similar to the one I present, although with less emphasis on rule following, see P.M.S. Hacker, *Wittgenstein: Meaning and Mind*, Part II (Oxford: Blackwell, 1993), pp. 62-4.

[87] Wittgenstein, *Investigations*, § 259.

only if rule-governed. That is, 'S' is meaningful only if the diarist's application of 'S' has some regular pattern of use. Suppose, then, that the diarist's rule is this: when I have this particular S-sensation, I will write 'S' down on my calendar.[88]

So far the diarist has noted some particular sensation, S, named it 'S', and created a rule to govern the use of the word. Next Wittgenstein notes that if 'S' is to be rule-governed (and, hence, meaningful) there must be a correct and incorrect way to apply the rule. This is why to say that "I impress it on myself [i.e., the connection between word and sensation]" means "the process brings it about that I remember the connection *right* in the future."[89] The point is that unless there is a right and a wrong way to apply the rule, there really is no rule. (Consider whether it is really a rule if one could not fail to follow it.)

The proverbial nail in the coffin comes when Wittgenstein states "but in the present case I have no criterion of correctness. One would like to say: whatever is going to seem right to me is right."[90] The point is that for the private diarist there is no real distinction between actual rule following and seeming to follow a rule. There is no distinction here because for the private diarist these amount to the same thing. To see why, consider the difference between 'how things are' and 'how things seem'.

---

[88] Wittgenstein can already conclude that 'S' is not meaningful because a stipulated rule, to give meaning, is just as indeterminate as naming: see Wittgenstein, *Investigations*, §§ 81-7. For Wittgenstein, meaning requires that the usage of a term is rule-governed; *and* the usage of the term 'seeps back' to retroactively create a rule (more to follow on this in the section on 'meaning is use'): see Wittgenstein, *Investigations*, §§ 189-91. So, since 'S', as of yet, has no usage, there is not yet a rule, and, therefore, 'S' is not meaningful. But, to not beg the question and to be as charitable as possible to the internalist, I suppose that the diarist can create this rule.

[89] Wittgenstein, *Investigations*, § 258.

[90] Wittgenstein, *Investigations*, § 258.

First. note that 'how things seem' means something like 'how things are from a particular point of view' (to say. for example. 'that seems sour to me' just is to say 'from my point of view that is sour'). Now suppose that certain restrictions are put in place such that the distinction between how things are and how things seem is being considered from a singular point of view—suppose this is the only possible view available. Under these conditions 'how things are' will also mean 'how things are from a particular point of view' precisely because it is the only possible view available. That is, the particular point of view determines 'how things are' because the particular point of view is the only one such that 'how things are' and 'how things seem' coincide. In other words, considered from the singular point of view the distinction between them collapses.

The final piece of the puzzle is to notice that the private diarist is in precisely this situation. Recall that the diarist's rule was 'write 'S' in the calendar whenever I have sensation S'. But if this is an actual rule, there must be a right and wrong or a correct and incorrect application of the rule. That is, there must be a 'how things actually are/how things seem' distinction, if there is a genuinely correct and incorrect application of the rule. For the diarist's term to be rule governed is for there to be a distinction between the diarist's actually having sensation S (how things are) and the diarist's merely thinking that some new sensation is S (how things seem). But by the definition of the private language the diarist's point of view is the only one available. And this means that the distinction needed to sustain the diarist's rule collapses. To actually have the sensation and to think one has the sensation amount to the same

thing from the diarist's singular point of view. So, there is no criterion of correctness because all the diarist can do is compare 'how things seem' (i.e. 'how things are from one particular point of view') with 'how things seem'. This leads Wittgenstein to conclude "whatever is going to seem right to me is right."[91]

Now, it may be objected that the temporal element in the case of the diarist avoids these problems—that is, the diarist's point of view at different times is supposed to underwrite the 'how things are/how things seem' distinction.[92] But if we consider the point, since it is such an important one, put slightly differently, I think this problem can be avoided altogether. Suppose the diarist claims, "my use of the term '*S*' is rule-governed." What this means, considered *qua* part of a private language, is "from my point of view the use is rule governed." But all this is just to say, "it seems to me that I am following a rule" (recall that 'how things seem' cashes out as how things are from a particular point of view). So, when the appropriate private language restrictions are in place there is no distinction between genuinely following and seeming to follow a rule—and this, notice, needs no reference to time. That is. "here we can't talk about 'right'."[93]

---

[91] Wittgenstein, *Investigations*, § 258.

[92] Wittgenstein, perhaps in anticipation of such an objection, argues that this still does not constitute an objective check. It is "as if someone were to buy several copies of the morning paper to assure himself that what it said was true."(§ 265) The point is that this sensation at a later time cannot be invoked to underwrite the 'how things are/how things seem' distinction because everything that could be done in comparing the new sensation to the original has already contributed to the thought that the new sensation is like the old. So, the diarist cannot objectively check if the sensation, at this later time, is like the original because anything that the diarist could do as a check is a repetition of whatever made the diarist think it was the same sensation in the first place—and this constitutes no check at all.

[93] Wittgenstein, *Investigations*, § 258.

For the diarist, the only determination of whether 'S' is being used correctly is that it seems that (i.e. the diarist thinks that) it is being used correctly. In other words, the diarist cannot fail to use the word correctly because thinking the word is being used correctly (or similarly thinking that it is the same sensation) is all that is available. And Wittgenstein remarks, "to think one is obeying a rule is not to obey a rule."[94] The private diarist, then, is not obeying a rule. So, in this private language, there is no correct usage of 'S' because a genuine rule involves both right and wrong application—of which there is none. That is, 'S' is not rule-governed and therefore has no meaning. And, since 'S' fails to be meaningful, Wittgenstein concludes that the type of private language under consideration is impossible.

I summarize Wittgenstein's argument in the private diarist example, then, as follows:

(1) The private language term 'S' is meaningful only if its usage is rule-governed.
(2) The usage of 'S' is rule governed only if there is a genuine distinction between actual rule-following and only seeming to follow a rule.
(3) For the private diarist (because of the private language restrictions) there is no such distinction.
(4) Therefore, the usage of 'S' is not rule-governed (from (3) and (2)).
(5) Therefore, the private language term 'S' is not meaningful (from (4) and (1)).

Moreover, this will hold for any private language term, so long as the private language is such that is restricts the distinction between actual rule-following and seeming to follow a rule to one private point of view. That is:

(6) A private language is not possible.

---

[94] Wittgenstein, *Investigations*, § 202.

Having concluded Wittgenstein's private language argument, what ought one now say about internalist theories of meaning? They too are not possible. Recall that for Locke ideas are 'hidden in our breasts'. Similarly, for Russell, sense-data are objects that only the individual perceiving them can be acquainted with. Moreover, these objects are supposed to give meaning to the terms that refer to them. So, internalism cannot be right precisely because these objects are private, just like the sensation S. Take any term one likes under the Lockean or Russellian account of language, and substitute it for 'S' in the above argument. Since the referent of such a term is a private object, it follows that there is no distinction to be drawn between the correct and incorrect usage of the term (i.e., whenever Locke or Russell think they are using the term correctly they must be using it correctly). Suppose that 'cat' refers to a Lockean cat-idea. Since the cat-idea is private, all that regulates the usage is one's thinking that the usage is correct. In other words, there is not really an incorrect usage of 'cat'. Similar considerations will apply to the terms that correspond to Russell's sense-data. As the diarist example demonstrates, any attempt to establish meaning by naming a private object *must* fail. Therefore, the terms of Locke's and Russell's internalist accounts cannot be meaningful. And yet it is precisely their meaningfulness that internalism sought to explain.

Meaning is Use

Thus far, we have seen Wittgenstein attack internalism on two distinct fronts. In the discussion on naming, we saw that something mental could not determinately fix meaning as this presupposes precisely what is in need of explanation, and in the

diarist example, we saw that names referring to private mental objects cannot be meaningful. Yet the internalist still has room to manoeuvre here. All Wittgenstein has done, the internalist may respond, is argue for a certain scepticism about meaning. But, since communication does occur—we do take our words to be meaningful—internalism is still the best theory available. It is the only theory available, and so the internalist may lay claim to being the proverbial 'only kid on the block'. But this simply misses much of the force of Wittgenstein's argument. Throughout *all* of the preceding discussions, while methodically attacking various versions of the internalist programme, Wittgenstein has simultaneously been constructing a positive account of language. This account is the doctrine that 'meaning is use'. The idea that the use of a term determines its meaning has been implicit through the sections above, and we need now merely make it explicit.

For Wittgenstein "philosophy may in no way interfere with the actual use of language; it can in the end only describe it."[95] The description that the internalist offers is found wanting. Internalists claim, by and large, that knowing what words mean consists in knowing the names of objects—language 'hooks onto the world' by way of naming, pointing and ostensive definitions. Language involves the alignment of particular mental states with specific terms. But Wittgenstein demonstrates that this is not enough for understanding (nor for meaning). A speaker can know the name of an object, yet not understand—the speaker does not yet have enough knowledge to use the word correctly. That is, nothing in the act of naming itself makes it the case

---

[95] Wittgenstein, *Investigations*, § 124.

that the word is used one way rather than another. So, Wittgenstein concludes:

> For a *large* class of cases—though not for all—in which we employ the word "meaning" it can be defined thus: the meaning of a word is its use in the language.[96]

Consider again the primitive builders' language in which 'Slab!' means 'bring the speaker a slab'. The point of the discussion was to demonstrate that if speakers know only the names of objects, they do not yet know the meanings of these utterances. This is because these speakers do not yet know that 'Slab!' is used as a command to get someone to bring a slab. It is the particular *use* of 'Slab!' that makes it mean what it does. Imagine, now, a similar scenario, in which learners are taught ostensively, but they go on to use 'Slab!' only as an inquiry into whether or not some object is a slab. In this new scenario, despite identical naming procedures, one would like to say that 'Slab!' means 'is this a slab?' What does the work, in this case, is the use—it is the use that the speakers are making of the term that confers meaning upon the term. For similar reasons, the builder's 'Slab!' and the English 'Bring me a slab' mean the same thing. The name of the object or some underlying mental state is not what make the utterances mean the same thing—they cancel out as irrelevant. It is the usage that the builders and English speakers make of the terms that bestows them with the same meaning.

This idea that use determines meaning was also at work in the private diarist discussion, particularly with respect to rule-following. For Wittgenstein, saying that the use of a term is rule-governed is not to say that meaning is determined *by* a rule, as

---

[96] Wittgenstein. *Investigations*, § 43.

if "the rule... traces the lines along which it is to be followed throughout the whole of space."[97] Rather than the rule determining how one continues to use a term, the rule is determined by one's continuous usage of a term. Wittgenstein considers some new mathematical formula '$x!2$'. Next he claims that "if by '$x!2$' you mean x^2, then you get *this* value for y, if you mean $2x$, *that* one... now ask yourself: how does one *mean* the one thing or the other by '$x!2$'?"[98] The point here is that the use someone makes of this formula will determine what it means. If a person uses '$x!2$' to generate $\{1, 4, 9, 16...\}$ then it means x^2, but if it is used to generate $\{2, 4, 6, 8...\}$, then it means $2x$. So, when Wittgenstein asks, "what is the criterion for the way the formula is meant?" the answer is "it is, for example, the kind of way we always use it, the way we are taught to use it."[99] Some predetermined rule or abstract notion of meaning does not, and cannot, dictate the future uses of words; rather the uses one makes of words retroactively 'seep back' to determine meaning. Now, when Wittgenstein states that speakers "obey the rule blindly,"[100] he is simply noting that we do not properly see the rule, since we are, in effect, creating it as we go. So, in the private diarist example, when Wittgenstein claims that for a word to be meaningful it must be rule-governed, what he means is that the usage of a term will determine what the correct and incorrect applications are.

Consider one final example that clearly demonstrates how, for Wittgenstein,

---

[97] Wittgenstein, *Investigations*, § 219.

[98] Wittgenstein, *Investigations*, § 190.

[99] Wittgenstein, *Investigations*, § 190.

[100] Wittgenstein, *Investigations*, § 219.

the meaning of a term is determined by its use. In the course of the private language

argument Wittgenstein discusses a beetle in a box:

> Suppose everyone had a box with something in it: we call it a "beetle".
> No one can look into anyone else's box, and everyone says he knows
> what a beetle is only by looking at *his* beetle.—Here it would be quite
> possible for everyone to have something different in his box.[101]

Now, *qua* part of the private language argument, the point of the example is that the

private beetle is irrelevant to the meaning of the term 'beetle':

> The thing in the box has no place in the language-game at all; not even
> as a *something*: for the box might even be empty.—No, one can 'divide
> through' by the thing in the box; it cancels out, whatever it is.[102]

Again, Wittgenstein is not advocating scepticism about meaning. It is not as if the

word has no meaning; it is only that the private object is not a relevant feature of that

meaning. But Wittgenstein's positive spin here is that 'beetle' does have a meaning.

It is meaningful precisely because we use it in a particular way. That is, 'beetle'

means 'whatever is in the speaker's box', or something of the sort, because that is

how the word is used. In short, Wittgenstein not only criticizes the internalist

position, but also offers a positive account of language whereby the meanings of

various terms are determined by ways in which they are employed by a speaker—that

is, by their use.

Before discussing the continuity between this position and Putnam's, it is

necessary to establish that the Wittgensteinian account is an externalist one.

Externalism is defined, in part, as a rejection of internalism. That is, in a certain

---

[101] Wittgenstein, *Investigations*, § 293.

[102] Wittgenstein, *Investigations*, § 293.

sense, externalism just is the rejection of the idea that meaning is determined by the psychological states of a speaker. Both the naming discussion and the private diarist example clearly demonstrate Wittgenstein's externalism, in this negative sense. But, recall from the introduction that externalism also offers a positive line of argumentation. If meaning is not determined by the psychological states of a speaker, then what it is determined by is a speaker's various transactions and interactions with the social and physical environment in which the speaker is situated. Does Wittgenstein satisfy this aspect of externalism? His idea that use determines meaning, I take it, makes this very same point. 'Slab!' means 'bring me a slab' because of the manner in which it is used to interact with slabs and others in the environment. 'Beetle', similarly, means what it does because of the way it is used in talking to other people. So, to say that meaning is determined by use just is to say that meaning cannot be determined outside of the various contexts, circumstances and practices in which a term is used—that is, outside of speakers' interactions with their environment.

But is the view that use determines meaning necessarily externalist? Here commentators disagree. Some read Wittgenstein as advocating a kind of 'community view of language' or a 'social convention theory of meaning'. Barry Stroud describes this view as the identification of 'meaning is use' with the idea that "the correctness of a person's application of an expression can be measured only by its conformity to the

general practices of the community."[103] Read thusly, we are to understand Wittgenstein as arguing that the meaning of a term depends on its use within a linguistic community—where 'use' is nothing more than the term's standard usage amongst community members. A speaker uses a term correctly when that use conforms to the term's conventional use within a particular linguistic community. Saul Kripke, for example, advocates such a view.[104] Kripke argues that Wittgenstein must be read this way because it is only by introducing a community of language users that one's utterances can be meaningful. That is, only once there is an established linguistic community (and a term is regularly used amongst them) can there be something like a correct and incorrect use of a term. So, for Kripke, 'meaning is use' is read in this social conventionalist way because an appeal to the linguistic community is necessary to underwrite the 'how things are/how things seem' distinction that was problematic in the case of the private diarist.

But, this, I suggest, can only be half right. Wittgenstein does indeed suggest that one needs reference to a linguistic community in order for one's utterances to be meaningful. (I concede as much in describing Wittgenstein's 'use' as involving interactions with one's social environment.) That is, other speakers are needed to underwrite the 'actually are/only seem' distinction. But to claim that correctness or incorrectness depends on conformity to a linguistic community is, I suggest, un-

---

[103] Barry Stroud, "Wittgenstein on Meaning, Understanding, and Community," *Wittgenstein— Towards a Re-Evaluation: Proceedings of the 14th International Wittgenstein-Symposium*, eds. Rudolf Haller and Johannes Brandl (Vienna: Verlag Hölder-Pichler-Tempsky, 1990), p. 34.

[104] Kripke, pp. 96-8.

Wittgensteinian. Alternatively. I think that any construal of 'use' that does not make

explicit reference to a speaker's physical environment. the contexts in which

utterances are made. etc. (i.e.. that does not view 'use determines meaning' as

externalist) cannot be an adequate interpretation of the later Wittgenstein.

If Wittgenstein's transitional period (the period in between the *Tractatus* and

the *Investigations*) is any indication, then I think Wittgenstein must be read as an

externalist. In *The Blue Book* Wittgenstein writes:

> Someone says, "Mr. N. will come to see me this afternoon"; I ask "Do
> you mean him?" pointing to someone present, and he answers "Yes".
> In this conversation a connection was established between the word
> "Mr. N." and Mr. N. But we are tempted to think that while my friend
> said, "Mr. N. will come to see me", and meant what he said, his mind
> must have made the connection.[105]

Here Wittgenstein mildly chastises us for thinking that 'meaning something' is a

problem—the puzzlement arises only because we look to the mind as the arbiter of

meaning. But Wittgenstein's answer is clear. The physical environment, the person

himself, has a role to play in determining meaning. To put it as Putnam might, the

fact that 'Mr. N.' means Mr. N. has everything to do with the fact that 'Mr. N.' is used

to refer to Mr. N. If 'Mr. N.' were not used to refer to *him* (i.e., Mr. N.), then it would

mean something else. That is, Mr. N. himself is a determinant of the meaning of 'Mr.

N.'

This externalist strand of thought continues into Wittgenstein's *Remarks on*

the Philosophy of Psychology, where he writes:

---

[105] Ludwig Wittgenstein, *The Blue and the Brown Books: Preliminary Studies for the
'Philosophical 'Investigations'* (New York: Harper and Row, 1965), p. 39.

> A child does not *have* to first use a primitive expression which we then replace with the usual one. Why shouldn't he immediately use the adult expression which he has heard several times? It really doesn't matter how he "guesses" that this is the right expression, or how he comes to use it. The main thing is that no matter what the preliminaries are, he uses the word the same way adults do: i.e., on the same occasions, in the same context.[106]

Here Wittgenstein appeals to contexts and occasions of utterances as factors that determine meaning. One should not be confused by the reference to using terms the way adults do. This may be how a child *learns* the expressions that Wittgenstein is concerned with, but using the expression as the adult does, does not play a role in determining meaning—the context and occasions do that. If the adults used the expressions in different contexts and on different occasions, then the child may still come to use them as the adult does; but the expressions will now have different meanings precisely because they are used in different circumstances from before. The child's meaning the same thing as adults do with an expression (i.e., following the conventional use) does not determine meaning—the meanings of *both* of their expressions are determined by the occasions and contexts of the utterances. So, for Wittgenstein, a term's usage in contexts—while interacting with one's total environment (both social and physical)—is what determines meaning.

This idea—that one's surroundings, the environment, and particular contexts are determining factors of meaning—is one that Wittgenstein never abandoned. In the *Investigations* he maintains that if we take a word, term, expression, etc. and cut it "out of its context" or "out of its surroundings" it is "no longer part of this

---

[106] Ludwig Wittgenstein, *Remarks on the Philosophy of Psychology*, eds. G. H. Von Wright and Heikki Nyman, trans. C.G. Luckhardt and M.A.E. Aue, vol. II (Chicago: Chicago UP, 1980), § 241.

language."[107] This suggests, to my mind at least, that the environment in which a speaker is situated is simply *too* important for Wittgenstein to warrant an omission in any interpretation of 'use determines meaning'.

But, moreover, and I think more importantly, if we consider the role that 'use determines meaning' plays in the context of the *Investigations*, then Wittgenstein *must* be considered an externalist. Kripke is right in thinking that use determines meaning is supposed to account for the meaningfulness of our terms, in light of private language considerations.[108] Again, Wittgenstein is not sceptical about meaning—he is just sceptical about meaning occurring *this* way (i.e., involving private mental objects and the like). Use determines meaning, then, is intended to offer a way out of the fly-bottle—it shows how our utterances can be (or *are*) meaningful. That is, 'meaning is use' points to what went wrong in the case of the private diarist; and, as an account of language, is meant to avoid the meaning scepticism that the private language argument, and the private diarist example in particular, is supposed to produce.[109] But if the view that use determines meaning is interpreted as Kripke (or other 'community interpreters') would have it, this offers only a partial solution.

---

[107] Wittgenstein, *Investigations*, § 584.

[108] In what follows, I do not intend sayings such as 'in light of the private language considerations' or 'to avoid private language results', etc. to imply that Wittgenstein's 'use determines meaning' is offered to allow private languages (i.e., as if *now*—with 'meaning is use'—one can speak a private language). Rather, I use these locutions to mean that 'use determines meaning' allows for the meaningfulness of our terms, in a way that avoids what went wrong—and what led to the conclusion that 'S' could not be meaningful—in the case of the private diarist (i.e., meaningfulness cannot occur if construed privately, *but* it can occur this way—with 'meaning is use').

[109] That is, if internalism is the correct account of language, then our utterances cannot be meaningful (recall the private diarist case using Lockean ideas); and the view that use determines meaning is offered as an account of language that can avoid this meaning scepticism.

Recall that the problem in the case of the private diarist is that the private language restrictions make it so that the distinctions between having a sensation and thinking one is having a sensation. actually following a rule and thinking one is rule-following. etc. are all being considered from a solitary point of view. The reason the private diarist's term '*S*' cannot be meaningful, I suggested. is because these distinctions, which are needed to sustain meaning. collapse from a single viewpoint. Now the 'community reading' of Wittgenstein wants to simply reinstate the distinctions. That is, since others are needed to maintain the how things are/how things seem distinction, 'use determines meaning' is read as the way a term is used by a linguistic community (i.e., in one's interactions with the social environment) in order to ensure that other speakers are available to constitute an objective check on correct and incorrect uses of terms. So, for Kripke, reintroducing other speakers into the mix reintroduces the correct and incorrect uses need to underwrite meaningfulness. But, as I alluded to above, this can only be a partial solution. This only solves half of the problem because none of these speakers has access to the supposed object of the diarist's term—namely, the sensation *S*.

Kripke reads Wittgenstein as offering "a picture of language based… on assertability conditions."[110] An utterance will be warrantedly assertible, for Kripke, only if the linguistic community accepts the utterance. But a necessary component of this reading is that rule-following, understanding, meaning, etc. depend on community agreement. A speaker's correct understanding requires that "he agrees in his

---

[110] Kripke. p. 74.

responses with the responses produced by the members of *that* community."[111] And, since meaning and understanding are correlative, meaning itself depends upon community agreement. That is, it is only with the introduction of a linguistic community that utterances are meaningful, as the community is necessary for genuine rule-following and correct and incorrect usage.[112] Returning to the problem at hand, this can only be a partial solution because there can be as many speakers in the linguistic community, or as much community agreement, as one wishes, yet without those speakers having access to the object that is referred to by '*S*', all of the private language conclusions still follow. Suppose that the diarist wanders around uttering '*S*' in conjunction with *S* experiences. Can only community agreement bestow meaning on the diarist's utterance?[113] (Suppose, for example, the community perceives an unrelated regularity in the diarist's use that is a purely coincidental.) The answer must be 'no', because the *real* problem, in the case of the diarist, is that since the object *S* under consideration is essentially private (by the definition of a private language), the other speakers do not have access to it, and so can never help underwrite the 'actually are/only seems' distinction. That is, no matter how many other speakers are introduced, since *S* is private it is *always* considered from a solitary viewpoint; and the sceptical conclusions of the private language argument follow. In other words, the object itself must have a role to play in determining meaning. So,

---

[111] Kripke, p. 96.

[112] Kripke, pp. 88-9.

[113] I qualify with 'only' community agreement because Kripke rejects Wittgenstein's contention that 'inner processes require outward criteria' (which I take to be, at the very least, suggestive of an externalist reading) as a probably false empirical claim: see Kripke, p. 103.

insofar as the view that use determines meaning is introduced as a way to avoid the meaning scepticism engendered by the private language argument, 'meaning is use' must be construed as an externalist position. For Wittgenstein, then, the use of a term determines its meaning where use involves *both* a speaker's physical *and* social environment.

Consider an example which should help to clarify the differences between these two readings. Imagine that Mary is a typical English speaker. Mary speaks just like the rest of the English community with one exception. Whereas we walk around and say things such as "I would like a cup of water", "The water is cold", "Water would really quench my thirst", Mary says instead "I would like a cup of cat", "The cat is cold", "Cat would really quench my thirst". That is, Mary uses the term 'cat' while interacting with water. Moreover, suppose that, by some strange quirk, water produces in Mary something like a Lockean cat-idea. Now if we take seriously the idea that use determines meaning, we must say that in Mary's idiolect 'cat' means water. And if even some of the preceding comments are right, then Wittgenstein too *must* say that Mary means water by 'cat'. This is because, for Wittgenstein, different surroundings, circumstances, contexts, and interactions with the environment produce different meanings. If someone uses the term 'cat' in different contexts and in different interactions with the environment, for example, as we do in interactions with cats, then the word will mean cat. But since Mary uses it in interactions with all and only water, her term 'cat' means water.

Notice that in the Mary example there is no need for conventional usage or

standard meanings. Mary is the only person to use the term 'cat' this way, yet it is perfectly obvious what she means. So, the standard or conventional use by a linguistic community of 'cat' is irrelevant to the meaning of her term. Mary's term means water because of the contribution of the environment—because she uses 'cat' while interacting with water. All that is required for Mary's term to be meaningful is that *her* use exhibit a regular pattern—not that it match the pattern of use of her linguistic community. This is not to say, however, that the other speakers are irrelevant to the meaning of Mary's term. Without any other speakers, Mary would not be able to sustain the how things are/how things seem distinction needed for her term to be meaningful—that is, she would be for all practical purposes in a private language situation. So, the linguistic community *is* needed for Mary's term to be meaningful. But notice that the linguistic community need not ensure that Mary's use of 'cat' conforms to the conventional usage of the community. All that the community is needed for is to ensure that Mary uses the term correctly relative to *her* use. That is, the community is needed only to ensure that there is a distinction between Mary using the term correctly and thinking she is using it correctly (i.e., to make sure, for example, that she applies it only to water). The role of other speakers, for Wittgenstein, then, is not to make certain that Mary's term is rule-governed as is their own (i.e., that she uses it like them), but only to make certain that it *is* rule-governed.[114] So, the meaning of Mary's term is determined by her interactions with

---

[114] Davidson's way of putting the point is that without others, one cannot individuate the terminus of one's own perceptual causal chain. The other speakers, then, (with their perceptual causal chains) help the individual to triangulate upon a common stimulus, such that what a word refers to can now be determinately fixed. Meaning then must have recourse to both the physical environment (or else the

her environment and the linguistic community ensures that it is meaningful.

I think after all this that I can safely conclude that Wittgenstein's position is an externalist one. (And, moreover, we have had occasion to further clarify what is meant by 'use determines meaning' along the way.) For Wittgenstein, the meaning of a term is determined by its use. The meaning of an utterance, for Wittgenstein, depends on the contexts in which it is used, on one's surroundings when using the utterance, etc. In short, the use of a term 'cashes out' as the use of a term while interacting with one's social and physical environment. This just is externalism; and as we shall see in the next chapter, this is precisely the account of language that Putnam offers as well.

---

causal chain goes inward and others cannot help fix it) and the social environment (or else the speaker cannot determine what object along the causal chain is the stimulus, in which case the distinctions would collapse again): see Donald Davidson, "Three Varieties of Knowledge," *A. J. Ayer: Memorial Essays.* Royal Institute of Philosophy Supplement 30, ed. A. Phillips Griffiths (Cambridge: Cambridge UP, 1991), pp 157-9.

3

Putnam

In the preceding chapter I discussed three particular areas in Wittgenstein's *Investigations* that I take to be overtly externalist. His remarks on naming suggest that a mental intention cannot be invoked to explain our linguistic practices because this presupposes the very ability that is in need of explanation. The private language argument demonstrates that an essentially private object (i.e., a 'hidden' psychological state) cannot 'enter into' meaning. And finally, I suggested that the idea that the use of a term determines its meaning is externalist because the usage of a term amounts to interactions with one's social and physical environment.

The purpose of the present chapter, and of my thesis in general, is to demonstrate that Putnam's semantic externalism is a continuation of the work begun by Wittgenstein. To that end, I undertake four tasks in this chapter. First, I examine Putnam's 'Twin Earth' thought-experiment and point to its affinities with the private language argument. Second, I demonstrate that Putnam's discussions on 'intentionality' are a continuation of Wittgenstein's remarks on naming. Third, I discuss semantic externalism in detail and emphasize points of similarity between it and Wittgenstein's view that use determines meaning. Finally, I take up a number of objections, raised by Rudd, to the effect that Putnamian and Wittgensteinian externalisms are distinct theories and I suggest that these criticisms depend on misreadings of both Wittgenstein and Putnam.

Twin Earth

It is in "The Meaning of 'Meaning'"[115] where Putnam introduces his now famous Twin Earth thought-experiment. The purpose of this example is, of course, to demonstrate that the mental state of a speaker does not determine the reference or meaning of a term. Lest we get ahead of ourselves, however, we ought to begin at the beginning.

Putnam starts by identifying what he believes are two unchallenged assumptions of traditional theories of meaning. He writes:

(I)   That knowing the meaning of a term is just a matter of being in a psychological state…

(II)   That the meaning of a term (in the sense of 'intension') determines the extension (in the sense that sameness of intension entails sameness of extension).[116]

These two assumptions, Putnam contends, are not jointly satisfiable, and, in particular, an entailment from (I) and (II) is false.

We begin the journey to Twin Earth with this entailment. According to the traditional theory, by assumption (I), understanding consists of being in a particular mental state such that knowing the meaning of a term, A, is a different mental state from knowing the meaning of a term, B. Similarly, knowing that M_1 is the meaning of A is a different mental state from knowing that M_2 is the meaning of A. That is, according to the traditional view, "if Oscar and Elmer understand a word A

---

[115] Hilary Putnam, "The Meaning of 'Meaning'," in *Mind, Language and Reality: Philosophical Papers*, vol. 2 (New York: Cambridge UP, 1975): 215-71.

[116] Putnam, "Meaning", p. 219: where 'intension' means something like meaning and 'extension' is the set of things that the term applies to.

differently. then they must be in a *different* psychological state."[117] To put the point slightly differently:

> (1) If Oscar and Elmer have the same mental state of 'knowing the meaning of *A*'. then they understand the same meaning or intension.

Now, by assumption (II), if terms *A* and *B* differ in extension, then these extensions are determined by two different meanings or intensions. This is because, according to the traditional theory, "two terms cannot differ in extension and have the same intension."[118] So, the same intension will determine only one extension. That is, in the case of Oscar and Elmer:

> (2) If they understand the same meaning or intension, then that determines one and the same extension.

It follows, therefore, that:

> (3) If Oscar and Elmer have the same mental state of 'knowing the meaning of *A*', then that determines one and the same extension.

It is this entailment (3), from (I) and (II), that Putnam takes to be false. As his Twin Earth example will demonstrate, it is possible for two speakers to be in qualitatively identical mental states, yet still mean different things by their words.

Before beginning an examination of the Twin Earth thought-experiment, it is important to consider the relevance of the above discussion to internalism. First, 'mental state' is understood as a state in the narrow sense—that is, a state that presupposes the existence of nothing other than the speaker to whom the state is

---

[117] Putnam, "Meaning", p. 222.

[118] Putnam, "Meaning", p. 218.

attributed.[119] In this way the mental state under consideration is exactly like Lockean ideas or Russellian sense-data. Second, Putnam notes that a term's extension is "its 'meaning' in the intuitive preanalytic usage of that term."[120] In other words, insofar as the mental state of a speaker determines extension, it determines the 'common-sense' meaning of the term. But finally, and perhaps most importantly, recall that 'meaning' and 'understanding' are correlative terms. That is, what one understands *in* understanding is a meaning. So, for Putnam, to say, "two speakers cannot be in the same psychological state in all respects and understand the term differently" just is to say, "the psychological state of the speaker determines the intension [meaning]."[121] If understanding is uniquely determined by the psychological state, then so too does the psychological state uniquely determine meaning. So, the traditional theory's assumption (I), that knowing the meaning of a term is a psychological state, is circumlocution for 'psychological states determine meaning'—and this is simply internalism.[122] That is, as the Twin Earth example is a refutation of assumption (I) it is also a refutation of internalism. Off to Twin Earth.

Putnam's rejection of the conjunction of (I) and (II) involves a story of science fiction.[123] Putnam imagines a planet, somewhere far off in the galaxy, that is very

---

[119] Putnam, "Meaning", p. 220.

[120] Putnam, "Meaning", p. 224.

[121] Putnam, "Meaning", p. 222.

[122] Putnam, "Meaning", p. 222. Here Putnam restates the assumptions of the traditional theory replacing (I) with "psychological state (in the narrow sense) determines *intension*.

[123] Putnam, "Meaning", p. 223.

much like Earth. In fact. this planet. called Twin Earth. is virtually indistinguishable from Earth. Any event that happens on Earth also happens on Twin Earth. There is a Twin Earth replica for each and every object on Earth. And every person has a Doppelgänger or duplicate that is qualitatively identical to its Earthian counterpart. These duplicates share a qualitatively identical personal history with their counterparts, even sharing mental states, experiences, lives, etc. In short, their subjective psychological states are identical. In fact. the Twin Earth that Putnam imagines *is* qualitatively identical to Earth; with one exception—water. The liquid on Twin Earth, which is also called 'water', is not H_2O, but a different liquid comprised of XYZ molecules. Moreover, this liquid is phenomenally indistinguishable from Earthian water—XYZ looks like water, tastes like water. etc. So, the only difference between Earth and Twin Earth is that the oceans, lakes, rains, etc. on Twin Earth contain, not water, but XYZ. This is all the relevant information one needs about Twin Earth.

Putnam next imagines that a space ship travelling from Earth visits Twin Earth. Noting the striking similarity to their home planet, these Earthian visitors will, at first. suppose that the term 'water' has the same meaning on both planets. That is, thinks Putnam, until the visitors discover that water on Twin Earth is actually XYZ. With this discovery, the visitors would then report the following to Earth: "on Twin Earth the word "water" means XYZ."[124] Similarly, if a Twin Earth ship travelled to Earth. it would at first falsely report that 'water' meant the same thing on both Earths.

---

[124] Putnam, "Meaning", p. 223.

But. upon discovering the chemical formula of water on Earth. the ship would report that "on Earth the word "water" means H_2O."[125] That is. 'water'. on Earth and Twin Earth respectively. means different things.

Suppose now that it is 1750 on both Earth and Twin Earth. and chemistry has not yet been developed. Suppose further that Oscar is a typical English speaker and Toscar is his Twin Earth duplicate. Now, *ex hypothesi*, Oscar and Toscar are qualitatively identical—that is. their appearance, history, mental lives, etc. are the same. Moreover. since chemistry has not yet developed (i.e., speakers do not yet know about the H_2O-XYZ difference), and since water is phenomenally identical on both worlds, it follows that for any thought (or belief or desire) Oscar has about water, Toscar will have the qualitatively identical thought. The mental state of understanding the term 'water', for example, will be the same for both Oscar and Toscar. In 1750 what Oscar understands by the term 'water' is the same as what Toscar understands by the term 'water'—their 'understanding of the term 'water'' is the same mental state. Now Putnam's point is just this: the extension of water on Earth was H_2O (and similarly, XYZ on Twin Earth) in 1750 just as it is now. So, in 1750. Oscar and Toscar are in identical psychological states, yet Oscar's term 'water' means H_2O and Toscar's means XYZ. That is, contrary to the entailment from (I) and (II). it is possible for two speakers to be in the same mental state, yet for the extensions of their terms to differ.[126]

---

[125] Putnam. "Meaning", p. 224.

[126] Putnam. "Meaning", p. 224.

But are the extensions *really* the same in both time periods? Since this example, at least, hangs or falls on this point, Putnam pauses to consider this question. An internalist may object that the extensions of the two terms are, in fact, different in 1750 from what they are now. In other words, the internalist rejoins, if the extensions of 'water' are the same on both Earths in 1750 (because the chemical difference is not known), then Putnam's conclusion does not follow. But this Putnam takes to be plainly false. For Putnam, a speaker's correct use of a term presupposes that the referent of the term (on that occasion) bears a 'sameness relation' to whatever the speaker and others in the linguistic community have referred to (with that term) on previous occasions.[127] For the term 'water' to be used correctly requires that the liquid is the same liquid that was referred to on other occasions by both the speaker and others in the linguistic community. If the liquids (i.e., XYZ and H_2O) are different now, then they were different in 1750; thus, XYZ is not the same liquid that Earth speakers refer to with 'water' and vice-versa. So, thinks Putnam, it is not as if the extensions of the terms were once the same, and with the advent of chemistry they diverged. Rather, the extensions were *always* different—what has changed is that in 1750 Oscar would have mistakenly supposed that XYZ was water, while in the present he would not. So, Putnam concludes:

> [Oscar and Toscar] understood the term 'water' differently in 1750 *although they were in the same psychological state*, and although… it would have taken their scientific communities about fifty years to discover that they understood the term 'water' differently. Thus, the extension of the term 'water' (and, in fact, its 'meaning' in the preanalytic usage of that term) is *not* a function of the psychological

---

[127] Putnam, "Meaning", p. 225.

state of the speaker by itself.[128]

Two speakers can be in identical mental states (i.e., understanding identical meanings) and have the extension of their terms differ, so the mental state cannot be what uniquely determines the extension.

However, in case Putnam is wrong about the extension of 'water' in the two time periods or the example is considered unnecessarily scientific, consider a further example. Suppose that the terms 'elm' and 'beech' are switched on Twin Earth—on Twin Earth 'beech' refers to elm trees and 'elm' refers to beeches. Suppose also that neither Oscar nor Toscar can tell elms and beeches apart. That is, the concepts they have in mind with respect to either tree are the same. Now since Oscar and Toscar are qualitatively identical, again *ex hypothesi*, every one of their mental states regarding elms will be the same (including the understanding of the term 'elm'). Yet despite the identity of their mental states, Oscar means elm by 'elm' and Toscar's 'elm' means beech. That is, mental states do not determine meaning. So, writes Putnam, "Cut the pie any way you like, 'meanings' just ain't in the *head*!"[129]

Putnam's Twin Earth, then, seems to lead to the same conclusion as the private language argument—mental states (if private or narrow) are insufficient to determine meaning. That is, meaning cannot function as internalists think it does. Recall that the mental states under consideration (i.e., in Putnam's example) are narrow mental states. They are states that depend on nothing other than the speaker who has or is in

---

[128] Putnam, "Meaning", p. 224.

[129] Putnam, "Meaning", p. 227.

that state. The ideas and sense-data of Locke's and Russell's internalism are some such narrow states. But notice that so too are the private mental objects of the private diarist. If the connection between the two is not obvious, compare the private and narrow states. Putnam writes:

> No [narrow] psychological state, properly so called, presupposes the existence of any individual other than the subject to whom the state is ascribed. (In fact…no [narrow] psychological state presupposes the existence of the subject's body even.)[130]

The narrow state, simply, presupposes nothing other than the mind of the speaker. For Wittgenstein, on the other hand, the words of the private language, recall, were "to refer to what can only be known to the person speaking; to his immediate private sensations."[131] Now I think this is already suggestive of the similarity between the private and narrow mental state because if the private state is knowable only to the speaker, then, it too, in a certain sense, presupposes nothing other than the mind of the speaker. That is, private sensations are instances of narrow psychological states— they are in fact the paradigm of a narrow state. If anything is going to qualify as a narrow state it ought to be our private sensations. One's subjective experience of a sensation, if anything, is traditionally assumed to presuppose nothing other than the sensation itself and one's mind. The similarity becomes more explicit, however, when Wittgenstein writes the following:

> *How* do I use words to stand for my sensations?—As we ordinarily do? Then are my words for sensations tied up with my natural expressions of sensation? In that case my language is not a 'private' one.

---

[130] Putnam, "Meaning", p. 220.

[131] Wittgenstein, *Investigations*, § 243.

Someone else might understand it as well as I.[132]

The point here is that if the mental state (i.e., in this case the subjective sensation) manifests itself in any external way, it is not the type of 'private' state with which Wittgenstein is concerned. But, similarly, a narrow state cannot owe its content to anything in the external world, or else it entails the existence of something other than the mind of the speaker in that state—namely, at the very least, the existence of the body of the individual and the behaviours that are manifestations of the state. What this means is that when Wittgenstein says that the language is, then, not private, he may just as well say that the mental state is, then, not narrow. So, insofar as private states are narrow and narrow states are private—that is, narrow states are private in the requisite sense—the object of their criticism is the same. So, although Putnam does not discuss private mental states and Wittgenstein does not discuss narrow states, both arguments are concerned with the mental state that underlies the meaning of a term—a state which is both narrow and private and is supposed, by the internalist, to determine meaning.

To get a better grasp of the similarity between Twin Earth and the private diarist, consider switching the examples. First, take Putnam's narrow mental state that accompanies 'water' and substitute it for S of the private diarist example. Suppose that the diarist, upon having an initial water-experience, undertakes to utter 'water' on the occasions of similar experiences. Again, perhaps the diarist attempts to impress an inner connection between 'water' and the accompanying mental state. The

---

[132] Wittgenstein, *Investigations*, § 256.

problem, for the internalist, is that all of the conclusions about 'S' are equally applicable to 'water'. Since the accompanying mental state is, *ex hypothesi*, narrow—that is, there is recourse to nothing beyond the mind of the speaker in the state—there is no distinction between using the word correctly and thinking the word is used correctly. Without access to anything external there is no difference between the speaker actually having another water-experience and seeming to have another water-experience. So, if 'S' is not meaningful, then neither is 'water'.

Now take a Wittgensteinian private sensation and substitute it into the Twin Earth example. People often claim that love and depression are similar experiences (particularly for young adolescents). They both manifest themselves in similar behaviours—e.g., lack of appetite, inattentiveness to one's surroundings, lack of sleep, etc. Suppose now that 'love' and 'depression' refer to private sensations *à la* Wittgenstein. Moreover, suppose that Oscar and Toscar have identical private sensations for love—the subjective feeling they call 'love' is the same. The catch, of course, is that on Earth 'love' refers to love, while on Twin Earth it refers to depression. That is, since Oscar and Toscar share the same private sensation, the private sensations of Wittgenstein's private language do not determine the meaning of the term 'love'. So, again, insofar as Putnam's narrow states just are Wittgenstein's private mental entities, the target of their attack is one and the same—the internalist's (private and narrow) mental state that determines meaning.

Before continuing to discuss the similarity in the conclusions of the two arguments there is an objection that ought to be considered. Now, I think, as a

demonstration of the continuity between the two positions. it suffices that the conclusions are relevantly similar and that the adversary (i.e., the internalist) is the same. However, to 'tighten' the similarity between Putnam and Wittgenstein I suggested that their private and narrow mental states are one and the same thing. My concern at the present moment is this: an internalist of a certain variety may insist that meaning is determined, at least partly, by the content of narrow mental states. This internalist wants to contend that the content of the narrow state, which partly determines meaning, is entirely a function of its conceptual role. Although somewhat over-simplified, to give its content we need only give a syntactical description of the concept's role in a language (or a language of thought)—its relations to other concepts, its role in (mentalese) sentences, the inferences it licences, etc. This view is internalist as it makes no appeal to the environment in which a speaker is situated, and it depends only upon what is inside the speaker's head (i.e., it involves computational or functional processes within the brain, or something of the sort), but it does not involve an obviously mental object, like a private sensation.

Now, at first glance this should not appear to be problematic. My concern, and all I require, is to show that *Putnam's* narrow states are private in the Wittgensteinian sense. However anyone else construes a narrow mental state is irrelevant. But there is a deeper problem lurking here: Putnam's argument may be able to refute such a theory, while it is not at all obvious that Wittgenstein's can—that is. the arguments are then different. Even if this is true, I doubt that it weakens my argument in any substantial way. I have never claimed that the arguments are

identical. only that they are similar in some very important aspects. Moreover. I have already stated that Wittgenstein's private sensations are instances of narrow mental states—that is. Wittgenstein's private language argument is a more particular version of the Twin Earth argument. In other words. one might expect Putnam's argument to be more broadly successful than Wittgenstein's.

Returning to the problem at hand, however, is it really the case that Wittgenstein's argument cannot work against a conceptual role narrow state? The conceptual role. recall. makes no appeal to a speaker's environment or anything outside of the head (read 'brain'). Is this not already private in the Wittgensteinian sense? Recall Wittgenstein's insistence that anything with behavioural or natural expressions was not private. Conversely, something without outward expressions is private. Similarly, he claims that non-private mental states "stand in need of outward criteria."[133] Again conversely, then private states have no outward criteria. But notice that these describe precisely the mental state whose content is determined by conceptual role. It has no outward expression and no outward criteria—that is, the conceptual role, if truly narrow. is private as well. I suggest, then, that the private language argument may be able to refute this kind of internalism, just as Twin Earth may. Although, as I suggest above, even if this is not the case and only the Twin Earth argument can refute the conceptual role semanticist, it does not detract from my claim that the two arguments are similar in important ways. It only shows that the

---

[133] Wittgenstein, *Investigations*, § 580. Here he says that 'inner processes' need outer criteria, but I call this inner process 'non-private' because one of the conclusions to be drawn from the private language argument is that mental states are not private.

Twin Earth argument is more general in scope, and, therefore, ought to be expected to handle more cases.

Returning to Putnam and Wittgenstein, as I mentioned earlier, the conclusions of these arguments are also relevantly similar. Recall Wittgenstein's discussion of the 'beetle in a box'. As part of the private language argument, the point of this example was to demonstrate how a hidden mental state played no role in determining the meaning of a term. According to Wittgenstein, everyone could have different objects (i.e., beetle) in their box (or no object at all), but the term 'beetle' had the same meaning—something like 'whatever is in the speaker's box'. As Wittgenstein put it, the beetle "drops out of consideration as irrelevant."[134] Now, if one takes the beetle to be a private mental object and the box to be our minds or heads, then, contra internalism, a mental state does not determine meaning because people with different mental representations or accompaniments will mean the same thing with their words.

Now consider Twin Earth again. Oscar and Toscar mean different things by their terms 'water' despite the fact that they are accompanied by identical mental representations or states (i.e., the 'water' mental state is their beetle). For Oscar 'water' means water (i.e., H_2O), while for Toscar 'water' means twater (i.e., XYZ). The psychological state that accompanies 'water', like the beetle before it, drops out as irrelevant. So, whereas Wittgenstein argues that different mental states are compatible with the same meaning, Putnam argues the contrapositive, that the same mental state is compatible with different meanings, because people with the same

---

[134] Wittgenstein, *Investigations*, § 293.

mental representation or accompaniment can mean different things by their words.

The similarity between Twin Earth and the private language argument, then, is simply this: both arguments assume the internalist position (i.e., that private mental objects determine meaning) and demonstrate that the mental state of a speaker does not determine the meaning of a term. For Wittgenstein, if a mental state is essentially private, then speakers cannot 'get out of their heads' to speak meaningfully. Whereas for Putnam, people do speak meaningfully, and regularly refer to things 'outside of their heads', such that the mental state cannot be what determines meaning. That is, in both cases, psychological facts about a speaker do not suffice to determine the meaning of a term. So, Wittgenstein, as well as Putnam, can conclude that 'meanings' (or the determinants of meanings) ain't in the head!

Intentionality

As we have had occasion to discuss, some internalist philosophers want to ground meaning and reference in terms of a speaker's intentions. So, for theorists such as Grice and Searle, a term's meaning is dependent upon a speaker's intentions while (or *in*) uttering a term. Somewhat simply, 'cat' refers to or means cat because that is what a speaker intends it to refer to. Putnam, however, like Wittgenstein before him, argues that this cannot be an informative or useful explanation of our linguistic practices and abilities.

First, note that intending is traditionally construed as a power or ability of the mind. Intentionality just is the mind's ability to refer or represent. But notice that it is also a mental state—it is the state of one's mind while exercising said power or

ability. What the internalist must decide is what type of mental state this intending is. For Putnam, mental states are only of two kinds. A mental state is either 'wide' or 'narrow'—either it presupposes the existence of objects and individuals other than the speaker and the state itself, or it does not.[135] The internalist dilemma, then, is this: is the mental state of intending a narrow state or is it a wide state?

Clearly, the intentional theorist cannot choose the former. That is, if the internalist construes 'intending' as a narrow mental state, then Twin Earth considerations apply. Suppose, for example, that both Oscar and Toscar intend 'water' to refer to *that* particular liquid (i.e., H_2O on Earth and XYZ on Twin Earth). Now, insofar as Oscar and Toscar are premised as qualitatively identical and their mental states are construed as narrow, they are in identical mental states. Oscar's intention is no different from his counterpart's. But again, 'water' means different things on Earth and Twin Earth respectively. So, the intention, construed narrowly, does not determine meaning because identical intentions are fixing different meanings or referents. That is, narrow mental states of intending "do not fix real world reference at all."[136] Twin Earth, I suggest, forces the internalist to the latter horn of the dilemma.

Suppose, alternatively, that 'intention' is conceived as a wide mental state. A wide state, recall, is one that presupposes the existence of things other than the individual to whom the state is attributed. Beliefs are such states. Oscar's belief that

---

[135] Putnam, "Meaning", p. 220.

[136] Hilary Putnam, *Reason, Truth and History* (Cambridge: Cambridge UP, 1981), p. 43.

there is water in the cup, for example, presupposes that there are cups and water, 'water' refers to water, 'cup' refers to cup, etc. Now, Putnam contends that intending presupposes reference in precisely the same way. The internalist desires to say that Oscar's term 'water' refers to water because that is what he intended. But, insofar as intending is a wide state, it already presupposes that 'water' refers to water. That is, "intentions are not mental events that *cause* words to refer: intentions… have reference as an integral *component*."[137] Intentions cannot be invoked to explain the ability to refer, or to mean this rather than that, because they have the very ability that is in need of explanation.

This is, of course, exactly what Wittgenstein tells us in the remarks on naming. Recall that, for Wittgenstein, naming can proceed (free of indeterminacy) only after one is 'master of a language'. That is, intending a name to apply to *that* object is possible only after one is sufficiently skilled in a great number of linguistic practices. So, since for Wittgenstein intending, naming, etc. are a part of or presuppose linguistic abilities, they cannot be invoked to explain those abilities. Putnam's point is the same. When Putnam speaks of intending, it is analogous to Wittgenstein's mental act of intending. Intentions cannot explain our linguistic abilities (i.e., our ability to mean *this* object by that name) because this very linguistic ability is part of what intending is. On both accounts, then, explaining reference or meaning in terms of intentions is circular.

---

[137] Putnam, *Reason*, p. 43.

Semantic Externalism

Thus far, we have considered two of Putnam's critiques of internalism. With the Twin Earth example Putnam argues that the mental state of a speaker does not determine meaning, and in the discussion on intentionality he argues that a mental intending cannot explain meaning or reference on pain of circularity. Moreover, I suggested that these criticisms were analogous to the earlier critiques of Wittgenstein. Like Wittgenstein's 'use determines meaning', however, Putnam too offers a positive account of our linguistic abilities, and of meaning in particular. This view is semantic externalism.

Putnam defines his semantic externalism as the view that "meaning is interactional. The environment itself plays a role in determining what a speaker's words... refer to."[138] Accordingly, for Putnam:

> The content of our words and thoughts is partly determined by our relations to things in our environment (including other people). The fact that what causes us to speak of *water* is water and not some other liquid has everything to do with the fact that the word *water* refers to water, for example.[139]

To see exactly what Putnam has in mind let us return to his Twin Earth example. Putnam's fundamental claim is that Oscar and Toscar, despite identical mental states, mean different things by the term 'water'. Part of the reason for this, recall, is that it is a presupposition of the correct use of 'water' that the liquid referred to bears a 'sameness relation' to the liquid referred to by the speaker and others in the linguistic

---

[138] Putnam, *Representation*, p. 36.

[139] Hilary Putnam, "Hilary Putnam," in *A Companion to the Philosophy of Mind*, ed. Samuel Guttenplan (Oxford: Blackwell, 1994), p. 511.

community on earlier occasions. Suppose that Oscar is being taught the meaning of the word 'water' ostensively—a speaker draws Oscar's attention to a glass of water (i.e., H_2O) and utters, "*This* is water." Putnam's point is that the 'this' is *de re*. For something to be correctly called 'water' it must be the same liquid as the stuff referred to as *this* on this particular occasion (or, in contrast with the twater of Twin Earth, *this* stuff in the actual world). So, for Putnam, "an entity x, in an arbitrary possible world, is *water* if and only if it bears the relation *same$_L$* [same liquid as]…to the stuff *we* call 'water' in the *actual* world."[140] In short, the Earth word 'water' cannot mean the twater of Twin Earth because it is not the same liquid that was referred to as '*this*' on Earth. And since the mental states of Oscar and Toscar are identical, the stuff referred to (i.e., the different liquids) must play a role in determining the different meanings of their term 'water'.

The use of the demonstrative 'this' suggests to Putnam that many more terms than usually thought have an 'indexical component'. Words like 'I', 'here', 'now', etc. are indexicals—their referents vary from utterance to utterance dependent upon context. So, for example, if Oscar and Toscar utter 'I would like a glass of water', the extension of the tokens 'I' differ with the two utterances (i.e., the one 'I' refers to Oscar while the other refers to Toscar). Notice, however, that the extensions differ, despite the fact that their mental state or concept associated with 'I' is identical. That is, thinks Putnam, no one supposes that the traditional theory holds for purely

---

[140] Putnam, "Meaning", p. 232.

indexical terms. [141] But Putnam's point is that the importance of '*this* liquid' to the meaning of 'water' indicates that "indexicality extends beyond the *obviously* indexical words."[142] So, the meaning of 'water' differs on Earth and Twin Earth because the kind of liquid referred to with '*This* is water' differs on each planet. Putnam argues, in other words:

> That words like 'water' have an unnoticed indexical component: 'water' is stuff that bears a certain similarity relation to the water *around here*. Water at another time or in another place or even in another possible world has to bear the relation same$_L$ to *our* 'water' *in order to be water*.[143]

The negative point to be gotten from all this is that just as the traditional theory is false of indexicals it is false of natural kind terms.[144] The positive point is that despite the fact that the traditional theory is not true for indexicals, they are still meaningful. Their meaning derives from the particular uses of the terms and the contexts in which they are uttered. 'I' refers to the person who said 'I', 'now' refers to the time of the utterance 'now', and 'here' refers to *this* place, precisely because that is how they are used. Now we need to extend this point just as indexicality itself was extended. That

---

[141] John Perry has argued that certain indexical belief states cannot be shared because they are essentially indexical, in which case the example does not pose a problem for the traditional theory as it is not a case of different extensions determined by the same mental state: see John Perry, "The Problem of the Essential Indexical," *Noûs* 13 (1979): 3-21. Most agree with Putnam, however, that the traditional theory is not well disposed to handle indexicals, and, moreover, Ruth Millikan has argued that Perry's so-called 'essential indexicals' are not actually indexical: see Ruth Millikan, "The Myth of the Essential Indexical," *Noûs* 24 (1990): 723-34.

[142] Putnam, "Meaning", p. 234.

[143] Putnam, "Meaning", p. 234.

[144] Putnam discusses only natural kind terms (e.g., 'water', 'tiger', 'lemon', etc.), but thinks that this analysis extends to include almost all nouns and a great majority of other parts of speech as well: see, Putnam, "Meaning", p. 242.

is. the same things hold true for natural kind terms. The fact that the traditional theory is problematic should not result in scepticism regarding the meaning of terms like 'water'. Rather. the meanings of these terms are determined by the use. contexts in which they are uttered. and. particularly in the present case. the physical environment in which the speaker is situated. That is, the environment—the liquid itself—has a role to play in determining meaning.

Consider a final example that demonstrates Putnam's point. Suppose that. in the middle of the night. Oscar and Toscar are switched. Oscar wakes up the next morning and upon his first encounter with water utters 'This is water'. Notice that Oscar's utterance is actually false—the liquid he encounters is not water, but twater. But, most importantly. his utterance is false despite the fact that if he had produced the same utterance on Earth. under identical circumstances, and while in the identical subjective state, his utterance would have been true. This change in truth-value occurs independently of anything internal to the speaker. The same is true of meaning. Oscar and Toscar's mental states regarding 'water' will always be qualitatively identical, yet the meanings of their terms are different. That is, a difference in the external world results in a difference in the meaning of their words. It is the contribution of the environment, then, that determines whether 'water' means water (H_2O) or twater (XYZ). So, the meaning of a term is dependent upon appropriate connections, interactions. and relations with things in the external world (including other individuals).

At this point we may lay down precisely the content of Putnam's semantic

externalism: first, the meaning of a term is not dependent upon the mental state of a speaker; and second, the meaning of a term is determined by a speaker's interactions with the social and physical environment in which the speaker is situated. Externalism asserts that the psychological states of a speaker do not suffice to determine meaning; rather, meaning is determined by factors *external* to the speaker. That is, semantic externalism just is Wittgenstein's view that use determines meaning.

Since much of the work demonstrating that use determines meaning is an externalist position was undertaken earlier, I think a few brief reminders are all that is required here. Recall that, in the preceding chapter, I considered two ways to interpret 'use determines meaning'. The first was a social conventionalist reading whereby the uses that a linguistic community makes of a term determines its meaning. This I rejected, however, because the only way terms can be meaningful (i.e., the only way to avoid private language results), is for both speaker and audience to have access to the same object. That is, a term can be meaningful only if a speaker can use it in interactions with others *and* the physical environment. So, the mental state of the speaker does not determine meaning; the use while interacting with the social and physical environment does. And this is just Putnam's point. The fact that Twin Earth speakers use 'elm' while interacting with what we call beeches, and 'beech' while interacting with elms, is what makes it the case that on Twin Earth 'elm' means beech, and 'beech' means elm—again, regardless of the mental states of particular speakers.

To get the similarity precise, contrast Wittgenstein and Putnam's positions

with internalism. For internalists. meaning is determined by the psychological state of a speaker. Similarly. understanding the meaning of a word is a mental state. But if understanding just is grasping meaning or having a particular mental entity before the mind. then the mental entity has to dictate how speakers will use the word. if they use it correctly (i.e., there is not properly understanding if the speaker *cannot* use the word correctly)—that is, for internalists, meaning determines use. But, for both Putnam and Wittgenstein this has things backwards. For Putnam. Oscar and Toscar (narrowly) understand the term 'elm' identically, but the one correctly uses it to refer to elms, while the other correctly uses it to refer to beeches. Similarly with 'water': the mental state does not compel the speaker to use 'water' in one definite way because Oscar and Toscar mean different things by their terms. Oscar means H_2O and Toscar means XYZ. Now, as a matter of fact they may one day get together and decide that their terms mean the same thing (i.e., they may decide that the phenomenal characteristics of the liquid is what is important). They could then go on to use the terms interchangeably. But they might not. They might choose to keep the terms separate, so that Earth 'water' means H_2O and Twin Earth 'water' means XYZ. The point is that it is the future use that will then decide what the term means. Nothing in the mental state forces one option over the other. Wittgenstein makes similar claims in his rule-following considerations. Nothing in a rule itself can compel us to use a word this way rather than that. The same point is found in the remarks on naming. The mental state of 'knowing the name of an object' could not, as yet, compel *any* future use. Similarly. we imagined speakers with identical mental states, yet one speaker

used 'Slab!' to mean 'bring me a slab' and the other meant 'is this a slab?' Again the mental state is not what determines future use. For both Wittgenstein and Putnam this has the cart before the horse.

Wittgenstein and Putnam argue, alternatively, that use determines meaning. Recall the speaker with the strange idiolect whose term 'cat' means water. The reason that 'cat' means water is that that is how the speaker uses the term—the speaker uses 'cat', that is, while interacting with water. Wittgenstein makes the same point in the discussion of the mathematical formula '$x!2$'. Nothing in the formula itself determines how a speaker ought to use the formula; rather the speaker's use determines what the formula means (i.e., whether it means x^2 or $2x$ depends on whether speakers use it to generate $\{1, 4, 9, 16...\}$ or $\{2, 4, 6, 8...\}$). In both cases, the use 'seeps back' to determine meaning; the meaning does not determine the use. The Twin Earth example is used to make the same point.

It is important not to let the chemical composition of the two waters confuse or blur the point; the use of 'XYZ' and 'H_2O' only serve to make explicit that two different liquids are being referred to. The terms mean different things because Earthian and Twin Earthian speakers, while interacting with others and the environment, use 'water' to refer to different liquids. The elm-beech example is equally unproblematic. Earthians and Twin Earthians agree that elms and beeches are different species of trees, and we suppose that they are sorted exactly alike on both planets. The reason that the terms mean different things is nothing mysterious having to do with the hidden 'real essences' or pragmatic decisions deciding which tree is

which. 'Elm' has a different meaning on Twin Earth simply because the speakers use the term in interactions with all and only beech trees. Consider a similar Twin Earth story using the strange idiolect discussed above. Suppose that. due to some strange brain wiring. all Twin Earth speakers are caused to have mental 'cat-ideas' upon interactions with water. Suppose also that. just like the speaker with the odd idiolect. Twin Earthians say things like "What a fine cup of cat" and "Wow! The cat is cold" when drinking water or jumping into lakes. Putnam, like Wittgenstein, has to say that on Twin Earth 'cat' means water. 'Cat' means water on Twin Earth because one's interactions with the environment (both social and physical) determine what one's words mean. That is, for Putnam, too, use determines meaning. So both Wittgenstein and Putnam can be construed as claiming that meaning is determined by use. And 'use' here means nothing more and nothing less than 'how one uses a term while interacting with the social and physical environment'. That is, use is identified or individuated in terms of speaker's interactions with the environment. So, insofar as meaning is determined by use, the social and physical environment in which the speaker is situated determines meaning. That is, 'meaning is use' is semantic externalism, and semantic externalism is 'meaning is use'.

Two Externalisms Reconsidered

I suggest that at this point I have satisfactorily made the case that Putnamian externalism is, to a large degree, continuous with the work of Wittgenstein. I have argued that the private language argument and the Twin Earth thought-experiment are, at the very least. *comparable* arguments—they both reject the idea that a speaker's

subjective psychological state determines meaning. Also, both philosophers reject the internalist modification that makes the 'something mental' an intention because this presupposes the linguistic abilities that the intention is invoked to explain. That is, Wittgenstein and Putnam share their anti-internalism. But, in the preceding section, I suggested that they also share a positive account of language. Since semantic externalism and the view that use determines meaning both make reference to the usage of terms in particular contexts, and by both accounts the meanings of utterances depend upon their use by a speaker in interacting with the physical and social environment, the two views amount to the same thing. In "Two Types of Externalism"[145], however, Anthony Rudd argues that this is wrong. He argues, first, that Putnam has become increasingly Wittgensteinian, to the point that he has now abandoned his own externalism, and second, that the Wittgensteinian and Putnamian positions represent two distinct species of externalism. I take these up in order.

Has Putnam become increasingly Wittgensteinian? The answer here must be yes. Since the writing of "The Meaning of 'Meaning'" Putnam has abandoned metaphysical realism and functionalism; he has become increasingly anti-reductionist, anti-theoretical and anti-scientistic; and in recent years he has become enamoured with the Wittgensteinian idea that if one side of a philosophical theory cannot be given a coherent sense, then the dichotomy itself is senseless.[146] But this, I take it,

---

[145] Anthony Rudd, "Two Types of Externalism," *Philosophical Quarterly* 47 (1997): 501-7.

[146] These attitudes and rejections of various philosophical positions can be found scattered throughout the papers collect in his *Words and* Life: see especially Hilary Putnam, "The Question of Realism," *Words and Life*, ed. James Conant (Cambridge, Mass: Harvard UP, 1994): 295-312.

cannot have anything to do with the latter argument, that the two externalisms are distinct, anymore than the fact that a sapling becomes more like an maple tree could entail that the sapling is a species distinct from the maple. So, although Putnam has *become* increasingly Wittgensteinian, this, by itself, cannot mean that he *was* un-Wittgensteinian. Rudd, I suggest, is not making this mistake, and the real point is that Putnam has now "effectively abandoned 'Putnamian externalism' altogether."[147]

In support of his claim, that Putnam has effectively abandoned his own position, Rudd cites a number of papers published after "The Meaning of 'Meaning'", with particular emphasis on "Aristotle After Wittgenstein".[148] In this particular paper, Putnam devises a thought-experiment in which a synthetic dog is created—a dog that is comprised of the right DNA, but is not descended from wolves.[149] Now we need to decide if this synthetic dog is a 'real' dog. Putnam's point with this example is that there is, as of yet, no correct answer. An evolutionary biologist may say 'no', a molecular biologist may say 'yes', and for the regular dog-owner on the street the scientific properties may be completely irrelevant to answering the question. The point is that one's interests need to be taken into account *before* the question can be answered correctly. Now Rudd takes this to be evidence that Putnam has abandoned his externalism.[150] But, clearly that cannot be right. This *might* demonstrate that

---

[147] Rudd, p. 502.

[148] Hilary Putnam, "Aristotle After Wittgenstein," *Words and Life*, ed. James Conant (Cambridge, Mass: Harvard UP, 1994)

[149] Putnam "Aristotle", pp. 76-7.

[150] Rudd, p. 506.

Putnam has become more of a pragmatist and/or more Wittgensteinian, but in no way can this be construed as an abandonment of externalism. In fact, in his Josiah Royce Lectures, written even later than the article Rudd cites, Putnam insists that he is still an externalist.[151] To my mind, at least, Putnam's change is evidence only of the fact that it has taken Putnam a while to sift the scientism, metaphysical realism, etc. out of his externalism. But, then, the externalism stood or fell independently of the reverential attitude to science all along.

Histrionics aside, Rudd's point is that if Putnam abandons these other views—if he abandons the idea that science can definitively answer questions like 'Is this a 'real' dog?' outside of human interests—then there is no correct answer to the H_2O-XYZ question, and the Twin Earth argument collapses. This leads directly to Rudd's first argument for why the two species of externalism are distinct. Rudd argues, borrowing from Laird Addis, that if Oscar and Toscar share a phenomenal concept of water, the extensions of their terms *could* be the same. The extension could simply be the set of all liquids that are clear, odourless, fill lakes, fall from the sky in the form of precipitation, etc. *no matter what* their chemical composition.[152] The point is, if we are not forced to describe the situation the way that Putnam does (i.e., as involving speaker's having qualitatively identical psychological states with different

---

[151] Hilary Putnam, "Mind and Body," *The Threefold Chord: Mind, Body, and World* (New York: Columbia UP, 1999), pp. 119-21.

[152] Rudd, p. 504; see also Laird Addis, *Natural Signs: a Theory of Intentionality* (Philadelphia: Temple UP, 1989), p. 90.

extensions), then the Twin Earth argument fails.[153] It fails precisely because Twin

Earth is no longer a case in which the same psychological state is determining

different extensions. Now, immediately after making this objection, Rudd states the

following:

> A Putnamian might reply that this misses this point: purely
> phenomenal concepts are inadequate because they are not sensitive to
> differences in real essence. But if this response is meant to be a
> metaphysical one, it would not be available to a Wittgensteinian.[154]

Now I concede that a Wittgensteinian cannot make this reply and I grant that a

Putnamian *might* reply in this manner. But Putnam himself need not make this reply,

nor does he. Putnam claims that:

> It could have turned out that the bits of liquid we call 'water' had *no*
> important common physical characteristics *except* the superficial ones.
> In that case the necessary and sufficient condition for being 'water'
> would have been possession of sufficiently many of the superficial
> characteristics.[155]

With only a cursory glance this would appear to support Rudd's 'Putnamian reply'.

That is, conversely, if there is a 'real essence', then that, rather than the superficial

characteristics, must determine meaning. But this, I suggest, is too quick. Putnam

immediately points to cases where the referents of terms do have important physical

characteristics that differ, but we classify them by superficial or phenomenal

characteristics nonetheless. That diseases with radically different etiologies are called

---

[153] This is why Putnam is supposed to have abandoned his semantic externalism: the dog discussion demonstrates that Putnam too thinks we are not forced to describe things 'one way' and if we do not have to describe things as Putnam did, then there is no argument for his externalism.

[154] Rudd, p. 504.

[155] Putnam, "Meaning", p. 241.

by the same name based upon their phenomenal qualities (i.e., their symptoms), and that the term 'jade' applies to two different minerals, are but two examples. Notice that these are cases in which both XYZ and H_2O would be called 'water'. That is, Putnam recognizes, admits, etc. that it is often the case that we classify things based on phenomenal or superficial characteristics, despite their having different and important 'real essences'. The point is that Putnam cannot claim that purely phenomenal or superficial concepts are inadequate because they do not track real essence, or something of the sort. That is, Putnam, like Wittgenstein cannot make Rudd's reply.

But this, it might be objected, gets Putnam out of the frying pan and into the fire. The point of the reply was to suggest a way for Putnam to avoid a possible objection to his Twin Earth argument. And I, in my zeal to portray Putnam as Wittgensteinian, have just deprived him of this reply (i.e., the reply is not available to Putnam, just as it is not available to a Wittgensteinian). That is, since Putnam no longer has recourse to this reply, we can describe the extension of 'water' as Rudd and Addis suggest; and the Twin Earth argument then fails. But, again, I think this is too quick because, as I will argue (and have argued) nothing in the Twin Earth argument depends on the real essence (i.e., the chemical composition) of water or twater.

Recall the two points of the Twin Earth thought-experiment: first, it demonstrates that meanings do not depend on the psychological states of speakers; and second, the meanings of terms depend on speakers' interactions with their social and physical environment. As I suggested in the preceding section, the chemical

composition of the waters has nothing to do with the argument—it serves only to mark that the two liquids are different. Leaving that aside, suppose that Rudd and Addis are correct and we are not forced to describe things as Putnam does. But equally, we are not forced to describe matters as Rudd would have us do either. What follows? Well, the same two points from the Twin Earth argument: meanings are not determined by the psychological states of speakers and are determined instead by the physical and social environment in which the speaker is situated. The very possibility that we *can* describe things as Putnam does shows that psychological states do not determine meaning. Suppose we *choose* to describe things as Rudd does, such that 'water' means the same thing in both places. The fact that there is a choice at all suffices to demonstrate that the mental state does not determine meaning. That is, if the mental state did determine meaning we would be *forced* into one or the other of the two meanings. This is why for both Wittgenstein and Putnam a 'mental meaning' cannot determine the future uses of a term. Imagine that at time t, Earthians and Twin Earthians discover that their respective waters are of a different chemical composition. Now after much debate it is agreed that at time t_{+1} a decision will be made about future uses of the term. The two choices are, (1) follow Putnam and say that they have different meanings or (2) follow Rudd and say that they mean the same thing. Notice that no matter what the decision, the psychological state drops out as irrelevant. What will determine the future meaning of the word is the future use after time t_{+1}. So, on this point, Putnam and Wittgenstein are in agreement; and the first claim from Twin Earth—that mental states do not determine meanings—still holds.

Suppose at t_{-1} the Ruddians win out and the word 'water' comes to mean both water and twater. So, 'water' comes to mean something like clear, odourless liquid, etc. no matter what its chemical composition. But the fact that the meaning of 'water' comes to include 'no matter what the chemical composition' has everything to do with the fact that 'water' is used to refer to two different chemical substances. That is, the use speakers make of a term while interacting with their environment determines the meaning of that term so that the liquids themselves partly determine meaning. That speakers use the term 'water' in interacting with both XYZ and H_2O is why the term means clear, odourless liquid of whatever composition. And this is just the second component of Putnam's externalism. Moreover, insofar as the use of a term determines its meaning this externalism is also a substantial point of agreement between Putnam and Wittgenstein. In short, the objection does not undermine Putnam's arguments, nor does it point to a difference between the externalisms of Putnam and Wittgenstein.

A final comment before considering another of Rudd's objections. Putnam's critics tend, by and large, to attack the XYZ-H_2O version of the Twin Earth argument. This is done, in part, I take it, because it is the example that Putnam himself draws most heavily upon. But I suspect that it is also partly because it is the easiest version to which one can object. The supposed reliance of the example on metaphysical realism and a kind of hyper-scientism makes it easy prey. But, as I have attempted to suggest, nothing in the argument depends on this. What does all of the work in the argument is that the terms are (or even could be) used to refer to different things quite

independently of the mental states of particular speakers. This generates both the negative thesis that mental states do not determine meanings and the positive thesis that meanings are determined by speakers' interactions with their social and physical environment (i.e., by their use). This was the point of the cat-water switching example; the fact that Twin Earthians use 'cat' in their interactions with water is why 'cat' means water. This is also my reason for mentioning, on a number of occasions, that the chemical composition is irrelevant. The point here is that since Putnam's Twin Earth argument in no way depends on either a strong realism or a hyper-scientism these cannot be invoked to demonstrate that Putnam's early externalism is un-Wittgensteinian.

Along similar lines, Rudd objects that "there is no reason to suppose that the practices of the scientific community should be normative for the rest of us."[156] Rudd's contention is that, even if it is granted that the scientific community gets to stipulate what 'water' refers to, this can have nothing to do with the meaning of the term; nor need the linguistic community follow the scientist's lead.

There are a number of related points to be made here. First, I happen to believe that as a kind of counter-factual matter of fact, if the Twin Earthian scenario ever actually occurred, then those in the ordinary linguistic community on Earth would accept the judgement of the scientific community and use the terms differently. Now Wittgenstein has insisted that his aim is to *describe* our linguistic practices. He writes that "philosophy may in no way interfere with the actual use of language; it can

---

[156] Rudd, p. 505.

in the end only describe it."[157] He urges us to look and see what our practices actually are—that is "don't think, but look!"[158] Rudd recognizes that this is Wittgenstein's position but thinks that the idea of "scientific classifications trumping everyday ones... seems just false as an empirical claim about what our classificatory practices are."[159] He cites Paul Churchland's lamentations that scientific physics and psychology have not replaced folk physics and psychology as evidence of this. Leaving physics and psychology aside, to argue that this is a false empirical claim is, I suggest, to look at our actual classificatory practices with blinders on. There are a great many cases of phenomenally similar things that are classified differently because we do allow the scientific community's classifications to trump purely phenomenal ones. Examples of convergent evolution are cases in point. Plantains are virtually indistinguishable from bananas, pyrite looks just like gold, and dolphins and ichthyosaurs are, from a purely phenomenal perspective, practically identical. Yet in each case we classify these things differently, and, moreover, use different terms to refer to them, in part, at least, on the scientific community's say-so. Lest we start off on a long digression let me return to the point at hand. The point is simply this: there is no reason a Wittgensteinian must necessarily reject Putnam's telling of the Twin Earth story. If our classificatory practices do in fact follow the scientific community, a Wittgensteinian can happily accept this. So, Putnam's reliance on the scientific

---

[157] Wittgenstein, *Investigations*, § 124.

[158] Wittgenstein, *Investigations*, § 66.

[159] Rudd, p. 504.

community cannot in and of itself make his account anti-Wittgensteinian.

Second, if Rudd is merely objecting that the scientific community has too large a role to play in the Twin Earth argument, then all of the previous considerations (regarding the first objection) apply. That is, nothing about the Twin Earth example, nor Putnam's externalism for that matter, depends on the science and/or the chemical compositions of water and twater. As I suggested earlier, the liquids are marked as XYZ and H_2O only to let *us* see that the term 'water' was being used to refer to two different liquids—and that is why the term means two different things.

Finally, part of Rudd's objection is correct—the stipulation of the scientific community can have nothing to do with the meaning of a term, nor, as he puts it, "should it determine the meaning of 'meaning'."[160] This is pure Wittgenstein. But if it is meant as a demonstration that Wittgenstein and Putnam's positions differ, it fails. It fails because this is a substantive (if not the main) point of agreement between the two. Recall that one of the points of both externalisms is that nothing—no mental meaning, no psychological state, no ostensive definition, etc.—can determinately fix our future uses of a term. That is, as I put it earlier, meaning does not determine use. For both Wittgenstein and Putnam the internalist has it backwards—the point of their externalisms is that use determines meaning. Putnam agrees with Wittgenstein, then, that a stipulation by the scientific community *now* cannot determine, drive or compel future uses of a term. That is, for both, the scientific community's stipulation cannot determine meaning. This, again, does not mark a difference between the two

---

[160] Rudd, p. 505.

externalisms.

Before concluding. there is a final issue to address. one which relates to my very last comments. Recall Rudd's initial objection that if Oscar and Toscar share a phenomenal concept of 'water', then their extensions could be the same. The point was that we were not required to tell the Twin Earth story the way Putnam does. I neglected to mention at the time that there is a way to construe the objection so that it beats Putnam's argument. At that time I chose to interpret this objection as meaning we had a choice whether to describe things Putnam's way or Addis and Rudd's way. I constructed a Twin Earth scenario that gave Earthians and Twin Earthians the choice to decide whether 'water' referred to both liquids or whether water and twater would be kept separate. Moreover, by allowing this choice I demonstrated that Putnam's argument (or its conclusions, at least) could still go through. But suppose that Rudd's claim is interpreted thusly: since Oscar and Toscar share a phenomenal concept of water, even now 'water' means something like 'clear, odourless liquid, etc. no matter what its chemical composition'. What follows? This now beats the Twin Earth argument. Oscar and Toscar's psychological states are qualitatively identical, but the extensions of their terms are now the same as well—that is, the extension includes both water and twater. So, Putnam has not produced a case where identical mental states yield different meanings, and his argument fails. Internalism stands.

Suppose this is the right interpretation of Putnam's critics and that they are correct. Unfortunately, Wittgenstein falls with him. That is, insofar as this is an objection to Putnam it is also an objection to Wittgenstein. To see why, consider what

is doing the work in this argument. The reason this succeeds as an argument against the Twin Earth example is that we suppose that it is true that the meaning of a term now determines how it ought to be used in the future. We suppose, that is, that given what 'water' means now, prior to the Twin Earth scenario, if a Twin Earth situation were to occur we would *have to* use water to apply to both liquids. But, if we suppose this is correct, then Wittgenstein too must be wrong—because he insists that meaning does not determine future uses; rather, use determines meaning. So, if we suppose that meaning determines use, then use does not determine meaning, and both Putnam's and Wittgenstein's accounts are mistaken. This I think, perhaps more than anything else, is a striking demonstration that Putnam's semantic externalism just is Wittgenstein's view that use determines meaning.

Conclusion

As I stated at the outset, the purpose of this thesis is to demonstrate the continuity that exists between the work of Wittgenstein and Putnam. More specifically, I intended to show that they share a conception of language—an account by which the meanings of terms are determined by a speaker's interactions with the social and physical environment in which the speaker is situated. That is, I intended to demonstrate that they share their externalism.

To that end, I set out to accomplish a number of tasks. In the first chapter I introduced internalism—the view that both Wittgenstein and Putnam oppose. As the purpose of this chapter was to provide a backdrop for later discussions, I examined a variety of positions and emphasized the emergence of two distinct and/or general internalist trends: (1) the idea that the referent of a term is a mental entity; and (2) the meaning of a term is determined by the speaker's mental state. In the next chapter I discussed the writings of Wittgenstein. During the course of my analysis of some of his later writings, I argued that Wittgenstein rejects both varieties of internalism that I mentioned above. I also suggested that although Wittgenstein was dismantling the internalist programme on one hand, he was simultaneously constructing a positive account of language—this was the view that 'use determines meaning'. I concluded this chapter by arguing that Wittgenstein's positive position was externalist because 'use' is to be understood as use while interacting with the social and physical environment in which the speaker is situated. In Chapter Three I arrived at the heart

of my argument. In this chapter I argued that Wittgenstein's and Putnam's works represented a continuous strand of anti-internalist thought—that is. Putnam, too. rejects the two varieties of internalism discussed in Chapter One. Furthermore, I demonstrated that Putnam's semantic externalism—the view that meaning depends upon the environment in which a speaker is situated—just is the Wittgensteinian view that use determines meaning. I concluded this chapter by responding to Anthony Rudd, who sees the accounts of language offered by Wittgenstein and Putnam as two distinct species of externalism.

Having accomplished my task, however, I think it perhaps more correct to say that my intention was to demonstrate that Putnam's early externalism was already Wittgensteinian. Yet I think this is easy to miss. When Putnam first introduced his externalism it was wed to his functionalism. his scientism, and his metaphysical realism—all very un-Wittgensteinian positions. But Putnam now rejects all of these latter positions. For Putnam, mindedness is now the possession of certain abilities, he is anti-theoretical on most topics, and he advocates a kind of modest or natural realism, which finds the metaphysical realist/anti-realist debate misguided from the start. etc. In other words, he is now a thoroughgoing Wittgensteinian. Since Putnam has subsequently abandoned all of these positions for Wittgensteinian ones, it becomes easy to think he has also abandoned *his* externalism for a Wittgensteinian variant. (As I conceded in an objection to Rudd, Putnam has become *more* Wittgensteinian; but this does not mean his externalism was un-Wittgensteinian then.) I have argued, however, that Putnam's externalism never depended on any of his other

views and, moreover, that it just is Wittgenstein's view that use determines meaning (whether he knew it or not). I would like to close with a few final comments that perhaps explain this Wittgensteinian trend in Putnam's thought.

As I suggested in the preceding chapter, if we take Putnam at his word, then he still advocates *his* brand of externalism. But, since he held his early externalism jointly with functionalism, scientism, etc. it seems odd, at the very least, that he has subsequently rejected these other positions but maintained his externalism. I suggest that this is not coincidental. Consider the case of perception.[161] Putnam, like most traditional analytic scientistic philosophers, held a representational view of perception. For Putnam, perception was best described causally—an object causally impinges upon an individual, which produces a 'visual experience' (like the early internalist 'picture before the mind'), and this experience was to be identified with some physical process in the brain. But, as Putnam began to realize, this makes it impossible to see how words refer to things outside our heads. That is, if our perception of the environment is indirect and meaning depends on a speaker's interactions with the environment, then meaning seems to be, in some sense, indirect as well—meaning would need to be mediated through the mind. But this is just the type of internalist story that Putnam's externalism rejects. Accordingly, Putnam rejects the representational theory and adopts instead a direct realist account of perception, which allows direct unmediated access to the objects in one's environment. In short, Putnam's rejection of the causal theory of perception is a

---

[161] Putnam describes his old causal theory of perception and why and how he came to reject it in favour of direct realism in Putnam, "Senses", pp. 12-25.

direct consequence of his semantic externalism. I suggest (although do not defend) that similar stories can be told for his rejections of metaphysical realism. functionalism. scientism. etc. So. quite simply. Putnam has become more Wittgensteinian because. in the years that have passed since his writing "The Meaning of 'Meaning'". he has recognized the consequences of externalism—consequences that Wittgenstein noticed all along.

Bibliography

Addis, Laird. *Natural Signs: a Theory of Intentionality*. Philadelphia: Temple UP, 1989.

Aristotle. *De Interpretatione*. In *Aristotle: Selections*, ed. W.D. Ross. New York: Scribner, 1938.

Berkeley, George. *A Treatise Concerning the Principles of Human Knowledge*. In *The Works of George Berkeley Bishop of Cloyne*, ed. T.E. Jessup, vol. 2. London: Nelson, 1949.

Chomsky, Noam. "Language and Problems of Knowledge." In *The Philosophy of Language*, ed. A.P. Martinich, 4th ed. New York: Oxford UP, 2001.

Dancy, Jonathan. "Foundationalism and Other Minds." In *Introduction to Contemporary Epistemology*. Oxford: Blackwell, 1985.

Davidson, Donald. "Three Varieties of Knowledge." *A.J. Ayer: Memorial Essays*, Royal Institute of Philosophy Supplement 30, ed. A. Phillips Griffiths. Cambridge: Cambridge UP, 1991.

Fodor, Jerry. *The Language of Thought*. New York: Crowell, 1975.

_____. *Psychosemantics*. Cambridge, Mass: MIT Press, 1987.

Grice, H.P. "Meaning." *Philosophical Review* 66 (July 1957): 377-88.

Hacker, P.M.S. *Wittgenstein: Meaning and Mind*, II. Oxford: Blackwell, 1993.

Hume, David. *A Treatise of Human Nature*, ed. L.A. Selby-Bigge. Oxford: Clarendon, 1949.

Hymers, Michael. *Philosophy and its Epistemic Neuroses*. Boulder, Co: Westview, 2000.

Kripke, Saul. *Wittgenstein on Rules and Private Language: An Elementary Exposition*. Cambridge, Mass: Harvard UP, 1982.

Locke, John. *An Essay Concerning Human Understanding*, ed. Raymond Wilburn. London: Dent, 1948.

McCulloch, Gregory. *The Mind and Its World*. London: Routledge, 1995.

McKinsey, Michael. "The Internal Basis of Meaning." *Pacific Philosophical Quarterly* 72 (1991): 143-69.

Miller, Alexander. *Philosophy of Language*. Montreal: McGill-Queen's UP, 1998.

Millikan, Ruth. "The Myth of the Essential Indexical." *Noûs* 24 (1990): 723-34.

Perry, John. "The Problem of the Essential Indexical." *Noûs* 13 (1979): 3-21.

Putnam, Hilary. "The Meaning of 'Meaning'." In *Mind, Language and Reality: Philosophical Papers*, vol. 2. New York: Cambridge UP, 1975.

_____. *Reason, Truth and History*. Cambridge: Cambridge UP, 1981.

_____. *Representation and Reality*. Cambridge, Mass: MIT Press, 1988.

_____. "Aristotle After Wittgenstein." In *Words and Life*, ed. James Conant. Cambridge, Mass: Harvard UP, 1994.

_____. "Hilary Putnam." In *A Companion to the Philosophy of Mind*, ed. Samuel Guttenplan. Oxford: Blackwell, 1994.

_____. "How Old is the Mind?" In *Words and Life*, ed. James Conant. Cambridge, Mass: Harvard UP, 1994.

_____. "The Question of Realism." In *Words and Life*, ed. James Conant Cambridge, Mass: Harvard UP, 1994.

_____. "Mind and Body." In *The Threefold Chord: Mind, Body, and World*. New York: Columbia UP, 1999.

_____. "Sense, Nonsense, and the Senses." In *The Threefold Chord: Mind, Body, and World*. New York: Columbia UP, 1999.

Rudd, Anthony. "Two Types of Externalism." *Philosophical Quarterly* 47 (1997): 501-7.

Russell, Bertrand. "On Denoting." In *Logic and Knowledge*, ed. Robert Charles Marsh. London: Allen & Unwin, 1956.

_____. *The Problems of Philosophy*. London: Oxford UP, 1959.

_____. *Introduction to Mathematical Philosophy*. London: Allen & Unwin, 1970.

_____. *The Philosophy of Logical Atomism*. In *The Collected Papers of Bertrand Russell*, vol. 8. London: Allen & Unwin, 1986

Schiffer, Stephen. *Meaning*. Oxford: Clarendon Press, 1972.

_____. *The Remnants of Meaning*. Cambridge, Mass: MIT Press, 1987.

Searle, John. *Speech-acts*. Cambridge: Cambridge UP, 1969.

_____. "Proper Names." In *The Philosophy of Language*, ed. A.P. Martinich, 3rd ed. New York: Oxford UP, 1996.

Stroud, Barry. "Wittgenstein on Meaning, Understanding, and Community." In *Wittgenstein—Towards a Re-Evaluation: Proceedings of the 14th International Wittgenstein-Symposium*, eds. Rudolf Haller and Johannes Brandl. Vienna: Verlag Hölder-Pichler-Tempsky, 1990.

Wittgenstein, Ludwig. *Philosophical Investigations*, trans. G.E.M. Anscombe, 3rd ed. Englewood Cliffs, NJ: Prentice Hall, 1958.

_____. *The Blue and the Brown Books: Preliminary Studies for the 'Philosophical Investigations'*. New York: Harper and Row, 1965.

_____. *Tractatus Logico-Philosophicus*, trans. D.F. Pears and B.F. McGuinness. London: Routledge, 1974.

_____. *Remarks on the Philosophy of Psychology*, eds. G.H. Von Wright and Heikki Nyman, trans. C.G. Luckhardt and M.A.E. Aue, vol. II. Chicago: Chicago UP, 1980.