



WITOLD M. HENSEL
UNIVERSITY OF BIALYSTOK

**WATERED DOWN ESSENCES AND ELUSIVE SPEECH
COMMUNITIES: TWO OBJECTIONS AGAINST PUTNAM'S
TWIN EARTH ARGUMENT**

Hilary Putnam (1975) famously contended that the extension of many linguistic expressions is underdetermined by speakers' psychological states taken in their narrow sense, or individuated on the assumption that no psychological state presupposes the existence of any object other than the subject of that state (methodological solipsism).¹ Putnam supported this claim by offering a series of Twin Earth thought experiments and appealing to the phenomenon known as the division of linguistic labor.

In this paper, I focus exclusively on the argument from Twin Earth. I claim that it rests on two assumptions and that these assumptions are highly contentious. Given that both assumptions must be accepted for the argument to work, the argument fails. The first assumption has to do with what Putnam called the logic of natural-kind terms, the second with the notion of a speech community. In what follows, I offer a brief description of Putnam's argument and then focus on each assumption in turn.

"Water" on Twin Earth

The story is familiar. Let there be Twin Earth, a planet that is exactly like Earth in every respect but one: the liquid filling the rivers and lakes on Twin Earth, though (almost) indistinguishable from H₂O, has the chemical composition XYZ. According to Putnam, if a spaceship from Earth visited Twin Earth and its crew discovered the difference between the two planets, the message they would send home would read:

¹ I dispense with the adjective "narrow" in the rest of the paper. The term "psychological state" will henceforth denote narrow psychological states.

On Twin Earth, "water" means XYZ.

The same statement would have been true in 1750, when scientists on either planet were unable to distinguish XYZ from H₂O. Therefore, in 1750, the sentence "Water is tasteless" would have been about H₂O, when uttered by Oscar, and about XYZ, when uttered by Twin Oscar, even if Oscar and Twin Oscar were in the same psychological state (their brains had exactly the same microstructure).

It follows, Putnam says, that knowing the meaning of "water" is not only a matter of being in the appropriate psychological state. Knowing the meaning of "water" also involves having the right kind of causal connections to the right sort of stuff in the world – in this case, to samples of H₂O for Oscar and samples of XYZ for Twin Oscar.

Natural kinds, indexicality and the *qua* problem

The Twin Earth thought experiment would not have been as persuasive as it was with any old word used in place of "water". It could not possibly work with the word "bachelor". And, presumably, a story about Twin pencils, with cores made of some mysterious substance rather than graphite, would not have evoked the response "On Twin Earth, 'pencil' means something else than on Earth", even at a time when *all* pencils on Earth had graphite fillings.

According to Putnam, there is a large class of words, which he calls natural-kind terms, that play an important role in explanations. Natural-kind terms include names of substances, physical magnitudes, animals and plants, as opposed to names of artefacts and other socially constructed objects, such as jobs. They are taken to feature in many inductive generalizations and lawlike statements.

Putnam maintains that they also display a special kind of logic. Namely, invoking a natural kind implies an appeal to a shared (and typically hidden) nature that accounts for manifest characteristics of the kind's members. Thus, any ostensive definition of a natural-kind term carries with it a defeasible empirical presupposition that the indicated sample of the term's extension bears a same-kind relation to most of the stuff to which the term has been applied on other occasions.

Putnam's discussion is somewhat confusing, though. An ostensive definition of any general term carries with it a defeasible presupposition of the type mentioned above. If I say "This kind of writing implement is called a pencil" and point to a pen that merely looks like a pencil, my

definition will not be valid, precisely because the pen I have indicated does not bear the appropriate same-kind relation to objects most members of my language community call pencils. Furthermore, one can define practically any non-empty general term *via* ostension: this is a bachelor, that is a bachelor, etc.

What is relevant to the Twin Earth argument is only that some general terms, including natural-kind names, such as “water”, have, as Putnam put it:

an unnoticed indexical component: “water” is stuff that bears a certain similarity relation to the water *around here*. Water at another time or in another place or even in another possible world has to bear the relation *same_L* to our “water” in order to be water. (Putnam, 1975, p. 152)

Therefore, what Putnam calls the logic of natural-kind terms should really be called the logic of indexicality. This is because, on Putnam’s account, some names of artefacts have an indexical component, whereas some natural-kind terms *do not*. As to the former, recall Putnam’s discussion of Rogers Albritton’s live pencils example: If we discovered that pencils on Twin Earth were organisms, we would refrain from calling them pencils, unless of course we discovered that pencils on Earth were also alive (Putnam, 1975, pp. 161-162). “A Mercedes” would be a less controversial illustration.

There are two distinct reasons why it is not the case that all natural-kind terms are indexical on Putnam’s view. First, some names of natural kinds, such as “sand” and “air”, do not seem to presuppose any particular nature shared by all their referents. This is so, even though we standardly explain the referents’ manifest qualities by appealing to microstructure. Second, whether or not a word exhibits indexicality depends on language users. Indeed, according to Putnam, some natural-kind words that begin their career as equivalent to clusters of descriptions can subsequently become indexical (and, presumably, the other way around). For all we know, “water” may have initially meant something like “colorless, tasteless, odorless liquid that quenches thirst”.

To complicate things further, Putnam believes that many different senses of the word “water” coexist and many of those senses are indexical. This is because “*X* bears the relation *same_L* to *y* just in case (1) *x* and *y* are both liquids, and (2) *x* and *y* agree in important physical properties. . . . Importance is an interest-relative notion” (Putnam, 1975,

p. 157). Presumably, the fact that structural properties are “normally” considered to be important implies that, in its *core* sense, “water” denotes H₂O.² In some senses, samples of H₂O with impurities are water, in others not; in some senses, ice counts as water, in others not, etc. These differences in extension result from interest-relativity.

What is the relationship between the Twin Earth thought experiments and Putnam's account of indexicality of some natural-kind terms? I suggest that the experiments are taken to confirm the theory, though, of course, cannot establish its truth. Our referential judgments are regarded as the theory's explananda, and indexicality is supposed to account for them. This raises two kinds of questions: about the existence of the target phenomena (e.g., do our referential dispositions comport with the theory's predictions?) and about the theory's ability to explain them, if they exist. In this part of the paper, I focus on the latter kind of questions, so I temporarily grant that our responses to Twin Earth scenarios agree with Putnam's.

There are three major difficulties here. The first is that the world may not have the natural-kind structure required by Putnam's account. Secondly, even if the world has the appropriate natural-kind structure, there are an indefinite number of widely differing, mutually exclusive and equally intuitive construals of reference transmission, the choice of which profoundly affects extension. In other words, in light of the first two problems, Putnam's theory may fail to explain the target phenomena. Thirdly, there may be an account of meaning that explains the target phenomena at least as well as Putnam's account, but in an internalist way. In other words, Putnam's account may not be the best explanation of the target phenomena. Although all these worries are equally important, I will restrict attention to the first.

In a recent paper, Sören Häggqvist and Åsa Wikforss (forthcoming) argue persuasively that the Kripke-Putnam account of natural-kind terms relies on microessentialism, a view according to which objects or samples of substances falling under a single natural kind all share a common microstructure that explains their macroscopic properties and is necessary throughout modal space. But microessentialism, they contend, is at odds with our best philosophy of science. Therefore, the Kripke-Putnam thesis, which asserts that the

² Putnam is silent on what makes a particular sense *the core sense*.

extension of natural-kind terms is determined by the microessences of stuff present in speakers' environment, is incorrect.

The core of Häggqvist and Wikforss' argument is well-known to anyone familiar with the so-called *qua* problem and contemporary philosophy of science. Typically, there is no single structure underlying the properties of an object or substance; instead, there are numerous structures that do not fit very well with the natural kinds suggested by common sense. This is especially clear in biology. As Häggqvist and Wikforss poignantly remark, Devitt (2008) is the only author in the philosophy of biology who clings to a form of essentialism. Similar conclusions are being reached by philosophers of chemistry about chemical kinds (see, e.g., Needham, 2000, and Hendry, 2005).

Following Häggqvist and Wikforss, let me summarize what science tells us about the nature of water. First of all, the formula H_2O does not capture the structure of water, but rather its chemical composition, or, in this case, molar proportions. This is an important distinction, because structural isomerism implies that different substances may share a single composition. For example, propanol, isopropanol, or methoxyethane are all C_3H_8O (and large molecules of organic compounds have millions of isomers). Generally, then, chemical composition is not a good candidate for a substance essence – we need to dig deeper.³ The essentialist may respond by invoking molecular structure: surely, she will say, water is composed of H-O-H molecules, isn't it? Well, not quite. As Häggqvist and Wikforss point out, water is not usually molecular. Liquid water is composed of H^+ and OH^- ions as well as H-O-H molecules, all of which are in constant flux, forming polymers of different lengths at rates that vary with temperature and pressure. On this level of description, then, liquid water has an immense number of structures. And Häggqvist and Wikforss have barely scratched the surface. They haven't broached the subject of heavy, semi-heavy, heavy-oxygen or tritiated water. Nor have they mentioned the fascinating complexities of water in its other states, including different varieties of amorphous ice (LDA, HDA, VHDA).

³ Water happens not to have isomeric structure, but since other substances do, we should probably look for chemical essences at a lower level of organization than that of chemical composition.

If it is so hard to pinpoint anything even remotely resembling *the* microessence of water, then it should come as no surprise that similar difficulties arise, in much greater numbers, when we turn to biological kinds. It is, I think, no exaggeration to say that modern biology is a thoroughly anti-essentialist science. This anti-essentialism is clearly reflected in the philosophical literature devoted to the life sciences, so I am not going to dwell on it here. Instead, let me recall briefly an interesting study by Andrew Shtulman and Laura Schultz (2008), which suggests that naïve essentialist beliefs about biological species seriously impede people's ability to understand the principles of Darwin's theory of evolution through natural selection. Given the impact of Darwin's theory on contemporary biology, it is no wonder that biologically-informed researchers are so vehement in rejecting essentialism.

Häggqvist and Wikforss's criticism would be potentially devastating against a conception of natural kinds that took singularity of common structure to be necessary for natural kindness and indexicality. Putnam's theory is not that sort of theory, however. Here is a quote that confirms this: "But the local water, or whatever, may have two or more hidden structures – or so many that 'hidden structure' becomes irrelevant, and superficial characteristics become the decisive ones" (Putnam, 1975, p. 1961). The same idea appears in a recent defense of externalism by Daniel Korman (2016), who formulates the following "default conditionals" that are supposed to govern the semantics of "water":

(i) If water turns out to be compositionally uniform, then "water" expresses a concept that applies to all and only samples of that compositional kind with respect to all counterfactual situations.

(ii) If water turns out to have a highly disuniform composition, then "water" expresses a concept that applies to all and only samples of superficially water-like kinds with respect to all counterfactual situations. (Korman, 2016, p. 507)

Korman's default conditionals are useful, because they wear their shortcomings on their sleeve. First, they are glaringly incomplete. We need at least one more default conditional to handle kinds that are neither uniform in composition nor highly disuniform. Second, the term "highly disuniform" is vague.

We can take care of the first problem by following Putnam, who maintained that the extension of terms such as “jade” is determined by a disjunction of two hidden structures (there are two kinds of jade, he claimed). In fact, we have no other option, since given the apparent lack of natural kinds unified by a single microessence, any other move would amount to adopting internalism rather than externalism.

Now, the vagueness of Korman's phrase “highly disuniform” is a different issue. As I see it, externalists have only two options available to them. The easy way out would be to use the standard method of dealing with vagueness: i.e., draw the boundary between highly and non-highly disuniform kinds in an arbitrary manner. Say, at three or seven, to stick only to magic numbers. This would make the dispute between externalists and internalists a partly conventional and partly empirical disagreement, a matter of decision as well as of fact. Although I do not pretend to have insights into the ultimate nature of reality, I would not bet on the world turning out the way externalists expect it to be. So far, increasing scientific progress has been associated with ever more discoveries of new structures (see Taylor, Vickers, 2017, for an overview of the phenomenon of conceptual fragmentation in science).

The other option is to maintain, like Putnam, that the boundary separating highly disuniform kinds from merely disuniform ones is delineated in light of our interests. Generally, then, the picture Putnam is proposing is that the hidden structures that determine the extension of natural-kind terms are always filtered by our interests. Unless we are talking infinities, the number of structures relevant to explaining what we want to explain in light of a set of interests is bound to be smaller than the number of hidden structures listed in a long unsorted disjunction. Moreover, it is arguably an empirical issue (though in a pretty broad sense of the word “empirical”) whether a set of structures is relevant in light of a particular set of interests.

While adding interests to the mix is a step in the right direction, I do not think it will save externalism. Indeed, it will only exacerbate the *qua* problem.

As far as the determination of extension of natural-kind terms is concerned, interests enter the equation in at least two places: when the speaker (or group of speakers) chooses which characteristics of a kind of substance or object need explaining, and when the speaker (or group of

speakers) decides which kind of explanations are acceptable. Needless to say, this is a grossly oversimplified picture of what is really going on.

To illustrate: When you observe water, you regard some of its superficial characteristics as more important than others. For example, you may want to know why water is tasteless or how it is that fish and other animals can live in it, but, at the same time, you can remain unmoved by the fact that water solidifies and increases in volume when the temperature drops below zero. Indeed, most superficial characteristics of water will probably escape your attention altogether. The choice, to repeat, depends on your interests, broadly construed. But, once you have selected which properties of water to account for, your interests will also affect what type of explanations you will pursue and accept. For example, you may prefer functional explanations to mechanistic ones (see Lambrozo, Gwynne, 2014); or you may opt for observation rather than experimentation, because it is cheaper. Although choice of explanans and choice of explanandum are often interdependent, it is reasonable to keep them separate here.

Let me use a toy example to flesh this out. Imagine that you are walking through a jungle and encounter an unfamiliar object or substance – a tree, a shrub, an insect, a mammal, or some malodorous slime oozing from a rock. You study it for a while and decide to give it a general name. You say to yourself “I will call this kind of slime ‘shlaw””. You put some of the stuff into a bucket, take it to your village and show it to the shaman, who is visibly excited. Suppose that five superficial characteristics of shlaw become important to people from your village and ten more, though remarked upon, have been largely ignored.

The question to address is this: How have your personal interests and the interests of your community constrained the choice of hidden structures relevant to determining the extension of “shlaw”? Answer: It is hard to say, but probably not very much. First, the characteristics of shlaw that are of interest to you or your community will probably be poorly defined (What exactly *is* slime? What did you mean by “malodorous”?) and thereby amenable to a wide range of theoretical interpretations. This means that, more often than not, they will be discovered by future science to be *clusters* of properties rather than properties *per se*. Second, they will be diverse: each characteristic will most likely be explainable in terms of a different set of hidden structures. And, third, there will typically be a large number of explanatory

approaches acceptable in your community, with each theoretical perspective potentially picking out, via specific idealizations, slightly different structures as explanantia.

Generally speaking, it would seem that knowing the interests of a speech community can help us to identify structures relevant to determining the extension of natural-kind terms only if the community in question is scientifically advanced, for only in such communities can we expect the properties to be well defined and the interests to be sufficiently well articulated. But this is an illusion. The real trouble with interests is that they shift over time, even in scientifically advanced communities. Worse still, these changes of interests are completely unpredictable.

With that in mind, let us try to find out what the word "water" in its "core" sense might denote nowadays. Suppose our present interests and technological development, together with the world, succeed in determining a small set of microstructures underlying the superficial characteristics of water. Can we justifiably maintain that having one of these microstructures is constitutive of water? If so, then what are we going to say when our interests shift, our technology changes and, as a result, a different set of microstructures becomes the most plausible candidate for the nature of water? And if not, then how else should the nature of water be determined?

Given the changeability of interests over time, we can decide that the extension of "water", in its core sense, is determined by the world together with: (a) the interests of our ancestors who first used the word "water" indexically, or (b) our contemporary interests, or, indeed, (c) our future interests – say, the interests of the last generation of our speech community.

Option (a) is implausible, because, as I have already observed, our ancestors' interests were probably too poorly articulated to pick out a sufficiently small number of microstructures. Moreover, there is probably no way of discovering who those ancestors were, what interests they had, and how they used the word "water". And, last but not least, it would be impractical for us to adopt a notion of water that did not harmonize with our present interests.

Option (b) has the obvious advantage of harmonizing with our current interests. However, it does not really bring us much closer to solving the *qua* problem than do accounts of reference that make no

appeal to interests. As things now stand, there are simply too many candidate microstructures to choose from, even if we bring current interests to bear on the choice. Another problem with option (b) is that it is almost indistinguishable from descriptivism, as it practically amounts to asserting that water is identical to whatever satisfies our best current theory. And, just like descriptivism, it also fails to stabilize reference over time: assuming (b), the extension of “water” *has* likely changed since 1750.

Just like (a), option (c) blatantly ignores our present interests. And do we really want the extension of our natural-kind terms to get fixed by something in the future? Moreover, as far as I can see, option (c) can help us solve the *qua* problem only if we adopt a curious form of convergent realism. The convergent realism I have in mind asserts that science will eventually reduce rather than expand the set of microessences plausibly associated with the word “water” (and other natural-kind terms). This, as I remarked earlier, runs counter to the inductive record. Therefore, choosing option (c) would, in most probability, only exacerbate the *qua* problem.

Unfortunately, options (a-c) do not exhaust the possibilities. Not by a long shot. There are indefinitely many accounts we may explore, and many of them would be more appealing than options (a-c) discussed above. But, while I like churning out complex speculative theories as much as the next guy, I will spare myself and the reader the tedium of considering a host of increasingly nuanced accounts of reference. Instead, I will jump right ahead to the conclusions.

Note that options (a-c) are all unsatisfactory, because each tethers the extension of “water” to an arbitrary point in time and thereby imposes unwarranted constraints on acceptable microessences. Option (a) is overly conservative: if our interests are incompatible with the interests of our ancestors, the extension of “water” will probably differ from what we currently take it to be (it is also utterly insensitive to the progress of science). Option (b) is biased in favor of the present and blind to future scientific, technological and social developments. Option (c) anchors the extension of natural-kind terms at the random moment when our speech community will cease to exist. Readers who enjoy apocalyptic books and movies can immediately see the fault in that: what if our civilization collapses and its few survivors, though still speaking English, die out after living for three generations in Dark Age conditions?

All this implies that a plausible account of extension fixing for natural-kind terms must probably involve expanding the set of candidate microstructures associated with options (a), (b) or (c) rather than reducing it.

Let me illustrate this by considering an improvement on option (a). Suppose, for the sake of argument, that our ancestors who first introduced the word "water" as a hidden indexical had interests I_1 , whereas, at present, we have interests I_2 . Assume also that the set of microstructures constitutive of water as determined by I_1 and the set of microstructures constitutive of water as determined by I_2 have no common element. This means that the extension of "water" as determined by (a) is out of step with current use.

We can remedy this by positing that the extension of "water", though fixed in the past, includes the endpoints of all metaphysically possible trajectories of knowledge development as jointly determined by the natural-kind structure of the world and all possible combinations of human interests. Although much more plausible than option (a), this account yields an indeterminately large number of microessences. However, because the account's plausibility depends on the supposition that it cannot exclude any reasonably acceptable microessences, any credible account of extension fixing for natural-kind terms must satisfy the same desideratum.

To summarize: Given what we know about science, the number of microstructures that can explain the superficial properties of objects or stuff falling under a single natural kind is probably too large to determine the extension of any natural-kind term. It is so, even if we specify the same-kind relation by appeal to interests.

The notion of a speech community

Microessentialism is not the only controversial presupposition of Putnam's Twin Earth thought experiments. A second, though frequently unnoticed, assumption that is involved has to do with the notion of a speech community. This sociolinguistic aspect of Putnam's reasoning was first brought out by Eddy Zemach (1976).

Zemach observed that Putnam's externalist formulation of the imagined report sent from the spaceship back to Earth relies crucially on how speech communities are individuated. If we are liberal and accept

the English speaking inhabitants of Twin Earth as members of our speech community, then the message should read:

We have discovered that there are two kinds of water: H₂O and XYZ.

Since, in Putnam's story, there is exactly as much XYZ as there is H₂O, "water" should presumably refer to a disjunction of H₂O and XYZ. Putnam's description of the situation makes sense only if Twin Oscar does not belong to the same speech community as Oscar.

But what possible reason could we have for excluding Twin Earthians from our speech community, given that *ex hypothesi* the only thing distinguishing them from us is that they happen to inhabit a slightly different environment? Can Putnam mark the distinction between speech communities without begging the question against internalists and excluding Australians, South Africans, or the English?

Zemach is skeptical. Attempts to define language in terms of a speech community are frequently circular, because, more often than not, a speech community is itself characterized as a group of people who speak the same language (see Wardhaugh, 2006). Zemach's worry, then, is that Twin Earthians belong to a different speech community than Earthians, because they speak a different language, and we know that they speak a different language because the word "water" applies to XYZ in Twin English and to H₂O in English.

Zemach's worry is justified. Putnam does not offer any reasons why we should respond to the imagined discovery of XYZ on Twin Earth by saying "The word 'water' on Twin Earth means XYZ" rather than by saying "There are two kinds of water". And the differences between American English and Australian English are both more numerous and more linguistically significant than the alleged difference between English and Twin English. In fact, as Zemach suggests, the idiolects of Oscar and Twin Oscar are probably more similar to one another than the idiolects of Hilary Putnam and any other speaker of American English.

Putnam, however, is not committed to admitting Australians, South Africans, or the English into his speech community. He merely needs to specify a non-question-begging, intuitive method of *excluding* Twin Earthians. Such a method seems available.

It is no profound insight that people belong to speech communities by virtue of communicating with each other using language, among other things. We can exploit this observation to formulate a necessary condition for membership in a speech community:

if a person belongs to a speech community, she must have communicated via language with another member of that speech community. Consequently, two completely isolated groups of people cannot form a single speech community.

The proposed necessary condition does not appeal to the notion of extension or to a particular notion of language and so it is not open to the charge of question-begging. Another advantage is that it may well be intuitive. If it is, then we should expect our spontaneous judgments about the extension of natural-kind terms to vary according to the extent of posited verbal interactions between speakers. As an exercise, consider the following two variations on Putnam's original story:

(1) In a galaxy far, far away, there is a planet that is almost exactly like Earth. It is inhabited by people that look like exact atom-for-atom replicas of us, but the distance between Earth and the galaxy far, far away is so great that it precludes any causal interaction of the sort necessary for copying. The only difference between Earth and its twin, call it Twin Earth, is that the liquid that fills the rivers, lakes, and seas on Twin Earth, though phenomenologically indistinguishable from H₂O, has the molecular composition expressed by the chemical formula XYZ. Assuming the story is true, does the word "water" (a) refer to H₂O in English and to XYZ in Twin English, or (b) are there two kinds of water, i.e. "water" means H₂O or XYZ?

(2) In a nearby galaxy a long time ago, there was a planet that was almost exactly like Earth. As a result of a cosmic coincidence, it was even inhabited by people that looked like exact atom-for-atom replicas of us. The only difference between Earth and its twin, call it Twin Earth, was that the liquid that filled the rivers, lakes, and seas on Twin Earth, though phenomenologically indistinguishable from H₂O, had the molecular composition expressed by the chemical formula XYZ. About a thousand years ago, a spaceship from Earth visited Twin Earth, and, at the same time, a spaceship from Twin Earth visited Earth. Having discovered each other's *Doppelgängers*, Earthians and Twin Earthians began travelling back and forth, talking on the radio, writing letters, etc. Only recently, and to their great astonishment,

have they discovered the difference between the two planets. Assuming the story is true, does the word "water" (a) refer to H₂O in English and to XYZ in Twin English, or (b) are there simply two kinds of water, i.e. "water" means H₂O or XYZ?

Your answers will count as evidence for the intuitiveness of Putnam's externalism if your confidence in (a) is noticeably higher in scenario (1) than in scenario (2). I confess that my own responses agree with Putnam's.

But even if most readers' answers to scenarios (1) and (2) agreed with Putnam's and mine, this would merely establish that we seem to share some beliefs about speech communities. It would not secure the stronger conclusion that the beliefs in question are intuitive in any interesting sense of the word. Since intuitive beliefs are standardly construed as strongly influenced by our biological makeup, it is useful to think of intuitions as deeply engrained domain-specific assumptions about the world. What makes these assumptions interesting is the strong causal connection between their etiology and our evolutionary history. This is why intuitions are taken to be practically universal across cultures.

It is doubtful, however, that there is a folk theory that relies on an intuitive notion of a speech community. We cannot claim the same kind of familiarity with the inner workings of speech communities as we can with the behavior of water, animals, and individual people. We hardly make – and we practically never *have to* make – any inductive generalizations involving speech communities in everyday life. Consequently, it is very difficult to identify any relevant features shared by speech communities besides the necessary condition proposed above. Lastly, unlike the words "water" and "animal", the phrase "speech community" has a distinctly theoretical ring to it. In sum, Noam Chomsky's (2000, p. 148) famous criticism of Putnam's account applies directly to the notion of a speech community:

We can have no intuitions about the question, because the terms *extension*, *reference*, *true of*, *denote*, and others related to them are technical innovations, which mean exactly what their inventors tell us they mean: it would make as little sense to explore our intuitions about tensors and undecidability, in the technical sense.

This suspicion is borne out by how the term “speech community” has been used in empirical linguistics (see Patrick, 2002, and Wardhaugh, 2006, pp. 119-132). The first modern definitions of “speech community” appealed primarily to uniformity of linguistic behavior. For example, Bloomfield (1933) explicitly assumes that “within certain communities successive utterances are alike or partly alike”, and then adds, by way of a definition, that “any such community is a speech community” (Bloomfield, 1933, pp. 153-154). Bloomfield’s emphasis on uniformity is later echoed in Chomsky’s “ideal speaker-listener, in a completely homogenous speech community (Chomsky, 1965, p. 3; see also Chambers, 1980). Interestingly, according to Patrick (2003), Bloomfield explains both external boundaries and internal variation in terms of speaker interactions: “a speech-community is a group of people who interact by means of speech” (Bloomfield, 1933, p. 42) and “differences of speech within a community are due to differences in density of communication” (Bloomfield, 1933, p. 46). In other words, on Bloomfield’s view, Twin Oscar would belong to the same speech community as Oscar and the similarities between their utterances would remain inexplicable. Needless to say, Chomsky and his followers would whole-heartedly agree.

The advent of sociolinguistics, ushered in by William Labov’s presentation at the 1962 annual meeting of the Linguistic Society of America (see Chambers, 2002, p. 5), did not change much when it came to the theorists’ lack of reliance on speaker interaction. Although some sociolinguistic approaches appear to be consistent with Putnam’s externalist account, because they either retain the interaction condition (Gumperz, 1968)⁴ or characterize speech communities in terms of geographic location, most accounts in the field appeal to criteria that have nothing to do with density of communication. Perhaps the most influential such criterion invokes shared norms of utterance production and evaluation. For example, Labov writes (1972, p. 120-121):

The speech community is not defined by any marked agreement in the use of language elements, so much as by participation in a set of shared norms. These norms may be observed in overt types of evaluative behavior, and by the uniformity of abstract patterns of variation which are invariant in respect to particular levels of usage.

⁴ See the next paragraph.

A striking feature of the many notions of a speech community explored in sociolinguistics is that none of them unequivocally classifies Oscar and Twin Oscar as members of different speech communities. Even notions closest to Putnam's proposal fail to do so. For example, Gumperz's definition, according to which a speech community is "any human aggregate *characterized by regular and frequent interaction* by means of a shared body of verbal signs and set off from similar aggregates by significant differences in language usage" (Gumperz, 1968, p. 381, emphasis added) either treats Oscar and Twin Oscar as members of the same speech community or, at best, leaves the matter unsettled, because Americans do not interact with Twin Americans, and yet the utterances made by Oscar and those made by Twin Oscar are linguistically indistinguishable.

In fact, even the most noncommittal definition of a speech community I know of, which stipulates that a speech community is merely "some kind of a social group whose speech characteristics are of interest and can be described in a coherent manner" (Wardhaugh, 2006, p. 119), would arguably be of no use to Putnam, because Oscar and Twin Oscar share all verbal dispositions, being linguistically indistinguishable from each other.

Of course, it is fairly easy to modify Putnam's story so that Twin English becomes a distinct language from English. Just introduce a sufficient number of differences in pronunciation and perhaps syntax. But there are two problems with this move. First, it is now unclear whether Oscar and Twin Oscar are in the same psychological state, because, at the very least, their brains are no longer identical. Second, the string of words represented as "Oscar would like a glass of water" should not count as a single sentence, but rather as two: one in English and one in Twin English. Indeed, if Putnam insists that Oscar and Twin Oscar belong to two different speech communities, the most accurate report sent back to Earth should read:

People on Twin Earth use a word that sounds like the English word "water" and applies to a substance that looks and behaves exactly like water, but its chemical composition is XYZ.

Yet the fact that two different words may denote two different substances is old news. It is hardly a profound insight that the word "water" means water, whereas the word "fire" means fire.

Externalists can shift gears, however, and insist that the Twin Earth story yields a desired conclusion concerning words, not meanings. The conclusion would be that knowledge of words does not supervene on psychological state, because Oscar and Twin Oscar, though psychologically indistinguishable, know different words. Alas, this argument faces a similar problem to the previous one. Namely, in order for the externalist conclusion to follow, Oscar and Twin Oscar must be in different brain states if one is to belong to a different speech community than the other. And although the externalist can reply that people who are in two different brain states may well be in the same psychological state, she will need an additional argument for the psychological irrelevance of neurological properties underlying linguistic differences.

Externalists can also point out that sociolinguistics is not a mature field, and it has not yet produced a good enough notion of a speech community. It is therefore possible that a mature sociolinguistic theory will recognize Twin Oscar as belonging to a different speech community than Oscar. But this misses the point. For, regardless of how sociolinguistics will develop, the important thing is that the notion of a speech community is the kind of concept that is shaped by the investigator's interests. As Patrick puts it (2002, p. 593):

we ought not to assume SpComs [speech communities – W.M.H.] exist as predefined entities waiting to be researched or identify them with folk notions, but see them as objects constituted anew by the researcher's gaze and the questions we ask.

Ultimately, then, both externalism and internalism are viable positions in so far as their choice is informed by the researcher's interests. If, however, we choose not to ignore current scientific practice when assessing philosophical positions, then externalism appears to be the less plausible alternative.

REFERENCES

- Chambers, J. K. (1980). Linguistic Variation and Chomsky's "Homogenous Speech Community". In M. Kinloch, A. B. House (eds.), *Papers from the Fourth Annual Meeting of the Atlantic Provinces Linguistic Association* (1-32). Fredericton: University of New Brunswick.
- Chambers, J. K. (2002). Studying Language Variation: An Informal Epistemology. In J. K. Chambers, P. Trudgill, N. Schilling-Estes (eds.), *The Handbook of Language Variation and Change* (3-13). Oxford: Blackwell (1st edition).
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (2000). *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.
- Devitt, M. (2008). Resurrecting Biological Essentialism. *Philosophy of Science*, 75, 344-382.
- Genone, J., Lambrozo, T. (2012). Concept Possession, Experimental Semantics, and Hybrid Theories of Reference. *Philosophical Psychology*, 25 (5), 714-741.
- Gumperz, J. J. (1968). The Speech Community. In D. L. Sills (ed.), *International Encyclopedia of Social Sciences* (381-386). New York: Macmillan.
- Häggqvist, S., Wikforss, Å. (forthcoming). Natural Kinds and Natural Kind Terms: Myth and Reality. *The British Journal for the Philosophy of Science*.
- Hendry, R. F. (2005). Lavoisier and Mendeleev on the elements. *Foundations of Chemistry*, 7, 31-48.
- Korman, D. Z. (2016). What Externalists Should Say About Dry Earth? *The Journal of Philosophy*, 103(10), 503-520.
- Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Lambrozo, T., Gwynne, N. Z. (2014). Explanation and Inference: Mechanistic and Functional Explanations Guide Property Generalization. *Frontiers in Human Neuroscience*, 8, Article 7.
- Needham, P. (2000). What is Water? *Analysis*, 60(265), 13-21.

- Patrick, P. L. (2003). The Speech Community. In J. K. Chambers, P. Trudgill, N. Schilling-Estes (eds.), *The Handbook of Language Variation and Change* (573-597). Oxford: Blackwell (1st edition).
- Putnam, H. (1975). The Meaning of "Meaning". In K. Gunderson (ed.), *Language, Mind, and Knowledge* (131-193). *Minnesota Studies in the Philosophy of Science*, vol. VII. Minneapolis: University of Minnesota Press.
- Shtulman, A., Schulz, L. (2008). The Relation Between Essentialist Beliefs and Evolutionary Reasoning. *Cognitive Science*, 32, 1049-1062.
- Taylor, H., Vickers, P. (2017). Conceptual Fragmentation and the Rise of Eliminativism. *European Journal for the Philosophy of Science*, 7(1), 17-40.
- Wardhaugh, R. (2006). *An Introduction to Sociolinguistics*. Oxford: Blackwell (5th edition).
- Zemach, E. (1976). Putnam's Theory on the Reference of Substance Terms. *The Journal of Philosophy*, 73 (5), 116-127.

ABSTRACT

WATERED DOWN ESSENCES AND ELUSIVE SPEECH COMMUNITIES: TWO OBJECTIONS AGAINST PUTNAM'S TWIN EARTH ARGUMENT

The paper presents two objections against Putnam's Twin Earth argument, which was intended to secure semantic externalism. I first claim that Putnam's reasoning rests on two assumptions and then try to show why these assumptions are contentious. The first objection is that, given what we know about science, it is unlikely that there are any natural-kind terms whose extension is codetermined by a small set of microstructures required by Putnam's indexical account of extension determination. The second objection is that there may not be a plausible concept of a speech community whose adoption would classify Oscar and Twin Oscar as members of different speech communities and, at the same time, render Oscar and Twin Oscar as being in the same psychological state. I contend that Putnam's argument fails because both objections are justified.

KEYWORDS: externalism; Twin Earth argument; natural-kind terms; qua problem; interest relativity; speech community