



# ‘Water’ and ‘Water’: On Twin-Earth and the Metaphysics of Words

**ABSTRACT:** *Putnam’s Twin-Earth thought experiment has been hugely influential as an argument in favor of semantic externalism. In this article, I argue that the Twin-Earth thought experiment relies on some previously unnoticed metaphysical assumptions about how to individuate words. My aim is not to argue that semantic externalism is false. Rather I aim to show that Putnam’s thought experiment is only effective as an argument for semantic externalism if we also are committed to certain additional highly controversial and/or implausible claims within the metaphysics of words. I close by arguing that a similar argument for semantic externalism by Burge also relies on unnoticed metaphysical assumptions in the metaphysics of words.*

**KEYWORDS:** Metaphysics of Words, Putnam, Twin-Earth, Semantic Externalism

## I

Putnam’s Twin-Earth thought experiment is likely to be familiar to most readers, but a brief retelling will be useful. Putnam asks us to imagine that there is, in addition to Earth, a further planet which we can call Twin-Earth. Twin-Earth is, by stipulation, an exact duplicate of Earth in almost every way. Because of this, while on Earth there is Oscar, on Twin-Earth there is Oscar’s exact duplicate, Twin-Oscar. Oscar and Twin-Oscar are molecule-for-molecule duplicates.

The only difference between Earth and Twin-Earth is that while on Earth the rivers, lakes, and oceans contain  $H_2O$ , the substance in Twin-Earth’s rivers, lakes, and oceans is not composed of hydrogen and oxygen but is instead composed of some elements entirely missing from Earth, having the molecular structure XYZ. Despite this difference, the XYZ in Twin-Earth’s rivers, lakes, and oceans behaves in exactly the same ways as  $H_2O$  does on Earth. It flows freely downhill, it is refreshing to drink, and the people on Twin-Earth use XYZ for all the same purposes as the people on Earth use  $H_2O$ .

Now suppose that Oscar and Twin-Oscar both live in 1750. Being exact duplicates, the progress of the relevant sciences on Earth and Twin-Earth are identical, and hence the molecular structure of the substance in the rivers, lakes, and oceans on both is entirely unknown to any person on Earth or Twin-Earth. No-one on Earth knows that water has the molecular structure of  $H_2O$ , and no-one on Twin-Earth knows that the molecular structure of the substance in their rivers, lakes, and oceans is XYZ.



Both Oscars, being exact duplicates after all, at some time, use the word ‘water’. Does this word have the same meaning for both Oscar and Twin-Oscar? According to Putnam, they do not. This is because while Oscar and Twin-Oscar are exact duplicates, and hence the intension of the word is the same for them both, they live on Earth and Twin-Earth, and by stipulation there is no substance with the molecular structure  $H_2O$  on Twin-Earth, and no substance with the molecular structure XYZ on Earth. The extension of the word ‘water’ on Earth and Twin-Earth differs. On Earth, ‘water’ picks out  $H_2O$ ; on Twin-Earth ‘water’ picks out XYZ. Putnam therefore concludes that the meaning of the word ‘water’ cannot be given by only the intensions—roughly the concepts associated with the word or the Fregean ‘sense’—of the speakers as Oscar and Twin-Oscar are in the same psychological state as they are exact duplicates. The difference in meaning between the word ‘water’ on Earth and Twin-Earth can only be explained by semantic externalism—roughly the view that the semantic properties of a word are, at least in part, dependent on factors external to the speaker. Intension does not determine extension as ‘water’ on Earth picks out  $H_2O$ , while ‘water’ on Twin-Earth picks out XYZ.

Putnam provides other thought experiments in addition to the one involving the term ‘water’. These are variations of the same point though. In each, we are introduced to two Oscars (or some other people), who speak English and Twin-Earth English respectively, do not know about some more recently discovered structural feature of the relevant substance (aluminium for instance), and have identical psychological states when they use the relevant word for that substance. However, in each case, Putnam argues that ‘the psychological state of the speaker does not determine the extension (or the “meaning,” speaking preanalytically) of the word’ (1975: 143) because the word picks out differing things on Earth and Twin-Earth. Semantic externalism is therefore correct: the meaning of the word ‘water’ cannot be solely determined by the internal states of individuals and must instead also depend, at least in some way, on external factors. The same word can have distinct meanings due to the impact of these external factors. Or, as Putnam famously remarks, this shows that “meanings’ just ain’t in the head!” (1975: 144).

Putnam’s thought experiment continues to be highly influential, both directly on the debate about semantic externalism, but also in extensions into various other topics. Numerous objections and responses have been put forward over the years, and (at least according to Google Scholar) the article that introduces Twin-Earth and Twin-Oscar is now approaching 10,000 citations.<sup>1</sup> Clearly, Putnam’s thought experiment must have some intuitive pull to still be as influential as it is nearly 50 years after publication.

What I want to focus on here, though, does not bear on whether semantic externalism is right or not. What I want to focus on is whether Putnam’s formulation of the Twin-Earth thought experiment, and later a similar argument from Tyler Burge, rests on some assumptions about the nature of words, and more specifically assumptions within what we can call the metaphysics of words. Nothing I

<sup>1</sup> There are too many important works to list here. For just some work that discuss it with respect to semantic externalism, see Kallestrup 2011, Matsui 2021, and Wikforss 2008.

raise in the rest of the article will aim to undermine semantic externalism more broadly. But, if I am correct, then we do at least have some reasons to be more sceptical about using Putnam's thought experiment as a major part of our arguments for semantic externalism. If I am correct, then to accept Putnam's argument will carry with it some highly controversial commitments in the metaphysics of words, commitments that I think we (independently of the issue of semantic externalism) should reject. Putnam might be right in his conclusion, but his argument to that conclusion may not be as strong as it first appears.

## II

The metaphysics of words is concerned with the ontological status and nature of words. It asks the question of what words are, metaphysically speaking. Work in the metaphysics of words has not, however, (normally) directly addressed questions over the semantics of words. That is, work in the metaphysics of words has not been concerned with what some particular word *means* or why it means what it means, and instead, the literature is concerned with issues of word change, how to individuate words, and what makes two particular words (or word-tokens) instances of the same word (or tokens of the same type) (see Miller 2020a for an overview). My aim is to show why these metaphysical considerations are relevant for Putnam-style semantic arguments.

The relevance of this work for Putnam's thought experiment arises from the fact that the thought experiment rests on the intuition that 'water' as uttered by Oscar and 'water' as uttered by Twin-Oscar are the *same* word. That is, for the thought experiment to work, it needs to be the case that Oscar utters a token of the word 'water', and that Twin-Oscar utters a (distinct) token of the *same* word. That Putnam takes these to be tokens of the same word can be seen in what he takes the aim of his article to be. He is interested, as he states it, in whether 'the word 'water' has the same meaning on Twin-Earth and on Earth' (1975: 140). Putnam's answer to this question is no, but he assumes it to be a question about 'the word 'water'', and more specifically a question of whether the meaning of the word 'water' varies between Earth and Twin-Earth. He assumes that Oscar and Twin-Oscar, when they are speaking, utter instances of the same word. Without this assumption, the conclusion that meaning 'ain't in the head' does not follow, as we would instead only arrive at a conclusion about two different words having different meanings. If 'water' as uttered by Oscar, and 'water' as uttered by Twin-Oscar, are instances of different words, then the difference in their meaning is, at least *prima facie*, unimportant to securing the claim that semantic externalism is correct.

But is it reasonable to assume that 'water' as uttered by Oscar, call this token  $w_1$ , and 'water' as uttered by Twin-Oscar, call this distinct token  $w_2$ , are tokens of the same word? On what basis must we individuate words such that  $w_1$  and  $w_2$  are instances of the same word? I will argue that Putnam could accept some accounts of how to individuate words, but they are independently implausible accounts of what it is that makes two instances of a word instances of the same word. I will also argue that more promising ways to individuate words are not open to Putnam as they would not secure the required claim that  $w_1$  and  $w_2$  are instances of the same word.

Before directly discussing issues about how to individuate words, it is worth noting that I am assuming here, as Putnam does, that words are types of which  $w_1$  and  $w_2$  are then tokens. There are views defended in the metaphysics of words (e.g., Bromberger 2011; Miller 2021a, 2021b) which hold that there are no genuinely existing word-types, arguing that only word-tokens exist. I will discuss what these ‘nominalist’ views might mean for Putnam’s thought experiment later, but for now I will assume that word-types exist – i.e., the entity that the phrase ‘the word “water”’ refers to which is distinct from particular tokens of that word – and assume that our questions relate to whether we have good reasons for thinking that  $w_1$  and  $w_2$  are tokens (or instances) of the word-type (or simply word) ‘water’.

Let us then consider some options about how to individuate words to consider on what grounds we can safely hold that  $w_1$  and  $w_2$  are tokens of the same word and whether Putnam might be able to appeal to them. A first option is to hold that  $w_1$  and  $w_2$  are tokens of the same word if they mean the same thing. The idea here would be that it is some semantic facts about words that secures their identity – ‘cat’ is a different word from ‘dog’ because they mean different things – and hence whether two tokens are instances of the same word will depend on whether those tokens reflect those semantic facts.

Such an account would face an immediate objection on the grounds that synonyms are distinct words despite having the same meaning. ‘Doctor’ and ‘physician’, at least in some communities of speakers, mean the same thing, yet we would not want to hold that they are the same word. Of course, such a response relies on there actually being perfect synonyms, which can be doubted. Perhaps more seriously, any semantic account of word individuation would need a way to handle the wide range of meanings that many words have. Dictionaries list various meanings that, intuitively, the same word can have, and on this account, we would need to specify which of these meanings is relevant to the identity of the word. Choosing any one meaning would seem arbitrary.

Perhaps in response we could attempt to build in some flexibility into the idea of using semantics to individuate words. We need not think that every instance of a word contributes to the meaning of the sentence it is part of in *exactly* the same way. This would allow certain minor differences, such as the variation caused by what we intuitively take to be the same word caused by grammatical differences across distinct tokens. This would allow us to hold that these are tokens or instances of the same word despite not contributing the *exact* same semantic content because their semantic contribution is sufficiently similar.

Or we might identify a word with some set of meanings. Tokens of the same word could therefore differ, just so long as the word itself is associated with this set of meanings. However, holding that the instances of the same word need only be ‘sufficiently similar’ in semantic content leads to issues of vagueness over what would constitute ‘sufficient similarity’, and allowing that words might be associated with multiple meanings would not solve the issue of how words can (and do) change their meaning over time. Furthermore, the particular tokens of some type will not have all of those multiple meanings. A word type cannot therefore be individuated by some set of meanings without denying that some (or even all) tokens that are intuitively of that type are not in fact tokens of that type. These are

importantly not simply epistemic issues. What we need is a *metaphysical* account of when it is the case that two words are instances of the same word, and not an account of why we might think that they are instances of the same word. Building flexibility into an appeal to semantics will either underestimate the fluctuating ways that individuals speak over time and the variation across speakers or become so general as to be uninformative.

Perhaps there are responses or adjustments to the semantic approach that would solve these concerns.<sup>2</sup> However, for this discussion, we need not rely on such concerns to show that appealing to semantics will not secure what Putnam needs. The semantic externalist cannot appeal to semantic facts to secure that  $w_1$  and  $w_2$  are instances of the same word precisely because they argue that there is a significant difference between the meaning of those instances. The thought experiment is intended to show that despite Oscar and Twin-Oscar having the same intensions, the extension of  $w_1$  and  $w_2$  are different, and hence  $w_1$  and  $w_2$  have different meanings. To hold that  $w_1$  and  $w_2$  are instances of the same word *because* they mean the same would thus undermine the conclusion that Putnam wishes to draw from the thought experiment. Putnam cannot appeal to meaning claims to secure that  $w_1$  and  $w_2$  are instances of the same word while also trying to conclude that due to their different extensions,  $w_1$  and  $w_2$  have different meanings.

More generally, it is worth noting that while appealing to semantic facts might align with some of our initial intuitions about words and might align with some ways that we talk about words, there is no defence of a semantics-based view in the literature on the metaphysics of words. The view is widely considered to be false, partly for the reasons noted above.<sup>3</sup> If Putnam, or the semantic externalist, wishes to defend this approach to securing why it is the case that  $w_1$  and  $w_2$  are instances of the same word, they would be committing themselves to a view that has been rejected within the literature that explicitly considers such issues. This last point is not a knockdown argument. Perhaps the existing views in the metaphysics of words are wrong. But a case would need to be made for this, and I cannot myself see how semantic facts could solve the issue of sameness of word more widely.

Might we instead appeal solely to the intensions (or 'senses') of the utterances to secure the claim that they are instances of the same word? By stipulation, Oscar and Twin-Oscar have the same intensions—the same 'concept' associated with the word 'water'—so this would ensure that  $w_1$  and  $w_2$  are instances of the same word. The problem here, though, is that using intensions more widely to fix the identity of words would result in us having to hold that speakers in the real world rarely (if ever) speak instances of the same word. Intensions are widely accepted to vary between speakers and it is common for there to be an appeal to some shared referent to avoid

<sup>2</sup> A reviewer suggests that we might hold that words are associated with multiple meanings in an analogous way to Homeostatic Property Cluster (HPC) theories of natural kinds. This might be promising, but such a view has not currently been developed, and as such it is hard to assess it without more details on how it would respond to these and other issues. For example, could such a view account for nonsense words, or functional words (such as articles, auxiliary verbs, and conjunctions) which might be taken to have no semantic properties associated with them.

<sup>3</sup> Rejections of a semantic approach are shared across the spectrum of metaphysical views on words. Hawthorne and Lepore (2011), Irmak (2019), Kaplan (1990, 2011), Miller (2020b), and Wetzel (2009) all have very different views about what words are, but all agree that words cannot be individuated by their meaning.

that variation. In this case, we would not be able to appeal to the shared referent to fix the sameness of two instances of the same word without begging the question in favor of semantic externalism, and hence we would be forced to hold that while  $w_1$  and  $w_2$  might be instances of the same word, there might be no such sameness relation holding between the utterances of ordinary speakers.

Further to this, and more specifically related to Putnam's aims with the Twin-Earth thought experiment, securing the sameness of  $w_1$  and  $w_2$  through intensions would also seem strange in that it would give some priority to intensions over extensions when it comes to saying when two instances are instances of the same word. Putnam's point—or part of his point as I take it—is to say that both intensions and extensions are important when trying to say what a word means, and so privileging one over the other with respect to the issue of sameness of word would be at least an unusual position for the semantic externalist to adopt.

Another suggestion is that  $w_1$  and  $w_2$  are instances of the same word because they have the same 'phonetic form'. In essence, they are utterances of the same word because they are pronounced the same, and we can secure that they are pronounced the same on the basis that Oscar and Twin-Oscar, being molecule-for-molecule duplicates, would pronounce  $w_1$  and  $w_2$  the same way. Appealing to phonetic form does admittedly have some intuitive pull. It is plausible that much of the time in our ordinary lives, we do use something like the phonetic form of words to assess whether those particular words are instances of the same word. Someone observing Oscar and Twin-Oscar may simply assume that  $w_1$  and  $w_2$  are instances of the same word because they sound the same, and some (such as Rowlands, Lau, and Deutsch 2020) have precisely used sameness of phonetic form to motivate the intuition that  $w_1$  and  $w_2$  are instances of the same word.<sup>4</sup>

Can we therefore hold that  $w_1$  and  $w_2$  are instances of the same word because they have the same phonetic form? Again, I think the answer is no. Even a cursory look at the literature on the metaphysics of words would show that this is also not a plausible option. Many papers in the literature on the metaphysics of words start with a dismissal of what has become known as 'shape-theoretic' nominalism, a view most often associated with Bloomfield (1933). Shape-theoretic nominalism holds that two words are instances of the same word if, and only if, they have the same phonetic (or orthographic) 'shape'. That is, if those words are pronounced (or spelt) the same. Such views have been widely rejected for several reasons, but it is sufficient here to note that the view fails to account for even slight variations in how a speaker, or speakers, talk (or write). For example, while we can imagine that Oscar and Twin-Oscar will have the same accent which secures that  $w_1$  and  $w_2$  might be pronounced in the same way, what about some other cases where Oscar or Twin-Oscar say 'water'? Would these other instances be instances of the 'same' word?

It does not take much imagination to see that they will not. For example, let us grant that  $w_1$  and  $w_2$  have the same phonetic form. But, now imagine that, at some later time, Oscar and Twin-Oscar, being exact duplicates, are both suffering from a cold, or are both tired, or both are hungover. In any of these cases, and many other

<sup>4</sup> Strictly speaking, Rowlands, Lau, and Deutsch appeal to the phonetic form of the sentence 'Water is wet', not to the phonetic form of the word 'water' specifically. This difference is not relevant here.

scenarios, Oscar and Twin-Oscar will utter  $w_3$  and  $w_4$  respectively, but will do so with a croaky voice, or may stutter over the word, or otherwise pronounce  $w_3$  and  $w_4$  in any other way than the way that  $w_1$  and  $w_2$  were pronounced. Whichever of these situations we imagine, it will not be the case that  $w_3$  and  $w_4$  have the same phonetic form.

If we used phonetic form to secure why  $w_1$  and  $w_2$  are instances of the same word, then we must now hold that  $w_3$  and  $w_4$  are instances of a different word. This is surely not something Putnam or other semantic externalists would want to commit themselves to. It would result in a view where, even in one sentence due to natural variance in the way that an individual speaks, there might be multiple words being tokened, contrary to our intuitions. If Oscar says, 'Water is wet and water is clear', then the semantic externalist will surely want to hold that 'water' and 'water' in that sentence are instances of the same word, but we cannot be sure unless we stipulate that Oscar does not vary their pronunciation across the utterance. I take Putnam's claim to be one that he would want to apply to all instances of 'water'. If we use phonetic form to secure that  $w_1$  and  $w_2$  are instances of the same word, then this is simply not possible.

Further issues with using phonetic form can be generated in other ways. Imagine a minor variation of Putnam's thought experiment. All is the same as in the original, except that when they utter  $w_1$  and  $w_2$ , Oscar and Twin-Oscar pronounce these words ever so slightly differently. One, say, says it with a slightly longer vowel sound than the other. This is a very minor variation, and it does not seem that such a small variation should matter to Putnam's main aim which is to argue in favor of semantic externalism. Yet, if we use phonetic form to explain why particular words are instances of the same word, even this would mean that, strictly speaking,  $w_1$  and  $w_2$  would not be instances of the same word, and hence the thought experiment fails.

Note, that as in the case of appeals to semantics, it is implausible to try to build some flexibility into an appeal to phonetic facts as it would be vague as to what level of flexibility would be allowed. And, just as above, this is again importantly not simply an epistemic issue. To secure that  $w_1$  and  $w_2$  are the same word, we need a metaphysical account of when it is the case that two words are instances of the same word, and not merely an epistemic account of why we might think that they are instances of the same word. Building flexibility into an appeal to phonetic form will underestimate the fluctuating ways that individuals speak over time and underestimate the variation across speakers. Phonetic form has rightly been rejected by all parts of the literature on the metaphysics of words, and it would not strengthen the semantic externalist's position to appeal to it to support Putnam's thought experiment.

Another option could be that words are individuated by their origins. Given that at least some semantic externalists also accept direct reference theories, this might be the most natural way for the semantic externalist to solve our issue.<sup>5</sup> The idea here will be that words are individuated in terms of their (unique) originating event, and

<sup>5</sup> Semantic externalism has been combined with direct reference theories and originalism perhaps most influentially by Kripke's (1980) concerning proper names, and Kaplan (1990).

two words are then instances of the same word if they are part of the same causal-historical chain of uses of that word back to that originating event.

Unlike appeals to semantic facts or to phonetic form, originalism about word individuation does have some defenders in the metaphysics of words literature (e.g., Millikan 1984; Sainsbury & Tye 2012, and Gasparri 2016). The versions of originalism presented in those papers and books offer a nuanced view of word individuation which can handle some of the most immediate objections, such as from cases where two distinct words appear to have the same origin, or where a single word might have two distinct origins. There do remain some very serious concerns for all versions of originalism. For example, Hughes has argued that a ‘new non-word-type  $W^*$  instantiated in an originating event must be sufficiently propagated by a linguistic community in order to be classified as an actual word-type  $W$ ’ (Hughes, 2023: 6). The problem, as Hughes goes on to argue, is that at least in the moment of origination, there can be no propagation of that new word precisely because it is new. This means that supporters of originalism, even in its more developed forms such as Stojnic’s (2021), cannot account for the difference between cases where there does seem to be an originating event – e.g., a child neologising the ‘non’-word ‘un-turnaround’ – and yet no new word is created, and cases where a new word is successfully originated.

However, for this discussion, we need not rely on such arguments. Even if originalism is the correct way to individuate words, then there is a simple reason why this will not secure that  $w_1$  and  $w_2$  are instances of the same word. The reason is that as Earth and Twin-Earth are themselves distinct and are distinct whether we structure the thought experiment as involving distinct possible worlds or merely distinct planets within the same possible world. Even if we accept an originalist account of word individuation, then under this view of words,  $w_1$  and  $w_2$  will be instances of different words as the causal-historical chain that  $w_1$  is part of will be distinct from the causal-historical chain that  $w_2$  will be part of. The word that  $w_1$  is an instance of will have a different origin from the origin of the word that  $w_2$  is an instance of simply in virtue of them being words on Earth and Twin-Earth respectively.

This will be the case even if we suppose, as Putnam does, that Oscar and Twin-Oscar are exact duplicates and hence that they have all the same personal history. The personal history of Oscar and Twin-Oscar is irrelevant to the question of what the origin of the word ‘water’ is on Earth and Twin-Earth. Indeed, even if it were the case that Oscar and Twin-Oscar coined the word ‘water’ on Earth and Twin-Earth respectively—even if the utterances  $w_1$  and  $w_2$  happen to be the originating events of the word ‘water’ on Earth and Twin-Earth—then the word ‘water’ on Earth and the word ‘water’ Twin-Earth would still be distinct words. Earth and Twin-Earth might be exact duplicates, but exact duplicates are not identical to each other. Just as Oscar and Twin-Oscar are exact duplicates but distinct, the words ‘water’ and ‘water’ on Earth and Twin-Earth might have origins that are exact duplicates but will remain distinct in virtue of the fact that Oscar coined one, and Twin-Oscar coined the other. Or, put another way, the words might have exactly resembling origins, but resemblance is not the same as identity.

This means that originalism cannot help us to secure the sameness of  $w_1$  and  $w_2$ . Even granting the correctness of originalism, the distinctness of Earth and Twin-Earth (and of Oscar and Twin-Oscar) will mean that it simply will not be the case that the word 'water' on Earth and the word 'water' on Twin-Earth will have the same origin, and hence  $w_1$  and  $w_2$  are not instances of the same word.<sup>6</sup> Note also that the same arguments will also apply against broader appeals to the history of words as a way to determine if two words are instances of the same word. Such accounts suggest that words are to be individuated via their history such that two words are identical 'if and only if they have the same history' (Irmak 2019: 12). For the same reasons that mean originalism fails to help Putnam, history will also fail. It is simply not the case that the word 'water' on Earth and 'water' on Twin-Earth share the same history. One, for instance, will have a history which includes previous utterances of 'water' by Oscar, and the other will have a distinct history which includes previous utterances of 'water' by Twin-Oscar. Their histories might be exact duplicates, but this is not enough to say that they are the same. In so far as  $w_1$  is an instance of the word 'water' on Earth, and  $w_2$  is an instance of the word 'water' on Twin-Earth, appealing to the history of a word also cannot help us to secure Putnam's claim that these are, in fact, the *same* word.

Where does this leave Putnam's argument for semantic externalism? I propose that at least if we assume that words and their instances stand in a type-token relationship—something we will reconsider shortly—then the argument rests on an unstable premise. There does not seem to be any plausible way in which Putnam can secure the claim that Oscar and Twin-Oscar are uttering instances of the same word, and hence it is not clear that the thought experiment does tell us anything about how the *same* word might vary its meaning due to external factors.

For completeness, it should be noted here that some who defend the type-token approach have themselves recognized that issues of the sameness of particular words might not have firm answers. For example, Wetzel defends the view that 'there is nothing interesting all and only uttered tokens of a particular word have in common other than being tokens of the word' (2009: 106–7). Hawthorne and Lepore have defended 'sloppy realism' and argued that 'there either are facts we may never know or simply no facts at all about the myriad borderline cases [of when particular words are instances of the same word] left unresolved by our capacity to settle questions in the area' (2011: 36). Both options are, at least technically, available to Putnam and the semantic externalist. They might, following such accounts, simply insist that  $w_1$

<sup>6</sup> What if the substance called 'water' on Earth was miraculously changed from  $H_2O$  to XYZ without anyone in the linguistic community noticing? An originalist may be able to respond in this case, on the assumption that 'water' spoken before ( $w_{t1}$ ) and after ( $w_{t2}$ ) this change can trace back to the same origin. We would need to hold that  $w_{t2}$  does not constitute the origin of a new word. If we are a direct theorist, then this becomes less plausible as we may view  $w_{t2}$  to be a new baptism for XYZ. I think that originalism is independently implausible, in part due to the concerns noted, and this could be used as an argument against semantic externalism on the grounds that it relies on an implausible metaphysics of words. But perhaps others will be more willing to accept these conditions. If nothing else, it becomes interesting that this argument can only be supported if combined with an originalist view of words, under certain conditions (including perhaps a rejection of some versions of direct reference theory), and only with a variation of Putnam's original thought experiment.

and  $w_2$  are instances of the same word, or hold that there are some facts that we might never know about which secures that they are instances of the same word.

Such approaches, though they might secure what Putnam's thought experiments need, are hardly going to be persuasive to those that have not already accepted both Putnam's thought experiment and the conclusions that he arrives at. For anybody that is sceptical of semantic externalism, an appeal to brute identity—à la Wetzel—or to sloppy realism—à la Hawthorne and Lepore—will look like cases of begging the question. The issue at hand is precisely whether or not  $w_1$  and  $w_2$  are instances of the same word and simply asserting that they are will not persuade anyone who is not already convinced.

### III

While type-token views remain the most popular sort of ontology in the literature on the metaphysics of words, there are other alternatives. Perhaps one of those views might provide some way to justify the claim that  $w_1$  and  $w_2$  are instances of the same word.

A first major alternative to mention is Kaplan's stage-continuant ontology. Under this ontology, particular utterances are 'stages' of words, and words themselves are 'continuants'. Words are therefore made up of 'interpersonal stages along with some more mysterious intrapersonal stages' (Kaplan, 1990:98; see Kaplan 2011). For Kaplan, words are temporally extended objects, with stages as their parts. Applied to Putnam's thought experiment, the issue would now be whether we have good reasons to maintain that  $w_1$  and  $w_2$  are stages of the same continuant.

Even before we consider what Kaplan himself says about this issue, it is worth noting how strange a metaphysics this would be for Putnam to adopt.  $w_1$  and  $w_2$  are utterances that take place in distinct possible worlds. If we accept that they are stages of the same continuant/word, this would make words inherently modal entities whose identity is given by what stages they are composed of, where a word is composed of stages from many (potentially infinitely many) possible worlds. Note that this is not the same as a mere commitment to transworld identity. Transworld identity secures that the same object exists in more than one possible world. Adopting Kaplan's ontology for words would go further, requiring us to hold that a word is *composed* of many stages across different possible worlds. It is then not only that the word 'water' is the same across worlds, but a stronger claim that the word 'water' only partly exists here in the actual world because it also partly exists in other possible worlds also (in the sense of having parts in those possible worlds). This would make words a particularly strange sort of entity, and one that many semantic externalists may not want to commit themselves to the existence of.

Returning to Kaplan, what does he say about what makes two stages stages of the same word (or continuant)? Kaplan explicitly rejects the idea that two stages are stages of the same continuant if they resemble each other. Instead, words are like families and, just like the people that make up families, the stages that make up words might resemble each other, but they need not. Instead, he holds the historical connection between stages that make them stages of the same continuant: two stages are stages of the same word if they 'descend from a common ancestor'

(Kaplan 2011:509). This historical connection, though it is necessary, is not sufficient for Kaplan. In addition, Kaplan argues that 'a sincere subject, intending to repeat a word that has been uttered by an examiner, will, indeed, utter that word' (2011:518). Thus, we can summarize Kaplan's view as being that for two stages to be stages of the same continuant they must both descend from a common ancestor *and* the speaker must intend to utter a stage of that word by 'repeating' that word.

Leaving aside broader objections that might be raised against Kaplan's stage-continuant ontology (e.g., in Hawthorne and Lepore 2011), it is unclear whether this account could be useful for Putnam. Although Kaplan has additional elements in his account, he centrally appeals to historical connection to some 'common ancestor'. But, analogously to the above discussion about originalism, it simply is not the case that  $w_1$  and  $w_2$  will share a common ancestor. The ancestors of  $w_1$  are stages that have been uttered (or inscribed, etc.) on Earth, while the ancestors of  $w_2$  are stages that have been uttered (or inscribed, etc.) on Twin-Earth. The ancestors of  $w_1$  and  $w_2$  may be exact duplicates, but, just as I argued above, this is not enough to secure the claim that  $w_1$  and  $w_2$  are instances of the same word. Exact duplicates are, by definition, qualitatively identical but numerically distinct, and hence even on Kaplan's account,  $w_1$  and  $w_2$  will be stages of distinct continuants (or words).<sup>7</sup>

Another major alternative to type-token views are nominalist views. As noted above, nominalism has historically been taken to be associated with 'shape-theoreticism', and the idea that what makes two words instances of the same word is that they have the same phonetic (or orthographic) 'shape'. More recently, though, other forms of nominalism have been developed which can avoid the objections raised against shape-theoretic nominalism. These newer forms of nominalism still deny the existence of words, or, more precisely, deny the existence of words *qua* (abstract) types. For these nominalists, only word-tokens exist—particular words such as the particular ink patterns on this page. Nominalists then often posit resemblance relations between word-tokens and invoke 'archetypes' or 'collections' as non-ontologically committing replacements for word-types. Any type-talk that we might engage with in ordinary language or in our scientific theorising should, for the nominalist, be taken to be 'mere' talk. Type-talk is not ontologically committing and 'no explanatory work will be done by picking out some one abstract entity as the sign type. That's to say, it might be that reifying sign types would be explanatorily superfluous' (Cappelen, 1999:100).

As above, we will not spend too long here discussing why the nominalist defends their ontology (see Bromberger 2011; Miller 2021a, 2021b for examples of nominalist views), nor what arguments might be raised against nominalism (e.g., Wetzel 2000). Rather, we will focus on whether nominalism might provide a solution that Putnam and the semantic externalist might be able to adopt to secure their claim that  $w_1$  and  $w_2$  are instances of the same word.

<sup>7</sup> Interestingly, much of the discussion of Kaplan's account in the literature has focused on criticisms of his appeal to the intentions of speakers; see Cappelen (1999) and Munroe (2016). My view is that these objections are significant and do provide good reasons for us to reject any appeal to intentions. However, we need not engage with that issue here. Kaplan's metaphysics cannot help Putnam here.

Whatever the merits of nominalism more broadly, we can immediately see that it cannot straightforwardly provide a way to solve the issue facing Putnam and the semantic externalist. While the nominalist will argue that invoking resemblance relations, collections, and other mechanisms can do much of the explanatory work that we want a metaphysics of words to do, the nominalist simply cannot accept any claim that two words are instances of the *same* word. This is not possible as there are, strictly speaking, no *words*, at least in the sense that we normally understand this claim as being about whether they are tokens of the same word-types. The nominalist can hold that  $w_1$  and  $w_2$  *resemble* each other in various ways, and that they might be members of the same collection, but this is not the same as holding that there is an *identity* relation between  $w_1$  and  $w_2$  where they are, in one sense, the *same* word.

Is this resemblance enough for Putnam? Might we hold that  $w_1$  and  $w_2$  are merely *resembling* words, and from that conclude that ‘meanings just ain’t in the head’? I think the answer is still no. If we were to try to retell the thought experiment but instead hold that Oscar and Twin-Oscar utter merely resembling words, then the intended conclusions simply do not follow. The thought experiment only supports semantic externalism if they are the *same* word. If the claim is that  $w_1$  and  $w_2$  are resembling words, then the difference in their meaning cannot secure the claim that semantic externalism is correct. If  $w_1$  and  $w_2$  merely resemble each other, then the difference in their extension does not tell us anything about the nature of meaning except that different words might have different meanings—something that surely no-one would deny.

It might be responded here that resemblance is enough if, granting nominalism for the moment, we hold that Oscar and Twin-Oscar are in the same psychological state when they utter  $w_1$  and  $w_2$ . However, as I will discuss in more detail below, if  $w_1$  and  $w_2$  are not instances of the same word then we have no reason to hold that Oscar and Twin-Oscar are in the same psychological state. This is because, if  $w_1$  and  $w_2$  are not instances of the same word, as they are not for the nominalist, then Oscar and Twin-Oscar will not be in the same psychological state because the word that they know or are uttering is part of each of their distinct psychological state. Oscar’s psychology contains knowledge of (the meaning of)  $W_1$  leading them to utter  $w_1$ , while Twin-Oscar’s psychology contains knowledge of (the meaning of)  $W_2$  leading them to utter  $w_2$ . The intensions of Oscar and Twin-Oscar are therefore different, and hence the semantic externalist conclusion can be resisted.<sup>8</sup> An appeal to nominalism will not save Putnam’s thought experiment, however well motivated the view might be elsewhere.

## IV

There are a couple of lines of response that need to be handled at this point. First, I have used some quotes from Putnam to support the idea that the Twin-Earth thought experiment involves considering whether ‘water’ on Twin-Earth and ‘water’ on

<sup>8</sup> This is not to deny that resemblance is sufficient for other situations where we want to say that we are uttering the ‘same’ word, nor is this intended to be a critique of nominalism about words. Indeed, in my view, the opposite is the case (see Miller 2021a, 2021b, 2022, 2024). But I argue resemblance is insufficient for Putnam’s argument.

Earth are instances of the same word, but we might think that this talk of 'the word 'water'' from Putnam is mere loose talk, and the sameness of 'water' and 'water' on Earth and Twin-Earth is not important. The 'sameness' of the word(s) might be merely tangential to the fundamental issue. Putnam's argument is about the intension and extension of the words, irrespective of whether or not they are, metaphysically speaking, the 'same' word.

This response, I think, is right but only to a certain degree. My aim in this article is not to show that externalism is false, and so nothing here is intended to show that we *cannot* accept externalism in light of positions within the metaphysics of words. Rather, my aim is to show that Putnam faces a dilemma. On the one hand, we could hold that 'water' and 'water' on Earth and Twin-Earth are instances of the same word, but if they are the same word, then we need to know what grounds that sameness, and I have argued that there is no promising route. On the other hand, we could deny that they are the same word. The sameness of 'water' and 'water' is irrelevant if all that Putnam needs is that the psychological or conceptual states of the Earth and Twin-Earth speakers are not sufficient to generate either meaning of either word. Putnam does not require that 'water' and 'water' are instances of the same word.

However, if 'water' on Earth and 'water' on Twin-Earth are different words, in what sense can Oscar and Twin-Oscar be 'psychologically' the same as Putnam requires for his argument. The entire thought experiment relies on Oscar and Twin-Oscar being psychologically the same, but if they know different words then they cannot be psychologically the same. Quite simply, Oscar and Twin-Oscar would know different words and as what words we know is a matter of our psychology, Oscar and Twin-Oscar must be psychologically different.<sup>9</sup>

Putnam's supporters might respond here that Putnam needs only to say that Oscar and Twin-Oscar have the same 'narrow psychology', not psychology more broadly. We could then hold that a difference in relational properties such as knowing a word does not entail a difference in psychology, at least in this 'narrow sense'. However, this is unsatisfying. To take knowing a word to be a relational property suggests that knowing a word requires a relation to some *thing*—the word. What then, metaphysically, is this word? Putnam's view again seems to rely on some unspecified metaphysics of words, and I have already argued that it is unclear what metaphysics of words would allow us to say why 'water' and 'water' are the same word.<sup>10</sup>

<sup>9</sup> Note the connection with nominalist views in the metaphysics of words here. For nominalists, it is simply always metaphysically the case that two speakers utter different words, even if they are (exactly) resembling. If nominalism is right, then it follows that speakers know different, albeit resembling, words, and on the grounds that knowing the meaning of a word is psychological state, we cannot hold that Oscar and Twin-Oscar are psychologically the same.

<sup>10</sup> Might we respond that knowing a word is not a psychological state? Perhaps. But this is, I argue, an assumption that I am sharing with Putnam. Putnam does not deny that Oscar and Twin-Oscar are in certain psychological states that relate to their knowing the meaning of the word 'water'. Indeed, Putnam gives the example of 'knowing the alphabet' as an example of a psychological state (1975: 136). Furthermore, even if we hold that knowing a word is not a psychological state, it is common for philosophers to hold that knowledge 'incorporates' a psychological state (see Nagel 2013). Knowledge of the word being incorporated within a psychological state suffices for both Putnam's argument and mine, so the arguments can be simply rephrased.

Furthermore, Putnam does not deny that knowing the meaning of a word is a psychological state—he wants instead to deny the claim that knowing the meaning of a word is *just* a matter of being in a certain psychological state. Given this, we can rephrase the problem facing Putnam. The thought experiment requires that Oscar and Twin-Oscar are psychologically the same. What Putnam means by this is that they have the same intensions (or roughly concept) associated with the word ‘water’—both have the psychological state ‘term ‘W’ is associated with intension I’ (or something similar). My claim is that Putnam is not entitled to this claim without argument. He is not entitled to the claim that both Oscar and Twin-Oscar have the psychological state ‘term ‘W’ is associated with intension I’ rather than the view that Oscar has the psychological state that ‘term ‘W<sub>1</sub>’ is associated with intension I’ and Twin-Oscar has the psychological state ‘term ‘W<sub>2</sub>’ is associated with intension I’. Whether or not these are part of ‘narrow’ psychology is then irrelevant, *at least* to the Twin-Earth thought experiment, even if not to Putnam’s wider claims. In the relevant sense, it is simply not clear why we should think that Oscar and Twin-Oscar are in the same psychological state, and this only assumes (as Putnam himself accepts, cf. footnote 10) that knowing a word is a psychological state.

My argument only requires the view that the psychological states “W<sub>1</sub>’ is associated with intension I’ and “W<sub>2</sub>’ is associated with intension I’ are different, while Putnam requires a way to show that ‘W<sub>1</sub>’ and ‘W<sub>2</sub>’ are the same word. Without that, if knowing a word is (part of) a psychological state, and Oscar and Twin-Oscar know different words, then it follows that they are in different psychological states. Oscar has the psychological state “W<sub>1</sub>’ is associated with intension I’ while Twin-Oscar has the psychological state “W<sub>2</sub>’ is associated with intension I’. And, once we are forced to hold that Oscar and Twin-Oscar are psychologically different, semantic externalism cannot be proven by the Twin-Earth thought experiment. The differences between the chemical composition of ‘water’ on Earth and Twin-Earth might be *one* reason we could propose as to why ‘water’ on Earth and ‘water’ on Twin-Earth mean different things, but we might also propose that they mean different things because Oscar and Twin-Oscar are different psychologically in a relevant way: they know different words.<sup>11</sup> Oscar and Twin-Oscar might still be ignorant that they use the words ‘water’ on Earth and ‘water’ on Twin-Earth to refer to different entities, but this is compatible with various forms of semantic internalism.<sup>12</sup>

Again, I am not defending the view that to know the meaning of a word is just a matter of being in a certain psychological state. My argument here takes no position on whether this is true or not. What I rely on here is not a claim about semantic

<sup>11</sup> Does this make two speakers of a language different psychologically if they know different words, or speakers of English different psychologically to speakers of French? In a simple sense, yes. If knowing a word is part of our psychology, and I do not know words that other speakers (of English or another language) know, then they differ psychologically from me. The importance of such differences may not be significant in most cases, but a reminder here that Putnam requires Oscar and Twin-Oscar to be psychologically the *same*, at least in the relevant ways. My suggestion is that those relevant ways, given the topic under discussion include psychological states involved in ‘knowing a word W’ or ‘associating term ‘X’ with intension ‘I’.

<sup>12</sup> Of course, this difference in referent cannot be used to secure semantic externalism alone without begging the question against semantic internalism.

externalism or internalism, but only that to know a word is a part of a person's psychological state. If Oscar and Twin-Oscar are supposed to be in the same psychological state, then that should mean that they know the same word. I have argued here that Putnam provides no ground for why we should accept that 'water' on Earth and 'water' on Twin-Earth are the same word, thereby undermining his argument. Nor can Putnam respond that 'water' on Earth and 'water' on Twin-Earth are different words, without undermining the crucial premise that Oscar and Twin-Oscar are in the same psychological state.

Second, in ordinary contexts we simply assume some notion of linguistic identity between words. Is Putnam not entitled to the same assumption? That is, perhaps Putnam does not need to provide an argument for why 'water' and 'water' are the same word, just as ordinary speakers need not have answers to the question of when it is that they utter the same word. Perhaps we could even support this because we hold that there simply is no coherent metaphysics of words.<sup>13</sup>

Maybe those who respond in this way are right. Maybe there is no coherent metaphysics of words, but I at least think that there is and the existing debates in philosophy over such matters suggest others do too.<sup>14</sup> I also would agree with this response to a degree in that I do not think that we need a metaphysics of words in order to engage in ordinary conversations, just as we do not need a metaphysics of composition and ordinary objects to think that I am sitting on a chair currently, or a metaphysics of time to understand the claim that I must teach a class tomorrow.

But, insofar as Putnam is wanting to arrive at a conclusion about the nature of meaning, and meaning is a property of linguistic objects, questions about the nature of those linguistic objects can be relevant here. Putnam's conclusion is one within the metaphysics of meaning, and so it would be strange to deny that other metaphysical issues cannot be raised in this context. In ordinary contexts, I agree that we are entitled to assume some notion of linguistic identity across words, but in a philosophical context, it is unclear that we are entitled to that same assumption, especially as we are not so entitled to other identity claims in philosophical (or metaphysical) contexts without argument.

A third response might be that the (or at least a) conclusion of Putnam's argument is that a difference in meaning secures that 'water' and 'water' are instances of different words, contra my suggestion that he needs them to be instances of the same word. Putnam is therefore in fact presenting a case where what appear initially to be instances of the same word are in fact not instances of the same word. But this does not work due to the problems already noted above. If part of Putnam's conclusion is that words are individuated by their meaning, then, as we have noted, we have good independent reasons for thinking that his conclusion is false.

<sup>13</sup> Thanks to a reviewer at this journal for raising this concern.

<sup>14</sup> What about linguistics? Putnam is also known for his suggestion that we should simply ask the experts, so why should we not simply ask linguists what words are? Unfortunately, linguists are just as unclear about what words are as philosophers. For example, Haspelmath (2023) defends only a theory-internal definition, and not a metaphysically substantive one. Such work also does not tackle the distinct question that is at the heart of this paper about what, metaphysically, makes instances of a word, instances of the same word. I am open to linguistic contributions to this, but until such work appears, the task seems to be mostly occupying the minds of philosophers currently.

Appealing to meaning or semantic facts to individuate words might be intuitive, and might align with how we talk about words, but there are very strong reasons to reject the view as a *metaphysical* account of how to individuate words.

Again, we should stress this is *not* to deny that the psychological or conceptual states of the speakers are not sufficient to generate the meaning of a word. Semantic externalism could be true irrespective of whether ‘water’ on Earth and ‘water’ on Twin-Earth are instances of the same or different words. If I am right, this only shows that Putnam cannot respond to the objections I have raised here by holding that ‘water’ on Earth and ‘water’ on Twin-Earth are instances of different words, or by holding that the very difference in meaning of ‘water’ on Earth and ‘water’ on Twin-Earth shows that they are instances of different words.

## V

While Putnam’s thought experiment is perhaps the most influential, there are other (somewhat) similar thought experiments that hope to argue for semantic externalism. Burge (1979), for example, asks us to consider a patient, call them Helen, a competent speaker of English, who has an inflammation in her thigh, and states to her doctor ‘I have arthritis’. This sentence is, of course, false. Arthritis affects joints and cannot affect the thigh. Now, though, consider Twin-Helen, on Twin-Earth, who is, just as Oscar and Twin-Oscar were, a molecule-for-molecule duplicate for Helen. On Twin-Earth, though, when Twin-Helen says, ‘I have arthritis’, what she says is true as on Twin-Earth, the term ‘arthritis’ is used for a condition that can affect the thigh.

Like Putnam, Burge takes this to be an argument for semantic externalism. When Twin-Helen says ‘arthritis’, she says a term that is not extensionally equivalent with ‘arthritis’ as uttered by Helen. Burge holds that this difference cannot be explained by some fact about Helen or Twin-Helen as they are molecule by molecule duplications, and hence their ‘non-intentional, phenomenal experience is the same. [They have] the same pains, visual fields, image, and internal verbal rehearsals. The counterfactuality in the supposition touches only [their] social environment’ (1979:78) and hence ‘the word ‘arthritis’ [on Twin-Earth] does not mean *arthritis*’ (1979: 79, *emphasis in original*).

Does Burge’s argument rely on the same implicit metaphysical assumptions concerning the metaphysics of words as Putnam’s? I argue that it does. Burge himself states that Helen and Twin-Helen say and hear ‘the same words (word forms) at the same times’ (1979: 78). Burge argues that the difference in mental contents is attributable to differences in the social environment, but just as in the case of Putnam, a difference in the meaning is uninteresting if ‘arthritis<sub>1</sub>’ and ‘arthritis<sub>2</sub>’ are not instances of the same word, but Burge provides no clear metaphysics that secures this sameness. We cannot use meaning, phonetics, intentions, origins or any of the means surveyed above to explain why ‘arthritis<sub>1</sub>’ and ‘arthritis<sub>2</sub>’ are instances of the same word (tokens of the same type) for the same reasons as applied before. Burge argues that Helen and Twin-Helen will have all of the same psychological states—the same ‘internal qualitative experiences,’ ‘physiological states and events,’ ‘behaviorally described stimuli and responses,’ and ‘dispositions to behave’ (1979:

79), but we cannot simply assume that they have, as part of their psychology, knowledge of the same word as we cannot simply assume that 'arthritis<sub>1</sub>' and 'arthritis<sub>2</sub>' are instances of the same word. Burge and Putnam's arguments are similar in that both rely on identical psychological states. The difference between them comes from how it is that meanings arise 'outside' of the head as Burge stresses the importance of the social environment of a speaker rather than Putnam's narrower focus on referents. However, both arguments fail due to their undefended reliance on a metaphysical claim concerning the sameness of  $w_1$  and  $w_2$ , and 'arthritis<sub>1</sub>' and 'arthritis<sub>2</sub>'.

## VI

What might we conclude from this discussion? I think there are a few important conclusions that we can draw, but it is important to start with something that we cannot conclude from what I have argued here. We cannot conclude from this that semantic externalism or any other view that has made use of the Twin-Earth thought experiment is false. Nothing here should be interpreted as being an argument against semantic externalism in general. I have also not argued against the Twin-Earth thought experiment if we view it not as a formal argument for semantic externalism, but as more of an intuition pump that aims to more rhetorically persuade us that semantic externalism. Putnam's thought experiment might still be important as presenting a case where the referent of the term seems to be important when considering what the term means. At most, what I have said provides some reasons why we might be sceptical of Putnam's (and Burge's) thought experiment's status as an *argument* for semantic externalism. Semantic externalism, however, might be argued for in any number of other ways.

Despite these limits to my conclusions, I think this is still a very significant conclusion. Of all the arguments that are normally put forward for semantic externalism, Putnam's Twin-Earth thought experiment is amongst the most cited, if not the most cited, and Burge's is similarly popular. Few undergraduate students make it to the end of a philosophy degree without coming across Twin-Oscar and Twin-Earth. But, if I am right, then these thought experiments cannot be used to arrive at semantic externalism as it rests on at best highly controversial, and at worst implausible, assumptions about the metaphysics of words. The only way to secure the words being the same is by appealing to views widely rejected in the relevant literature, such as through appealing to the phonetic form of the words but accepting such views would greatly damage the persuasiveness of the thought experiment.

To close, I want to mention one further broader consequence of my claims in this paper. Putnam's Twin-Earth thought experiment has generated a cottage industry of literature arguing its strengths and weaknesses. If I am right, however, its usefulness is predicated on assumptions about the metaphysics of words. The metaphysics of words is a relatively new domain, and thus far work within it has been quite narrow focusing on purely ontological or metaphysical issues about words.<sup>15</sup> I suggest that

<sup>15</sup> Exceptions are Tarnowski and Glowacki (2022) and Miller (2019).

this case provides an important example of why the metaphysics of words might be of genuine importance for philosophy more broadly. Many philosophical arguments rest on claims about the nature of words. They rely on what words mean, and how we use them. I predict that at least some others of these will also rely on assumptions concerning whether two words are instances of the same word. While it is natural to talk like this, work in the metaphysics of words shows that accounting for the sameness of words is more difficult than it initially appears. This does not, of course, mean that we should think that all philosophical arguments that mention words are now to be thought to rest on controversial metaphysical assumptions. But it does mean that we should look to see if any others do, and we should insist that future arguments that rely on claims about words are not implicitly assuming controversial metaphysical positions about the nature of those words.

JTM MILLER  
UNIVERSITY OF DURHAM, UK  
[james.miller@durham.ac.uk](mailto:james.miller@durham.ac.uk)

**Acknowledgements.** A version of this paper was presented at the joint *Semantics and Philosophy in Europe 12* and *Ontology as Structured by the Interfaces with Semantics 3* conference in October 2023 in Nice. My thanks to the organisers and audience at that conference for incredibly helpful feedback on ideas in this paper. My thanks also to Anna Bortolan and the reviewers for this journal for their constructive discussions and suggestions.

## References

- Bloomfield, L. (1933). *Language*. New York: Henry Holt.
- Bromberger, S. (2011). 'What are Words? Comments on Kaplan (1990), on Hawthorne and Lepore, and on the issue', *The Journal of Philosophy*, 108(9): 486–503.
- Burge, T. (1979). 'Individualism and the Mental', *Midwest Studies in Philosophy*, 4(1), 73–121.
- Cappelen, H. (1999). 'Intentions in Words', *Nous*, 33(1): 92–102.
- Gasparri, L. (2016). 'Originalism about Word Types', *Thought*, 5: 126–33.
- Haspelmath, M. (2023). Defining the word. *Word*, 69(3), 283–297.
- Hawthorne, J. and Lepore, E. (2011). 'On Words', *The Journal of Philosophy*, 108(9): 447–85.
- Hughes, T.J. (2023). 'Possible words: generativity, instantiation, and individuation', *Synthese* 202(207): 1–27.
- Irmak, N. (2019). An Ontology of Words. *Erkenntnis* 84 (5):1139–58.
- Kallestrup, J. (2011). *Semantic Externalism*. New York: Routledge.
- Kripke, S. (1980). *Naming and Necessity: Lectures Given to the Princeton University Philosophy Colloquium*. Cambridge, MA: Harvard University Press.
- Kaplan, D. (1990). Words. *Aristotelian Society Supplementary Volume* 64 (1):93–119.
- Kaplan, D. (2011). Words on words. *Journal of Philosophy* 108 (9):504–29.
- Matsui, T. (2021). Inferentialism and semantic externalism: a neglected debate between Sellars and Putnam. *British Journal for the History of Philosophy* 29 (1):126–45.
- Miller, J.T.M. (2019), 'Natural Name Theory and Linguistics Kinds', *The Journal of Philosophy*, 116(9): 494–508.
- Miller, J.T.M. (2020a). 'The Ontology of Words: Realism, Nominalism, and Eliminativism', *Philosophy Compass*, 15(7): 1–13.
- Miller, J. T. M. (2020b). On the individuation of words. *Inquiry: An Interdisciplinary Journal of Philosophy* 63 (8):875–84.
- Miller, J.T.M. (2021a), A Bundle Theory of Words, *Synthese*. 198: 5731–48.
- Miller, J.T.M. (2021b), Words, Species, and Kinds, *Metaphysics*. 4(1):18–31.

- Miller, J.T.M. (2022), Sameness of Word, *European Journal of Analytic Philosophy*, 18(2): 2–26.
- Miller, J.T.M. (2024), How to Misspell ‘Paris’. *Philosophy*, 99(4): 511–37.
- Millikan, R. G. (1984). *Language, thought, and other biological categories: New foundations for realism*. Cambridge, MA: MIT Press.
- Munroe, W. (2016). ‘Words on Psycholinguistics’, *The Journal of Philosophy*, 113(12): 593–616.
- Nagel, J. (2013). ‘Knowledge as a Mental State’. *Oxford Studies in Epistemology*, 4:275–310.
- Putnam, H. (1975). The meaning of ‘meaning’. *Minnesota Studies in the Philosophy of Science* 7: 131–93.
- Rowlands, M, J. Lau, and M. Deutsch, (2020). “Externalism About the Mind”, *The Stanford Encyclopedia of Philosophy* (Winter 2020 Edition), Edward N. Zalta (ed.)
- Sainsbury, R. M., & Tye, M. (2012). *Seven puzzles of thought and how to solve them: An originalist theory of concepts*. New York, NY: Oxford University Press.
- Stojnić, Una (2021). Just Words: Intentions, Tolerance and Lexical Selection. *Philosophy and Phenomenological Research* 105 (1):3–17.
- Tarnowski, M., and Glowacki, M. (2022). ‘Words on Kripke’s Puzzle’. *Synthese* 200, 292.
- Wetzel, L. (2000). ‘The Trouble with Nominalism’, *Philosophical Studies*, 98: 3617–0.
- Wetzel, L. (2009). *Types and tokens: on abstract objects*. Cambridge: MIT Press.
- Wikforss, Å. (2008). Semantic Externalism and Psychological Externalism. *Philosophy Compass* 3 (1):158–81.