Honors Theses   University Honors Program

12-2021

Why Searle's Objection to Twin Earth is Problematic

Paul F. Scheeler
University of Dayton

Follow this and additional works at: https://ecommons.udayton.edu/uhp_theses

Why Searle's Objection to Twin Earth is Problematic

Honors Thesis

Paul Scheeler

Department: Computer Engineering

Advisor: Trip Glazer, Ph.D.

November 2021

Why Searle's Objection to Twin Earth is Problematic

Honors Thesis

Paul Scheeler

Department: Computer Engineering

Advisor: Trip Glazer, Ph.D.

November 2021

Abstract

The philosopher John Searle wrote that, "The fundamental question in the philosophy of language has always been: How does language relate to reality?" For much of history, people assumed the answer to this question was "the mind", and the debate was then about how the mind related language to reality. Starting in the 1970s, this view was challenged by the philosopher Hilary Putnam and others who claimed that meaning is not only in the mind but also involved the environment (a view known as semantic externalism). In this paper, I will give some background to Putnam's semantic externalism and will then defend his view against Searle's criticism by showing that Searle's view in fact requires externalism.

Disclaimer

I am not committed to the truth of the position I present here. For all I know, it could be wrong.

University *of*
Dayton

Table of Contents

1. Introduction

In this section, I'll give a compressed version of this paper. I'll then take a step back and explore the fundamental ideas of the philosophy of language in the next section. Next I'll turn to Putnam's argument, followed by Searle's objection to it. I'll then present reasons why I believe Searle's objection fails to dissipate the force of Putnam's argument.

Internalism[1] is the view that meaning is internal to the mind. The basic picture of internalism is that if you could look into someone's mind, you'd "see" the meanings there. Put another way, God could know the meaning of your words simply by knowing the information in your brain. He would not have to know anything about the environment, or your brain's relationship to the environment. Internalism was attacked starting in the 1970s by a number of prominent philosophers, including Hilary Putnam (1973) and Tyler Burge (1979). In this paper, I concentrate on Putnam's critique of internalism, in particular his Twin Earth argument for externalism and defend it against an objection by John Searle.

In the Twin Earth thought experiment, Putnam imagines that there is a "Twin Earth" just like our Earth except that what they call "water" has a different chemical constitution, which we'll abbreviate as XYZ (Putnam 1973, p. 285). This different chemical constitution does not change its sensible properties. Prior to the rise of chemistry, there was no way to distinguish between H_2O and XYZ. Thus, Twin Earthians had all the same experiences of XYZ that we had of H_2O. However, since meaning determines reference, Putnam argues that Twin Earthians must mean something different when they use the word "water" (1973, p. 290). But since there was no way to tell H_2O and XYZ apart (prior to chemistry), an Earthian and their counterpart on Twin Earthian had the same psychological states. Therefore, the idea that if you could look into someone's mind,

---

[1] The view here is now referred to as "semantic internalism" to distinguish it from other forms of internalism, including epistemic internalism and mental content internalism.

Eayou'd see the meanings there is false (Putnam 1973, p. 290). Meaning involves not just the mind but also its relationship to the environment.[2]

One of the most famous attacks against Putnam's work comes from John Searle (1983).[3] Searle states that internalism is the belief that the causal powers of the mind are sufficient to determine meaning (1983, p. 230). He argues that the mind sets conditions of satisfaction that constitute meaning (1983, p. 207). The conditions of satisfaction for seeing a chair are that there is actually a chair in one's field of vision; if one is hallucinating, then the conditions of satisfaction are not satisfied, and one doesn't see a chair. Thus, conditions of satisfaction (according to Searle) decide what must be the case for a statement to be true or a name to refer. Searle claims these conditions of satisfaction differ depending on context. As said above, he claims (for the sake of argument) these conditions constitute meaning (1983, p. 207). For example, on Earth the conditions of satisfaction for the sentence, "This is water", depend on what causes the speaker's experience of water (H_2O). However, on Twin Earth, the conditions of satisfaction for the very same sentence, "This is water", depend on the speaker's causal relation to XYZ. Since the mind is what sets these conditions of satisfaction, Searle concludes that internalism is vindicated, and Putnam is mistaken (1983, p. 208).

In this paper, I argue that Searle's critique seems inadequate for two reasons. First, he sidesteps Putnam's argument by redefining internalism. Previously, the basic idea behind internalism is that if you could look into someone's mind, you'd see the meaning there. You don't have to consider anything *external* to the mind (such as the environment). But Searle identifies meaning with Intentional conditions of satisfaction. These conditions differ depending on context. In other words, it is not good enough to look into the speaker's mind. One must consider their environment as well as know what the conditions of satisfaction are. Moreover, Searle argues that the causal powers of the mind

---

[2] Many philosophers since have expanded upon Putnam's semantic externalism. For example, Burge (1979) expanded it to include social kind terms.

[3] Other notable critiques, which I won't explore here, include Crane (1991), Horowitz (2001), Pelczar (2009), Tappenden (2011), and Kotatko (2012).

are sufficient to fix meaning (the conditions of satisfaction), but this seems incorrect, even by his own account. The conditions of satisfaction are not set by the mind alone. If they were, then people in the same psychological state would set the same conditions of satisfaction. But Searle admits that people in the same psychological state set different conditions of satisfaction when they are in different *environments* (1983, p. 208). Thus, Searle's account not only seems to fail to show that meaning is not in the mind but also does not explain how the causal powers of the mind alone are sufficient to generate meaning. To understand these ideas more fully, let us now take some time to explore the key concepts of this paper in more detail.

2. Philosophical Background

It's a good philosophical practice to begin with the commonplace. Oftentimes when we're having a conversation, we run into a misunderstanding. Someone says something that doesn't make sense. Perhaps we don't know what they are trying to say or what they are talking about. In situations like this, we might say, "What do you mean?" or "What are you referring to?" In asking these questions, we've hit upon two basic notions in the philosophy of language: Meaning and reference.

Let's start with meaning. What is meaning? Most of us have a commonsense notion of meaning as a concept or "what's going on inside our head" when we use words. For example, if I say the word "apple," I might mean "a delicious, juicy fruit that contains many vitamins and grows on trees and whose seeds contain cyanide." And if I say, "Hand me the apple on the table," then I am *referring* to the apple. How do I refer to the apple? By getting my listener to look for the delicious, juicy fruit that contains many vitamins… on the table. My words refer to the apple because the apple is the only thing that fits the description.[4] That's why I can't say, "Hand me the goo goo gah on the table." My listener wouldn't know what to look for. It seems then that I refer to the apple because of the meaning of the name "apple." We might put the matter like this: "Meaning determines reference."

In fact, this was the conclusion that Gottlob Frege came to in his 1892 paper, "On Sense and Reference." Frege considers two kinds of cases. The first kind is when two names refer to the same thing. The second kind is when a name doesn't refer to anything. Let's examine the first case first.

In ancient Greece, people noticed that in the morning there was a star that was brighter than all the others. They called this star Phosphorus. People also noticed that in the evening there was a star that was brighter than all the others. They called this star

---

[4] Donnellan (1966) raises the interesting possibility of referring with false descriptions. It would take me beyond the scope of the thesis to respond to this issue here. Kripke (1970) and Putnam (1973) also critique this account, as we shall see later.

Hesperus. Frege noticed that the case we described above poses an interesting puzzle. Consider the sentence, "Phosphorus is Phosphorus." One does not need to know anything about Phosphorus to know that this sentence must be true.[5] Sentences of the form "A is A" are always true and do not inform us of anything. If you asked me, "What is Phosphorus?" and I replied, "Phosphorus is Phosphorus", you would probably feel I had not really answered your question. But now consider the sentence, "Phosphorus is Hesperus." Based on their astronomical knowledge, the Greeks believed that Phosphorus and Hesperus were different stars. But as it happens, Phosphorus and Hesperus are the same star. In other words, the name "Phosphorus" has the same referent as the name "Hesperus." Thus, this sentence reveals something to us: What we thought were two different stars are actually the same star.

How can "Phosphorus is Phosphorus" tell us nothing, but "Phosphorus is Hesperus" tell us something? Frege concludes it because Phosphorus expresses a *sense* (or meaning) (1948, p. 27; see also Lycan 2008, p. 33). When we think of Phosphorus, we think of the brightest star in the morning. When we think of Hesperus, we think of the brightest star of the evening. Thus, these names have both sense (meaning) and reference. For this reason, the sentence "Phosphorus is Hesperus" is both true (because the names have the same reference) and meaningful (because the words express different senses). We learn to connect these two meanings. So now, when we hear "Phosphorus" we think of the brightest star in the morning *and* in the evening.

The second case that Frege considers is sentences like, "Santa Claus has a reindeer named Rudolf."[6] Explorers who traveled to the North Pole failed to find any jolly fat man with red robes who traveled around the globe on a flying sled every Christmas eve to deliver presents to children. In short, Santa Claus does not exist. This means that the

---

[5] I'm sidestepping the question of whether sentences or propositions are truth bearers. See Burgress (2011) for a discussion of this issue. For the purposes if this paper, you may think of "This sentence is true" as abbreviating "The proposition which this sentence expresses is true", if you prefer.

[6] The actual sentence Frege considered was, "Odysseus was set ashore at Ithaca while sound asleep" (Frege 1948, p. 33).

name "Santa Claus" has no reference. The same can be said of "Rudolf". But we still know what the sentence, "Santa Claus has a reindeer named Rudolf" means. This gives further evidence that names must have a sense and not just a reference. In some cases (for example, the case of Santa Claus), names don't have a reference, and so a sense is needed to understand how the name can be used meaningfully.

As we saw, Frege concludes that names have both sense and reference (though sometimes they can fail to refer, as in the case of Santa Claus). Frege also reasons that names refer by means of their sense. Why does the word "Phosphorus" pick out the brightest star in the morning? Why doesn't it refer to Betelgeuse or to the cup of orange juice I had for breakfast? Because, according to Frege, "Phosphorus" expresses a sense (the brightest star in the morning), and that sense determines its reference. In other words, we arrive at the same formula we discovered above: Meaning determines reference.

We could also say this using two other important words in the philosophy of language. The first word is "intension," and it means something like "meaning." Suppose that the meaning of "apple" is as before "a delicious, juicy fruit that contains many vitamins and grows on trees and whose seeds contain cyanide." Then if you have a group of red apples, the intension of the group is "red delicious, juicy fruits…." A green apple can't be part of the group, because it isn't red, and a red strawberry can't be part of the group, because it isn't a fruit that contains cyanide. The intension of the group of red apples excludes green apples and red strawberries. This is because intension is what defines the group. The second word is "extension," and it means "all the stuff that belongs to the group." For example, the extension of red is all the things that are red. So, we can reformulate "Meaning determines reference" as "Intension determines extension."

We might ask ourselves why we'd want to use fancy words when we could use simple words. One reason is that the philosophy of language has been heavily influenced by mathematics (and has in turn heavily influenced mathematics). There's an important branch of mathematics known as "set theory." Set theory is a theory about sets, which are

collections of objects.[7] For example, if I have two red apples that I bought at the store earlier, I have a *set* of apples. We can see now one reason why intension and extension are important: they help us describe sets. The extension of this set is two apples. The intension of the set is "the red apples I bought at the store earlier." The intension is what makes the two red apples part of the set, and not the other apples I saw at the store but didn't buy. The intension also determines the extension. There are only two objects in the world that satisfy my description, "The two red apples I bought at the store earlier," and these objects are the extension of the set.[8]

Notice earlier that I associated an intension with a *description*. I said that the intension of the set of two red apples is "the red apples I bought at the store earlier." This description is called a *definite* description because it picks out a specific set of apples, rather than any set of apples, in virtue of the word "the." Sentences like, "The person who wrote this essay" and "The orange juice I had for breakfast" are also definite descriptions, since they start with "the" and pick out definite objects.

Bertrand Russell (1905) built on Frege's ideas of sense and reference. He argued that proper nouns like Phosphorus and Hesperus are disguised definite descriptions. Consider the way we explained Frege's idea of sense. The sense of "Phosphorus" was "the brightest star in the morning." This is, as we can now see, a definite description. Russell also used Frege's reinvention of logic to build a theory of definite descriptions. One of the fundamental assumptions of this theory are that words should be understood in terms of declarative knowledge (for example, descriptions). This emerges from a fundamental assumption guiding philosophers of language at this time: Meaning is in the head.

---

[7] Set theory originates with Georg Cantor (1874). Russell offered an early version of set theory that avoided paradoxes found in Frege's logic (1903). Modern set theory generally relies on the work of Ernst Zermelo and Abraham Fraenkel (Bagaria 2019). The modern axioms of set theory are thus eponymously known as ZFC (where C indicates the Axiom of Choice).

[8] Set theory has been very important in the history of mathematics and physics. Much of mathematics and physics relies on set theory. While the broader relevance of intension and extension has a lot to do with set theory, set theory itself will not concern us here. Instead, we will concentrate. the underlying notions of meaning and reference, intension and extension.

John Searle critiques Frege's position that proper names express a sense and Russell's position that proper names are disguised definite descriptions. He writes that proper names "refer to and do not describe a particular object" (1958, p. 167). Instead, he argues that we *use* descriptions to teach proper names when we cannot point at the object. For example, if Aristotle were still around, we could teach someone how to use the name Aristotle by pointing at him and saying, "This is Aristotle." But since we cannot do this, we have to use a description—or a cluster of descriptions. Searle allows for this so-called "cluster" for two reasons: (1) We may associate some false descriptions with Aristotle and still be able to refer to him; and (2) Any one of the descriptions we apply to Aristotle may be insufficient to determine reference. Searle writes, "To use a proper name referringly is to presuppose the truth of certain uniquely referring descriptive statements, but it is not ordinarily to assert these statements or even to indicate which exactly are presupposed" (1958, p. 171). Thus, according to Searle, one succeeds in referring because one presupposes certain definite descriptions about the bearer of a proper name, but the proper name is not a disguised definite description. This became known as the "cluster theory" of reference: Speakers refer by virtue of clusters of descriptions (Lycan, 2008, pp. 38-39).

For many years, the basic ideas of Frege and Russell went unchallenged. Whether names were disguised definite descriptions (Russell) or referred in virtue of clusters of descriptions (as in Searle), there was little question that names referred because of something going on inside the speaker's head. Further, there was no question that the meaning of names and other words was inside the speakers head. To see why this commonsense assumption would be challenged, we need to investigate one more topic: indexicality.

Recall from earlier that Searle says there are two ways to teach someone a name: "by ostension or description" (1958, p. 168). Ostension, roughly speaking, means "pointing at." To teach someone the name Aristotle, one could point at the man named Aristotle, and say, "This is Aristotle." In doing so, one uses the word "this." Words like "this," "she," and "now" are called *indexicals* (Braun 2015). Expressions that use indexicals are

called *indexical expressions*. For example, "This is Aristotle" is an indexical expression, since it uses the indexical "this."

Indexicality raises an interesting question. What is the meaning of words like "this," "she," and "now"? It's hard to give a definition of the word "this." It's something like what's in front of us, what we're gesturing at, or maybe what we're thinking about. But what "this" is depends on context. I might say "Take this" and mean "Take the pen I'm holding" or "Take the paper I'm holding" depending on what I happen to be holding. The sentence, "Take the pen I'm holding" is also indexical. If *I*, the author, say it, it means "Take the pen Paul Scheeler is holding." But if someone else says it, the meaning seems to change.

Philosophers of language traditionally assumed that indexicals were special, and that names like "water" or "aluminum" had nothing to do with indexicality. They also assumed, as we have been assuming, that people referred to water because they had some meaning in their head—such as a description or a concept—associated with the word "water." If you could peek inside their brain, you'd see the meaning there. These two assumptions would be critiqued in a surprising way by the philosopher Hilary Putnam.

3. Putnam's Critique

In *Meaning and Reference*, Putnam puts forward a thought experiment to demonstrate that meaning is not in the head (1973). The thought experiment considers the case of a hypothetical place called "Twin Earth":

> For the purposes of the following science fiction example, we shall suppose that somewhere there is a planet we shall call Twin Earth. Twin Earth is very much like Earth: in fact, people on Twin Earth even speak *English*. In fact, apart from the differences we shall specify… the reader may suppose that Twin Earth is *exactly* like Earth. He may even suppose that he has a *Doppelganger*—an identical copy—on Twin Earth… One of the peculiarities of Twin Earth English is that the liquid called "water" is not H_2O but a different liquid whose chemical formula is very long and complicated. I shall abbreviate this chemical formula simply as XYZ. I shall suppose that XYZ is indistinguishable from water at normal temperatures and pressures (1973, p. 285).

Thus, when an Earthian says "water", they refer to H_2O, but when a Twin Earthian says "water", they refer to XYZ. Putnam then asserts that when Twin Earthians in 1750 (before the chemical composition of water was known) say the word "water", they are in the same psychological state as we Earthians are (in 1750) when we say "water" (1973, p. 285). By psychological state, Putnam means states like memory and belief (1973, p. 285). In other words, psychological states have to do with our experience and our knowledge. One way to express this knowledge is through definite descriptions, as we saw in Russell's descriptivism.

Putnam then asks us how we ought to make sense of the fact that an Earthian and their twin on Twin Earth can be in the exact same psychological state and yet refer to different kinds of things (H_2O and XYZ, respectively). According to the status quo postulates:

(1) "[K]nowing the meaning of a term is just a matter of being in a certain psychological state (in the sense… in which states of memory and belief are 'psychological states')" (1973, p. 285).

(2) "[T]he meaning of a term determines its extension" (1973, p. 285).

Yet, Twin Earthians are in the same psychological state as Earthians when they use the word "water," but their objects of reference differ. Twin Earthians refer to XYZ, while Earthians refer to H_2O. If this is so, then either (1) or (2) must be false. Putnam argues it is more intuitive to retain (2) and reject (1). If meaning does not determine reference, then what is meaning? But if this is so, then meaning—the reference-determiner—involves more than psychological state. Putnam expressed this conclusion famously as, "Cut the pie any way you like, 'meanings' just ain't in the head!" (1973, p. 287). By this he meant that meaning is some function of psychological state *and* environment.

It is worth noting that this is not a fight over the meaning of words (in particular, the word "meaning"). The point is not that Putnam is choosing to define the word "meaning" in an unusual way, but that meaning is not determined by psychological state alone. The environment gets a vote. Another way of saying this is that words like "water" have an "unnoticed indexical component" (Putnam 1973, p. 290). The meaning of the word, like the meaning of "this" and "now," is indexed to the environment. Putnam thus considers his position as part of a larger point: the meaning of indexical words "varies from context to context" (1973, p. 290).

4. Searle's Objection

One of the more famous responses to Putnam comes from Searle (1983). To understand Searle's response, we must first understand the notions of Intentionality and conditions of satisfaction. Intentionality is "the power of minds and mental states to be about, to represent, or to stand for, things, properties and states of affairs" (Jacob 2019). Searle capitalizes "Intentionality" to distinguish it from "intentionality". If you want to eat an apple, you might form an intention to eat an apple. In this sense, you have intentionality. But capital-I-Intentionality is more general. For example, when you look at an apple, your visual experience of the apple is *about* the apple (even if you have no *intention* to eat the apple or do anything with the apple).

People in Intentional states have Intentional content. When you think of your best friend, your thoughts are *about* your best friend. Another way to say this is that the content of your thoughts is your friend. Since your thoughts are *about* your friend, we can also say that the Intentional content of your thoughts is your friend. Likewise, the Intentional content of your visual experience of the apple is the apple.

It is important also to distinguish between the terms "Intentional" and "intensional." As we saw just now, Intentionality means "the power of minds and mental states to be about" (Jacob 2019). Earlier, we saw that the term "intensional" has to do with what must be the case to be a member of a set. For example, the intension of "apple" was, in our case, "a delicious, juicy fruit that contains many vitamins and grows on trees and whose seeds contain cyanide." (There is another meaning of "intension", but it will not concern us here.)

Also crucial to Searle's account is the notion of conditions of satisfaction. If someone orders you to leave the room, their utterance sets certain conditions of satisfaction. Their command will be obeyed if and only if you leave the room. Thus, the conditions of satisfaction of their command are that you leave the room. In a similar way, the visual experience of an apple sets conditions of satisfaction. If you see a red apple on a wooden table a few feet from you, then the conditions of satisfaction of the visual experience are

that there is a red apple on a wooden table a few feet from you. If you are dreaming or hallucinating and there is no such apple, then the conditions of satisfaction are not satisfied. Now that we have explored the background of Searle's approach, let us turn to his critique.

In Searle's book, *Intentionality,* he launches a full-scale assault against semantic externalism, the view that non-mental facts help define meaning. His criticism of Putnam's arguments centers on the assumption that indexicality presupposes externalism. Searle concedes, at least for the sake of argument, that Putnam's Twin Earth refutes that meaning is just a description based on perceptual experience. Thus, Searle grants that water is not defined as "a clear, colorless, odorless, tasteless liquid" or any other definition based on sensible properties like "clear" or "colorless" (1958, p. 204). Rather, water is whatever bears the sameness relation to the stuff we drink in water cups, and which fills oceans and rivers, and so on. Searle allows that water has a "hidden indexical component."

Searle denies, however, that this has any special implications. He writes, "two people can be in type-identical mental states… and their Intentional contents can still be different; they can have different conditions of satisfaction" (1958, p. 207). For the purposes of argument Searle grants that meaning or intension, in the sense that Putnam uses the term, is cashed out in terms of Intentional content. Thus, he claims that two people can have type-identical mental states and yet mean different things, but this does not entail externalism.

To further develop his view, Searle uses the term "causal self-referentiality." By this, he means that the conditions of satisfaction refer to the Intentional state itself: Whatever is causing this *perceptual experience*. The conditions of satisfaction of the perceptual experience ("Whatever is causing…") refer to the Intentional state ("this perceptual experience"). This self-referentiality is causal because it involves causation: "Whatever is *causing* this perceptual experience." Note also that Searle believes the power to set causally self-referential conditions of satisfaction is a mental power.

Searle also uses the notion of type-identity. This allows him to talk about "sameness" in a more nuanced way. If you have two apples, these apples are in once sense the same (they're both apples) and in another sense different (they're not the same apple). The first kind of sameness is called type-identity. The apples are type-identical. Type-identity contrasts with token identity, another kind of sameness. An individual apple is token-identical with itself, since it is the same as itself. In contrast, the two apples are not token-identical, since they are different tokens (there are two apples, not one).

Now we are in a position to understand the main thrust of Searle's argument. How can an internalist account explain how someone on Earth (call them "Jones") and their Twin on Twin Earth (twin Jones) have the same psychological state, yet mean different things (have different Intentional contents)? Searle writes,

> [T]heir Intentional contents can be different because each Intentional content is causally self-referential.... The indexical definitions given by Jones on earth of "water" can be analyzed as follows: "water" is defined indexically as whatever is identical in structure with the stuff causing *this* visual experience, whatever that structure is. And the analysis for twin Jones on twin earth is: "water" is defined indexically as whatever is identical in structure with the stuff causing *this* visual experience, whatever that structure is. Thus, in each case we have type-identical experiences, type-identical utterances, but in fact in each case something different is meant. (1958, pp. 207-208).

Searle finishes, "[I]n each case, the conditions of satisfaction established by the mental content (in the head) is different because of the causal self-referentiality of perceptual experiences" (1958, p. 208). He concludes that internalism can account for differences of meaning between twins with type-identical psychological states. This is because the power to set conditions of satisfaction for Intentional content is a mental power, and mental powers exist in the head. Further, these conditions of satisfaction depend on what

causes the Intentional state they belong to: H_2O causes the perceptual experience of water, so the conditions of satisfaction for that experience involve H_2O.

5. My Position

Searle's objection attempts to show that Putnam's premises are compatible with Searle's internalist theory of Intentionality. However, the reasons Searle gives for internalism seem to be inadequate. The first reason is that Searle defines Putnam's criticism out of existence rather than answering it. Putnam criticizes the view that "knowing the meaning of a term is just a matter of being in a certain psychological state" (1973, p. 284). His informal claim that "meanings just ain't in the head" must therefore be understood not as a separate claim but as a funny way of saying that knowing a meaning is not merely about being in the right psychological state. Searle's response to Putnam begins by accepting this fact, and then proceeds to redefine internalism. But we shall let this slide and show that even on Searle's own terms his objection to externalism is troublesome.

As we saw above, Searle's definition of internalism claims that the causal powers of the mind are sufficient to fix meaning (the conditions of satisfaction of Intentional content). According to this conception, people in identical psychological states (having the same perceptual experiences) might still be in different Intentional states, since different things in the world cause their perceptual experiences. In what follows, we'll see that this explanation is actually externalist.

To show that Searle's explanation is actually internalist, we must first examine his claim that the causal powers of the mind are sufficient to fix meaning. We'll sidestep the question of what it means for causes to be sufficient for some effect (such as fixing meaning). Instead, we will propose the more modest principle: There is no change in the effects without a change in the causes (effects supervene on their causes). For example, if you throw a basketball with exactly the same force in exactly the same conditions, then it will have exactly the same trajectory. By the same coin, if a basketball has a different trajectory than previously, then either you threw it differently or the conditions were different (maybe a fan was blowing).

Using the notion of type-identity, we'll now reformulate our principle as follows: Type-identical causes, acting in type-identical circumstances, give rise to type-identical effects.

For example, if you throw a stone off the Tower of Pisa, it will accelerate toward the Earth at 9.8 m/s (minus air resistance). If you throw another stone off the Tower of Pisa, it will accelerate toward the Earth at 9.8 m/s—and as long as the air resistance and gravitational force are the same, the acceleration will be the same.

Let us return to Searle's fundamental claim. When he claims that the causal powers of the mind are sufficient to fix meaning, he intends us to understand that these causal powers are *internal.* The picture that Searle intends us to form runs something like this: There is this system, the mind. Inside this system, there are perceptual experiences, psychological states, linguistic capacities, and meanings. Outside this system, there is the environment. The causal powers of this system fix the meaning of words. Searle grants that even if we could see the world through the eyes of God, we wouldn't be able to see the meanings in the mind. All the same, we would see the causal powers that fix meaning.

Now, we seem to have a problem. If we only consider the causal powers of the mind, we have type-identical causes (the causal powers of the mind) in type-identical circumstances (psychological states) giving rise to non-type-identical effects (meanings). For example, consider the case of Oscar and his counterpart on Twin Earth, Twin Oscar. Oscar and Twin Oscar have the same causal mental powers; they are (according to both Putnam and Searle) in type-identical psychological states; and yet the meaning of "water" for Oscar is H_2O (though he may not know that), and the meaning of "water" for Twin Oscar is XYZ. Therefore, the powers of the mind are insufficient to explain why the meaning is different. One must invoke causes external to the mind. Searle himself grants that the word "water" means one thing on Earth and another thing on Twin Earth, because the Intentional content of "water" is different—and it is different, unknown to Oscar and Twin Oscar in 1750, because different things in the environment cause their experiences. Thus, to determine, one cannot just look to the information inside a speaker's head. One must also consider the environment.

It may be objected that in one sense, Earthians and Twin Earthians are not in type-identical circumstances, and that according to our principle, there is no problem. (Type-

identical causes can give rise to different effects in different circumstances.) Earthians live on a world with oceans composed of H_2O, while Twin Earthians live on a world with oceans composed of XYZ. But these are facts about the mind's relationship to the *environment*, not the mind alone. For internalism to succeed, we must restrict ourselves to the resources of the mind, and the minds of an Earthian and their twin are in type-identical psychological states (ex hypothesi). I conclude, therefore, that Searle's internalist project seems to fail on its own terms.

Where does this leave us? Recall that Frege introduced the distinction between sense (meaning) and reference and claimed that meaning determines reference. Later philosophers tried to clarify on the insights of Frege. Though there were disagreements about the particulars of meaning and reference, they were "in house" disputes. There was no question that if you could only look into the speaker's thoughts, you would know the meaning of their words. Putnam's Twin Earth argument provides evidence against this assumption. Since Searle's objection to Putnam fails, we are left with the conclusions of Putnam: Meaning depends not just on our psychological state and mental powers but on the environment also.

6. Conclusion

In this paper, I argue that Searle's critique is problematic and therefore Putnam's argument for externalism still stands. I claim that Searle's account is problematic for two reasons: (1) It sidesteps Putnam's definition of internalism as the position that psychological state is sufficient to determine meaning; and (2) Searle's reformulated version of internalism appeals to causes outside the mind to fix the Intentional state (which Searle identifies with meaning); in doing so, it grants that the mind alone is not sufficient to fix meaning. Thus, Putnam's critique of internalism survives Searle's objection unscathed.

References

Bagaria, J. (2019, February 12). *Set theory*. Stanford Encyclopedia of Philosophy. Retrieved November 15, 2021, from https://plato.stanford.edu/entries/set-theory/.

Braun, D. (2015, January 16). *Indexicals*. Stanford Encyclopedia of Philosophy. Retrieved November 15, 2021, from https://plato.stanford.edu/entries/indexicals/.

Burge, T. (1979). Individualism and the Mental. *Midwest Studies in Philosophy*. https://doi.org/https://doi.org/https://doi.org/10.1111/j.1475-4975.1979.tb00374.x

Cantor, G. (1874). Über eine Eigenschaft des Inbegriffs aller reellen algebraischen Zahlen. *Journal Für Die Reine Und Angewandte Mathematik*, 258–262.

Collins, J. (2009). Methodology, Not Metaphysics: Against Semantic Externalism. *Proceedings of the Aristotelian Society, Supplementary Volumes*, *83*, 53–69. http://www.jstor.org/stable/20619129

Crane, T. (1991). All The Difference in the World. *The Philosophical Quarterly (1950-)*, *41*(162), 1–25. https://doi.org/10.2307/2219783

Davidson, D. (1987). Knowing One's Own Mind. *Proceedings and Addresses of the American Philosophical Association*, *60*(3), 441–458. https://doi.org/10.2307/3131782

Frege, G. (1948). Sense and Reference. *The Philosophical Review*, *57*(3), 209–230. https://doi.org/10.2307/2181485

Horowitz, A. (2001). Contents Just Are in the Head. *Erkenntnis*, 321–344.

Jacob, P. (2019, February 8). *Intentionality*. Stanford Encyclopedia of Philosophy. Retrieved November 11, 2021, from https://plato.stanford.edu/entries/intentionality/.

Koťátko, Petr (2012). Searle's Defence of Internalism. Organon F: *Medzinárodný Časopis Pre Analytickú Filozofiu* 19:93-106.

Kripke, S. (1980). *Naming and necessity*. Harvard University Press.

Lycan, W. G. (2008). *Philosophy of language a contemporary introduction* (2nd ed.). Routledge Taylor & Francis Group.

Donellan, K. (1966). Reference and Definite Descriptions. In (Martinich, A. 1996 ed) *The Philosophy of Language* (Third Edition, pp. 231–243). Oxford University Press.

Evans, G. (1996). The Causal Theory of Names. In (Martinich, A. 1996 ed) *The Philosophy of Language* (Third Edition, pp. 271–283). essay, Oxford University Press.

Putnam, H. W. (1996). Meaning and Reference. In (Martinich, A. 1996 ed) *The Philosophy of Language* (Third Edition, pp. 284–291). essay, Oxford University Press.

Strawson, P. F. (1950). On Referring. In (Martinich, A. 1996 ed) *The Philosophy of Language* (Third Edition, pp. 215–230). essay, Oxford University Press.

Pelczar, M. (2009). Content Internalism about Indexical Thought. *American Philosophical Quarterly*, *46*(2), 95–104. http://www.jstor.org/stable/20464442

Russell, B. (1905). On Denoting. *Mind*, *14*, 479–493.

Searle, J. R. (1983). *Intentionality an essay in the philosophy of mind*. Cambridge University.

Searle, J. R. (1958). Proper Names. *Mind*, *67*.

Tappenden, Paul (2011). A Metaphysics for Semantic Internalism. *Metaphysica* 12 (2):125-136.

Whitehead, A. N., & Russell, B. (1978). *Principia Mathematica*. Cambridge University Press.