UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di FILOSOFIA, SOCIOLOGIA, PSICOLOGIA e PEDAGOGIA APPLICATA (FISPPA)

CORSO DI DOTTORATO DI RICERCA IN: FILOSOFIA
CICLO: XXIX

ON THE SEMANTICS OF ARTIFACTUAL AND SOCIAL KIND TERMS

Tesi redatta con il contributo finanziario della FONDAZIONE CARIPARO

Coordinatore: Ch.ma Prof.ssa Francesca Menegoni
Supervisore: Ch.mo Prof. Massimiliano Carrara

Dottorando : Irene Olivero

Index

Introduction

Abstract. Hilary Putnam's proposal of extending the scope of his famous externalist semantic theory to artifactual (e.g. 'pencil', 'chair', 'television', etc.) and social kind terms (e.g. 'pediatrician', 'university', 'money', etc.) has opened an ongoing debate, which is the main focus and basis of the present work. By contrast, the same semantic account limited to words for natural substances and species (e.g. 'water', 'gold', 'tiger', etc.) has become very popular and quite widely adopted, since it seems to give a convincing explanation of how the semantics of these terms works. The first part of the present work is dedicated to a twofold purpose: on the one hand, I aim to argue that Putnam's argument about the semantics of artifactual (and social) kind terms fails in its goal; on the other hand, – on the strength of the considerations drawn out from this analysis – I point out that the main positions which characterize the aforementioned debate do not succeed in their intent either. In the second part, I focus on remarking that the criticisms of Externalism arise already even when accounting for some of usually believed steady cases. I aim to argue, in this respect, that all problems of Putnam's Semantic Externalism – as applied not only to artifactual and social kind words, but across the board – are a consequence of the fact that Putnam implicitly relies on a hyperrealist view on modality. Such a modal approach notoriously encounters several problems. In order to bypass those problems, Amie Thomasson proposes an alternative approach to modality: Modal Normativism. I aim to highlight that adopting a normativist approach commits us to abandon Putnam's Externalism and to embrace a hybrid theory of reference. Such view on reference combined with Modal Normativism – I intend to show – not only overcomes the semantic weaknesses of Putnam's Externalism, but it also accounts for those controversial cases emerged within the debate, as much as by the opponents of Putnam's semantic view in general. Such alternative approaches combined constitute – I therefore argue – a more tenable and attractive account of the semantics of our kind terms. Within the perspective I here advocate for the debates about artifactual and social kind terms, as much as in general about the controversial cases Externalism has difficulties to explain, can be seen – I suggest – as metalinguistic negotiations. With such a notion – advanced by Tim Sundell and David Plunkett and taken up by Thomasson – we gain the further advantage of accounting for the importance of the debate while still preserving the advantages of a deflationary approach.

The bottom line of my argumentation is based on the intention of highlighting the strict interrelation between semantics and the metaphysics of modality, which motivates my reasoned critique of Putnam's argument about words for artifactual (and social) kinds, as well as the alternative approach I propose to explain the semantics of all kind terms.

"Cut the pie any way you like, 'meanings' just ain't in the *head*!" (Putnam 1975, p. 227)

This is the slogan of a famous semantic theory, nowadays known as *Externalism*, whose two main exponents are Saul Kripke and Hilary Putnam. Such a theory was meant to defeat the traditional semantic approach, according to which what picks out the referent of a given word is exclusively the cognitive content which a competent speaker associates with that term. On this perspective, the reference of a word like e.g. 'water' is determined by the description of the properties commonly associated with it (such as "being a transparent, odorless, tasteless liquid normally found in oceans, rivers, rain, etc."). If the traditional semantic account does not distinguish between knowing the meaning of a term and grasping the thought expressed by it (so that knowing the meaning of a term coincides with what the speaker grasps when that term is uttered), Semantic Externalism insists on the dissociation between the meaning of a term and what a speaker knows about it. According to this latter view, the meaning and the reference of our words are determined by an objective relation between the speakers and the world, a relation that can be mistaken by or completely unknown to them (even to the fully competent ones). More specifically – on Putnam's account – the meaning and the reference of our words are determined by the real nature of their referents, whether or not we know or could possibly ever know what this nature is. Therefore, a word's extension is ultimately fixed independently of any conceptual content.

Such a theory of reference was originally conceived to explaining the semantics of words for natural substances and species, and has become very popular and quite widely adopted, since it seems to give a credible account of how the semantics of those terms works – definitely better than the traditional semantic theory. Putnam advocates for this semantic view on examples like the following: if we were to discover a liquid with identical superficial characteristics of water, but with different microstructure, we would not call it 'water', for it does not share the same nature as what we have paradigmatically called by that term. We would call 'water' – Putnam underlines – only each and every bit of liquid that has the microstructure H_2O (for that is the hidden structure of what is paradigmatically called 'water'), regardless the fact that also some other liquid may match the description of properties associated with the term. I will lay

out a brief overview of Externalism in Chapter 1. Those readers who are already familiar with this account can move forward to the following chapter.

The general focus of the present analysis is, however, the debate arisen from Putnam's proposal of extending his semantic view to artifactual (e.g. 'rheostat', 'pencil', 'oven', etc.) and social kind terms (e.g. 'university', 'money', 'pediatrician', etc.). He suggested this idea, albeit briefly, in his celebrated article of 1975, "The meaning of 'meaning'". There, he claims that the reference of artifactual and social kind terms is not picked out – as we may initially believe – by a set of properties associated with those terms, for it is fixed by the "nature" of the paradigmatic examples of the relevant kinds.

It is not hard to see why around such proposal have caused intuitions at odds and disagreeing views on the semantics of the words involved. It has indeed opened a still ongoing debate. Notwithstanding that the majority agree with Putnam as long as he restricts his analysis to natural kind words, many doubts have arisen about whether or not it can be applied also to the other kinds of terms. Specifically, this controversy originated with the thought experiments that Putnam laid out in order to prove his controversial thesis. Chapter 2 will be devoted to evaluating whether those arguments actually prove what Putnam intended them to prove, showing that they do not. The authors who engaged themselves in the debate have indeed taken side in favor or against the aforementioned extension of Putnam's theory, but nobody – as far as I know – has pointed out that Putnam's arguments themselves cannot prove anything in this respect. I then bring out the two requirements that a compelling externalist argument ought to satisfy.

It emerges already the strict interrelation between semantic and metaphysics, more specifically, the metaphysics of modality, underlined throughout and by the present work. In consonance, in Chapter 3 I begin by portraying a little sample of the common ontological and epistemological difficulties which arise around the kinds here at stake. I then sketch out the core positions taking part in the debate, pointing out how none of the theories advanced, neither on the descriptivist-internalist side nor on the externalist one, actually provides a convincing and objections-free defense of the position it conveys.

Chapter 4 is dedicated to the second part of my work: convicting Putnam's semantic approach of implicitly relying on a hyperrealist view on modality, showing how this

already problematic modal approach is also the source of the general semantic criticisms that Externalism has been accused of. In order to bypass the problems of a hyperrealist modal view, Thomasson proposes an alternative account: Modal Normativism (2013, *forth.*). This deflationary approach denies that modal statements are descriptive at all. The most basic function of what passes as basic claims of metaphysical necessity is, rather, to provide a very useful way of expressing constitutive semantic and conceptual rules in the object language, by which instructing and correcting those who are unsteady in using the terms occurring in those statements. This presupposes that there are rules of use which are meaning constituting. If we embrace this view – I shall argue – we are committed to dismiss Putnam's Pure Externalism, for we cannot accept that the meaning of our kind terms is entirely externally determined. Moreover, espousing this new approach to modality enables us to adopt a hybrid semantic theory, already proposed in order to overcome the aforementioned problems of which Putnam's Externalism in general has been accused. According to a hybrid theory of reference there must be some implicit but crucial descriptive elements in grounding a term's reference (at least the category of objects the word refers to is determined by us as a linguistic community). For example, if we are to coin a new term for certain things right in front of us, we need some guidance at least regarding which of the various different kinds it belongs to (natural, artifactual or social). Most of the times we even need some further conceptual content to tell us which are the relevant properties to seek if we are to refer, for instance, to a mammal rather than simply to a vertebrate. However, constitutive rules as conceived by Modal Normativism are also world-deferential, they can vary and be revised due to empirical discoveries, purposes or interests. All the empirical discoveries can be fully addressed by conditionalizing the constitutive semantic rules on empirical facts. Moreover, in the several cases in which the frame-level conditions expressed by the constitutive semantic rules are not met, whether to keep or forsake the term initially adopted becomes a matter of decision. We can use the constitutive semantic rules of those kind terms as a guide for their application, while still reserving some room to make decisions about all the uncertain examples. I conclude by proving how a hybrid theory of reference combined with a normativist approach to modality is superior in many ways to the previous semantic views that have

been advanced, for it handles much better not only the examples which made their case, but also all fuzzy cases.

In the conclusion I suggest that adopting these combined approaches leads to deflate the debate discussed in the present work, without however undermining the importance of the issues that arise around it. By adopting the notion of "metalinguistic negotiation", all the disagreements arose from the debate discussed here can be fruitfully seen as disagreements about how the terms involved *should* be used, moving the attention to the pragmatic level.

Lastly, I added an appendix to examine and criticize further possible defenses of Externalism, showing how the alternative I propose is still more tenable and attractive.

As this overview of my project clarifies, I support that semantics and the metaphysics of modality are strictly interrelated – they constantly influence one another. The present analysis emerges from an accurate and thorough analysis of Putnam's externalist account applied across the board and of the debate originated from it. This is why I leave aside a proper discussion of internalist or dualist positions, which do not embrace this view or leave this matter in the background.

My work focuses on advocating, in the end, for a unified account of the semantics of all kind terms, although on crucially different premises. Such account does not aim to give necessary and sufficient criteria of application of those words, but rather to point towards a radically different view of the entire issue. Putnam's account (and most of the positions presented within the debate) heavily relies on intuitions – a very problematic criterion, since it works only if the intuitions are met. My approach could represents a more viable strategy also in this respect.

Chapter 1

Semantic Externalism

In this chapter I lay out a brief overview of Semantic Externalism – the theory according to which the meaning and the reference of our terms are determined by an objective relationship occurring between the speakers and the world, a relationship about which speakers, even if fully competent, can be massively mistaken or of which they can be even utterly ignorant. For the sake of completeness, I sketch also Kripke's theory of proper names and natural substances and species, even though in the rest of the work I will only be concerned with Putnam's account of reference.[1]

1. Kripke's Semantic Externalism

1.1. Kripke on Proper Names

Saul Kripke's "direct reference" theory of proper names[2] aims to refute the image gave by the traditional theory, nowadays known as *Descriptivism* precisely because of its main conveyed idea, according to which names are *abbreviations* of definite descriptions. On a descriptivist theory, the relationship between the speakers and the world is *mediate* by *senses*. Classical descriptivists, like Gottlob Frege (1892) and Bertrand Russell (1905), basically claim that the sense associated with a name is the "mode of presentation" of the reference, it is the "path" towards the determination of its denotation, which is usually identifiable with a definite description. In other words, a particular use of a proper name refers by means of some cognitive content associated with that name – standardly, it is assumed, such cognitive content is a definite description (something like: "the teacher of Alexander the Great", "the ancient Egyptian

[1] Usually Putnam and Kripke's theory are associated under the label of "causal theories of reference", particularly due to Kripke's account of proper names' reference. I here distance myself from this label, for it does not sufficiently take into account the difference between the theory of proper names and the theory of kind terms. In this work I am only concerned about kind terms, and I do not attempt to extend the approach given to proper names as well.

[2] Note that, even though Kripke's account of proper names is known as "direct reference" theory, there is still a *mediation* in the determination of reference, played by the causal chain which link the speakers of a given linguistic community as well as by the intention of each speaker to refer to the same object to which the speaker(s) from whom she acquired that name was referring. I will not press this point here, since I will not be concerned with Kripke's semantic theory.

queen", etc.). [3] That definite description (or descriptive content) is what uniquely determines the reference of the name. In a nutshell, Descriptivism assumes the identity of a name with its definite description(s).[4]

In 1970 Kripke gave a series of conferences, later collected and published in the legendary *Naming and Necessity* (1980), meant principally to cast doubts on such a traditional view on reference, and to replace it with a different image.[5] Kripke primarily disputes the fact that, if we accept that the meaning determines the reference, we cannot understand the meaning in the sense of the descriptivists. Following the descriptivist account, identification across the worlds is guaranteed by means of purely qualitative characteristics. Accordingly, in different worlds, the same qualitative definite description could be met by different entities or even by no entity (in which case the existence of that object is simply denied). As a consequence, the same name can designate different entities in different worlds. In order to demonstrate his revolutionary thesis against this image, Kripke gives three types of arguments:

An example of a semantic argument. If – as Descriptivism holds – 'Cleopatra' is synonymous with "the ancient Egyptian queen", then to claim that Cleopatra is the ancient Egyptian queen should be a mere tautology.[6] Every competent speaker should know such identity only by knowing the name 'Cleopatra'. On the contrary, the claim

[3] The cognitive content is what a competent speaker has in mind when grasping that word, pursuant to which she understands that linguistic expression.

[4] Pretty soon it became clear that the traditional theory presented as such ran into obvious contradictions. The main difficulty is that, if a speaker associated with the name the wrong definite description, we would not want to say that such a name does not refer to anything or that it has no sense. In order to avoid such a complexity, it has been proposed to connect to the name not only one definite description, but rather a *family* or a *cluster* of descriptions. This solution has been advanced by L. Wittgenstein (1953), J. Searle (1958), and P. F. Strawson (1959). According to the strongest version of this theory, a name is synonymous with the cluster of descriptions connected to it, so that is sufficient that the name is associated with the disjunction of the properties specified by those descriptions. E.g. The name 'Aristotle' is synonymous with "the teacher of Alexander the Great", or "the ancient philosopher who was born in Stagira", or "the disciple of Plato", etc., but is never synonymous with the inclusive disjunction of those properties.

[5] Kripke specifies that he does not aim to give a detailed theory of meaning, but only an image of how things go in the determination of reference.

[6] A tautology is a needless repetition of an idea, especially in words other than those of the immediate context, without imparting additional force or clearness, as in "widow woman".

expresses the *fact* that she is the ancient Egyptian queen. That identity seems something that we learn.

An example of an epistemological argument. Typically, the only description connected to the name 'Gödel' is "the mathematician who discovered the incompleteness of arithmetic". "Suppose that Gödel was not in fact the author of this theorem. A man named 'Schmidt' […] actually did the work in question. […] when our ordinary man uses the name 'Gödel', he really means to refer to Schmidt, because Schmidt is the unique person satisfying the description?" [Kripke 1980, pp. 83-84]. We would not answer yes to this question – Kripke underlines. We would rather say: "The person who discovered that theorem was not Gödel but Schmidt". The speaker, even utterly competent, can be unaware of or wrong about what actually determines the reference of a certain name. Nonetheless, it is the objective relation occurring between the object and the world which gives the name's extension, not the cognitive content associated with it by the speaker. Indeed, once we are informed of our initial mistake, we are promptly willing to correct it.

An example of a modal argument.[7] If – as Descriptivism claims – 'Rosa Park' was synonymous with "the first lady of civil rights" or with "the mother of the freedom movement", then it should follow that the identity expressed by the proposition "Rosa Parks is the mother of the freedom movement" is analytic[8] and therefore necessary.[9] Yet, it does not seem necessary at all. Rosa Park could have sat in the back, not resisting segregation and so (unfortunately) not becoming the first lady of the civil rights. She would have still been Rosa Parks, though.[10] It does not seem that we need to have in mind a description (or a cluster of descriptions) of the properties which qualitatively

[7] Such arguments are based on our intuitions about what is or is not *possible*. Modality concerns the interaction of truth and falsity with possibility and necessity, namely with what can or cannot be otherwise.

[8] A proposition is analytic if it is true is virtue of its meaning, as in "All bachelors are unmarried men".

[9] Something is necessary if its contrary is logically impossible.

[10] Note that Descriptivism does not work even if we consider names as synonymous with a cluster of descriptions. To mention the same example, Rosa Parks could have been born in a different city, have had a completely different life, so not doing any of the things she actually did, and still be Rosa Parks. Or so Kripke argues.

identify a certain person in order to imagine counterfactual situations about that very person.[11] We simply refer to *her*.

Kripke reaches then the conclusion that proper names are *rigid designators*: "in every possible world [they] designate[s] the same object" (Kripke 1980, p. 48).[12] Definite descriptions do not give a name's reference, nor do they give its meaning, since meanings and references are determined by real factors, independent of the speakers' knowledge about the object designated by that name. Accordingly, the image about reference determination that Kripke proposes to replace the traditional one is the following:

> Someone, let's sat, a baby, is born; his parents call him by a certain name. They talk about him to their friends. Other people meet him. Through various sorts of talk the name is spread from link to link as if by chain. A speaker who is in the far end of this chain, who has heard about, say Richard Feynman, in the market place or elsewhere, may be referring to Richard Feynman even though he can't remember from whom he first heard of Feynman or from whom he ever heard of Feynman. He knows that Feynman is a famous physicist. A certain passage of communication reaching ultimately to the man himself does reach the speaker. He then is referring to Feynman even though he can't identify him uniquely (Kripke 1980, p. 91).

It is exactly in virtue of this causal and real connection among the speakers and between them and the referent himself that we can succeed in referring to *him*. The

[11] Note that sometimes the definite description is used to *fix* the name's extension – as happened with 'Hesperus'. However – Kripke highlights – this does not mean that the name is synonymous with *that* definite description. Hesperus could have situated in a completely different position in the sky or could have not been visible at night; yet, it would still be Hesperus – even though it could be named differently.

[12] Kripke specifies that when he says that a designator is rigid, that it designates the same thing(s) in all possible worlds, he means that, as used in *our* language, it stands for that thing(s) when we talk about counterfactual situations. He also does not imply that that thing(s) exist(s) in all possible worlds, but just that the name refers rigidly to it(them), if it(they) refer(s) at all.

determination of a proper name's reference has little or nothing to do with the cognitive content which speakers connect to that name. [13]

1.2. Kripke on Modality

By means of the types of arguments described in the previous paragraph, Kripke challenges also the conventional view on modality. Kripke disruptively argues that many of the truths we are used to think about as necessarily true because known *a priori* are instead contingently true – as the example of Hesperus shows (cf. footnote 11, this chapter). The name's reference is fixed by the definite description connected to it *a priori*; yet, what specified by that description is not necessarily but rather contingently true of Hesperus (that very star). A prioricity does not imply necessity. On the other hand, there are necessary truth which are *a posteriori* – as the same example illustrates. 'Hesperus' is a name of a certain heavenly body, which in fact happens to be the same as the body named 'Phosphorus'. Once the names' reference is being fixed, we use those name to rigidly refer to the same object in all possible worlds. However, "we do not know *a priori* that Hesperus is Phosphorus, and are in no position to find out the answer except empirically" (Kripke 1980, p. 104).

1.3. Kripke on Natural Kind Terms

Kripke widen the scope of the theses reported so far to the words for natural substances and species (such as 'water', 'tiger', 'gold', etc.): "According to the view I advocate, then, terms for natural kinds are much closer to proper names than is ordinarily supposed" (Kripke 1980, p. 127). In order to prove that, Kripke uses the aid

[13] As anticipated in footnote 2, Kripke's theory of proper names is a *causal-intentional* theory. There is a causal chain that connects the speakers of a linguistic community, in virtue of which, going back to the baptizer and through her to the referent herself, we refer to the *actual* referent of the name we use. For the mechanism of reference to work, "when the name is "passed from link to link", the receiver of the name *must* [italics mine], I think, *intend* [italics mine] when he learns it to use it with the same reference as the man from whom he heard it" (Kripke 1980, p. 96). As already specified, in the present work I will leave aside the discussion of proper names. I will sketch out more in detail Kripke's theory of natural kind words, which instead is an emblematic case of the traditional externalist theory I am here concern with.

of some thought experiments and examples, which can be ideally divided into two types of arguments:

Examples of arguments from error. *A real case*. We use 'gold' as a term for a certain kind of thing, which is thought to have certain identifying marks. However, some of those marks may not really be true of gold. We might discover that we were wrong about them. "Further, there might be a substance which has all the identifying marks we commonly attributed to gold and used to identify it in the first place, but which is not the same kind of thing, which is not the same substance. We would say of such thing that though it has all the appearances we initially used to identify gold, it is *not* [italics mine] gold" (Kripke 1980, pp. 118-119). This is what happened in fact when we discovered iron pyrites or fool's gold. *An imaginary case*. An interesting science fiction case on the same lines is the one of reptile-tigers. Kripke asks us to suppose that, in some part of the world, we find animals which, despite they look just like tigers, on examination were discovered not even to be mammals. They are found out to be very peculiar looking reptiles. Would we then conclude that some tigers are reptiles? No, we would rather say that, though those animals display the superficial features usually attributed to the species of tigers, they *are* actually *reptiles*. This is because they do not share the same internal structure of what we have called "the species of tigers" – with the paradigmatic instantiations of tigers. The term's extension cannot be given by the cluster of properties (even if conceived as an exclusive disjunction) usually associated with it. This is further proven – Kripke adds – by the fact that, if we discovered that *none* of the properties by which we identified tigers is true of that species, we would substitute all of those properties with the discovered ones.

Examples of arguments from ignorance. *A real case*. There might be cases in which we do not know what actually determines the reference of a natural kind term, for we have had (or we may ever have) no access to the internal structure of its extension yet. Nevertheless, it is that hidden nature what actually fixes the word's reference. This is what happened with gold and iron pyrites before the development of chemistry, when the fact that they are two different substances was precluded to us. We would not say – Kripke enhances – that, before 1750, 'gold' referred indistinguishably to gold and fool's gold. We rather maintain that 'gold' refers and has always referred to what has atomic number 79, despite the fact that we were not able to tell so. *An imaginary case*. A good

fictional example of radical ignorance is the one of water and XYZ imagined by Hilary Putnam. I will now turn to briefly lay out his semantic account.[14]

2. Putnam's Semantic Externalism about Natural Kind Words

2.1. Putnam's Twin Earth Thought Experiment and the Defense of Externalism

Putnam's well-known defense of Externalism about the semantics of words for natural substances and species goes like this:

1. Let us imagine that somewhere in the galaxy there is a planet, Twin Earth, that is exactly like Earth, except for the fact that its seas, oceans, rivers, etc. contain a liquid that is superficially indistinguishable from our water (it is colorless, tasteless, odorless, etc.), but has a very complicated chemical formula, abbreviated as XYZ.
2. This liquid is called 'water' by the Twin Earthlings.
3. What would happen if a spaceship from Earth ever visited Twin Earth? The supposition at first will be that 'water' on Twin Earth has the same meaning as 'water' on Earth. However, that belief will be corrected as soon as it is discovered that "water" on Twin Earth is not H_2O, but XYZ. Symmetrically, a spaceship from Twin Earth visiting Earth would report its findings in terms such as: "On Earth 'water' means H_2O".
4. But, what would happen if we roll back time to 1750, namely a time before the development of chemistry on either planet? Neither the speakers of Earth nor the speakers of Twin Earth could have known how to distinguish the two liquids from each other. Earthlings and Twin Earthlings would have had the same beliefs regarding the word 'water', namely they would have been in the same psychological state regarding that term.

[14] As specified many times, Kripke and Putnam together are usually considered the main exponents of Semantic Externalism. However, they developed such a theory simultaneously but independently, and this – in my opinion – has also brought peculiar and crucial differences in their accounts. Not to mention that Kripke's theory does not deal with artifactual and social kind terms. For the purposes of the present work, I will only deal with Putnam's account of reference.

5. Nevertheless, in 1750, just as in 1950, the reference of the word 'water' still would have been H_2O on Earth and XYZ on Twin Earth.
6. Hence, despite the fact that beliefs of the speakers of Earth and of the speakers of Twin Earth were the same, they were using 'water' with a different extension.

Putnam uses this imaginary case in order to prove Internalism false.[15] Such traditional semantic view rests upon the conjunction of the following two theses:

a. the meaning of a word is the cognitive content associated with it by a competent speaker, so that knowing the meaning of a word equals to being in a certain psychological state;
b. the meaning of a word (its intention)[16] determines its extension (or denotation)[17].

Putnam aims to prove that the those assumptions are mutually exclusive: if we choose a. we ought to refute b., and vice versa. As made clear by the experiment outlined above, he opts to reject assumption a., holding that it is perfectly possible for two speakers to be in the same psychological state (to associate the same cognitive content to a given word) and understand the same term differently.[18] We would not say

[15] Descriptivism, the traditional account of reference described previously, is labelled *Internalism* when not specifically applied to proper names. On a internalist account, what uniquely picks out the referent of a given word is the cognitive content which a competent speaker associates with the word. In particular, a natural kind term 'A' applies to x iff x has a weighted majority of properties P_1, P_2, …, P_n associated with x by the speaker. According to "modern" Internalism, the speaker can also be unaware of the cognitive content she connects to the term. What matters to the understanding of a word is the set of cognitive processes, even when not transparent, that the subject is maintaining. However, in this work I will consider only the "classic" Internalism, according to which the speaker (if not she herself, as a part of a linguistic community) is always aware of the cognitive contents associated with the words she uses.

[16] The intension of a term is the "concept" associated with the it (e.g: the intension of 'rabbit' = a large-eared, hopping animal, usually smaller than the hare, living in holes in the ground).

[17] The extension of a term: the sets of things the term is true of (e.g. the extension of 'rabbit' = the set of rabbits).

[18] This results even more evident if we reformulate this famous fictional case, highlighting its structure as a *reductio ad absurdum*: 1. Let us suppose that both Internalism's assumptions stand. 2. Earthlings and Twin Earthlings are in the same psychological state regarding the term 'water' (especially before 1750). 3. Then, for a., b. (the two internalist assumptions), 'water' has the same extension both on Earth and on Twin Earth. 4. But 3 is absurd! For we know that 'water' on Earth refers to H_2O and on Twin Earth refers to XYZ. 5. Therefore, we cannot accept both Internalism's assumptions: either knowing the meaning of a

– Putnam stresses – that a substance with identical superficial properties but a different microstructure is water. As Putnam puts it: "an entity x, in an arbitrary possible world, is *water* if and only if it bears the relation $same_L$[19] (construed as cross-world relation) to the stuff *we* call 'water' in the *actual* world" (Putnam 1975, p. 232). What really matters for the determination of a word's extension is the *actual nature* of its referents: its relation to the world, not to our mental states. However, this does not mean that speakers must know what this underlying trait (i.e. the nature of an entity) is.[20] It may take an indeterminate amount of scientific investigation to find out the conditions which are necessary and sufficient to claim membership in a kind.[21] Nevertheless, this fact does not imply that the meaning of the word in question will have changed in the interval. This is why the word 'water' would be ascribed the same extension both in 1750 and nowadays.[22] Following this reasoning, Putnam claims that natural kind words are *rigid* and *indexical*: *rigid*, because they refer to the same objects in every possible world in which they refer to anything at all; *indexical*, because they designate whatever shares the same nature with what we have paradigmatically called by those terms. This is the core of his semantic theory.

---

term is being in a certain psychological state or the meaning determines the extension. Internalism is false of the words in questions.

[19] Same liquid as.

[20] Putnam points out that, however, ignorance and error about membership criteria for a kind have limits. In order to be considered competent, a speaker ought to know at least the central features of the stereotype associated with the term she uses. Of course, this does not mean that then the operational definition is what analytically gives the word's extension. It is only to specify that a speaker's linguistic competence of a term like 'lemon' is verified by the fact that she would not apply it to an institution or to an algebraic structure.

[21] Putnam explains this concept by using the notion of *theoretical relation*: whether something is or is not the same substance as another substance may take an indeterminate amount of scientific investigation to determine. Moreover, even if a "definite" answer has been obtained, this answer is always defeasible: future investigation might reverse even the most "certain" example.

[22] Putnam means to defeat the traditional idea, according to which all natural kind words refer to "whatever satisfies the *contemporary* [italics mine] "operational definition" connected to them" (Putnam 1975, p. 235), for it "does not provide a necessary and sufficient condition for the application of any such word[s]" (p. 238). What matters for the determination of our natural kind words' reference is how the things actually are, not our epistemic knowledge about them. Such a reference is rather fixed by an ostensive definition pegged to a paradigm.

2.2. The Aluminum-Molybdenum Thought Experiment and the Sociolinguistic Hypothesis

In order to thoroughly make his case, Putnam uses the aid of a further thought experiment and an example. Let us suppose that – Putnam asks us – on Twin Earth, pots and pans are made out of molybdenum instead of aluminum, and that molybdenum is as common on Twin Earth as aluminum is on Earth, and that aluminum is instead as rare as it is molybdenum on Earth. Let us also suppose that molybdenum pots and pans are indistinguishable from our aluminum pots and pans. Finally, let us assume that the words 'molybdenum' and 'aluminum' are switched on Twin Earth. If a spaceship from Earth ever visited Twin Earth, the visitors probably would not suspect that "aluminum" pots and pans are made out of molybdenum, especially when the Twin Earthlings said they were. But – Putnam underlines – there is one important difference between this and the water-XYZ thought experiment: Earthling and Twin Earthling metallurgists can easily distinguish the two metals and tell that the two words are switched in the respective planets. The confusion of "aluminum" with aluminum involves only a part of the linguistic community. The common speaker can defer her possible lack of competence to the experts of her linguistic community.[23]

From this sorts of thought experiments or similar examples Putnam draws his sociolinguistic hypothesis, based on his Marxist background. According to this hypothesis, there is *division of linguistic labor*, which rests upon and presupposes the *division of labor tout court*. This phenomenon – Putnam explains – accounts for the fact that "the "average" speaker who acquires [a certain term] does not acquire anything that fixes its extension" (Putnam 1975, p. 229). Putnam insists that common speakers have to acquire words like 'gold', 'tiger', 'water', etc., but not the method of recognizing that the things to which they apply these terms are really gold, tigers, water, etc. They can rely on a special subclass of speakers: the *experts*, whose judgment they can trust in case of doubt. So, ultimately, this knowledge (the criteria for recognizing a term's extension) is possessed by the collective linguistic body, even though it is not possessed by each individual member of the body.

[23] This – as we shall analyze in the Appendix (§ 2) – this is a particular example of an argument from ignorance, conceived to convey the idea of the division of linguistic labor.

2.3. The 'Elm' and 'Beech' Example. A real case.

Let us say – Putnam confesses – that I cannot tell an elm from a beech. This means that I am in the same psychological state regarding the two words designating the two types of trees, since my concept of an elm is exactly the same as my concept of a beech. Yet, 'elm' in my idiolect have the same extension as in anyone else's idiolect, different from the extension of 'beech'. When I use the two words I succeed in referring to the two different types of tree. Such a difference in extensions, thus, cannot be brought about by a difference in the concepts *I* have of elms and beeches and the concepts the people who can tell an elm from a beech have. The extensions of those terms are determined by the nature of their referents, not by the mental content connected with them by the speakers.[24]

2.4. Putnam's Externalist thesis

Putnam believes then to have proved what follows:

> We have now seen that the extension of a term is not fixed by a concept that the individual speaker has in his head, and this is true both because extension is, in general, determined *socially* – there is a division of linguistic labor as much as of "real" labor – and because extension is, in part, determined *indexically*. The extension of our terms depends upon the actual nature of the particular things that serves as paradigms (Putnam 1975, p. 245).

A good theory of reference ought to take into account two contributions to the determination of the extension: the contribution of the society and the contribution of the real world, which have both been neglected by the traditional view.

The issues addressed so far can also be buttressed by resorting to Kripke's modal considerations (cf. 1.2). If Internalism was right, the truths expressed by statements such

[24] This is one more clear example of an argument from ignorance: Putnam – as he confesses – is not able to distinguish elm trees from beech trees; nevertheless the word 'elm' and the word 'beech' refer respectively to elms and beeches. Their reference is determined by the objective relation occurring between Putnam and the world, not by his mental states.

as “Water is H_2O”, “Gold is the element with atomic number 79”, “tigers have such and such DNA”, etc. should be analytic and hence true by definition, or *a priori*. By contrast, they are neither this nor that: they are discoverable by empirical research, and they are never incorrigible, for they can always turn out to be false. They are, if true, *metaphysically necessary* albeit *epistemically contingent*. In other words, they are, if true, necessarily true, and yet *a posteriori*.

Chapter 2

Putnam on Artifactual Kind Terms

Putnam was not content with showing that his semantic theory applies to terms for natural substances and species. He argues that Externalism can be extended to the great majority of all nouns, including artifactual kind terms (e.g. 'pencil', 'chair', 'television', etc.). The aim of this chapter is to examine the argument that Putnam outlined for this purpose, in order to evaluate whether it proves what Putnam intended it to prove.[1] Many authors have indeed taken sides in favor or against such extension (cf. next Ch. 3, § 2), but nobody – as far as I know – has pointed out that Putnam's argument itself cannot prove anything in this respect.[2] I will start by drawing attention to the structure of this argument, in order to note its similarity to the one used to show the indexicality and rigidity of natural kind terms (section 1). I will then explain why Putnam's argumentation does not succeed in its intent, and the requirements that, instead, a good externalist argument about artifactual kind words ought to satisfy (section 2). Lastly, I will outline Putnam's suggestion about social kind terms, concisely convicting it of the same fallacy (section 3).

1. Putnam's Externalism about Artifactual Kind Words

If one can be persuaded that natural kind words have an externalist semantics, terms like 'pencil', 'bottle', 'pillow', etc. are instead typically believed to be defined by conjunctions, or clusters, of properties. That is to say, they are typically believed to have internalist reference.[3] To disprove this thesis, in a short section of the aforementioned article, Putnam argues that "'pencil' is just as rigid and indexical as

[1] Putnam tried to demonstrate that his semantic Externalism applies further to social kind terms (i.e. terms like 'pediatrician', 'lawyer', etc.). In this chapter I will focus mainly on his argument about artifactual words, showing only briefly that it also does not work to secure an externalist semantics to terms for social kinds.

[2] Putnam's claim of extending Externalism to other kinds of words has open a still ongoing debate, the main participants of which include: S. Schwartz (1978, 1980, 1983); H. Kornblith (1980, 2007); J. Nelson (1982); D. Putman (1982); B. Abbott (1989); C. Elder (1989, 2007); A. L. Thomasson (2003, 2007), D. Marconi (2013).

[3] As remarked in the previous chapter (footnote 15), in the present work I intend "Internalism" as a sort of Internalism-Descriptivism, according to which the reference of our kind terms is given by a description (or a cluster of descriptions) specifying the properties associated with those kinds, for this is how the authors engaged in the debate understand that notion.

'water' or 'gold'" (Putnam 1975, p. 243). He uses once again the aid of a little science fiction – a thought experiment modeled on his Twin Earth argument:

> Imagine this time that pencils on Earth are just what we think they are, artifacts manufactured to be written with, while pencils on Twin Earth are organisms *à la* Albritton.[4] Imagine, further, that this is totally unsuspected by Twin Earthlings – they have exactly the beliefs about "pencils" that we have about pencils. When we discover that, we would not say: "some pencils are organisms". We would be far more likely to say: "the things on Twin Earth that pass for pencils aren't really pencils. They're really a species of organisms".

To call attention to its structure and especially to its similarities to the argument about natural kind terms, we shall draw out all of the passages of this symmetrical reasoning:

1. Let us imagine that somewhere in the galaxy there is a planet, Twin Earth, that is exactly like Earth, except for the fact that the things which appear to share all of the aspects of our pencils are actually not artifacts but a species of organisms.
2. These objects are called 'pencils' by the Twin Earthlings.
3. What would happen if a spaceship from Earth ever visited Twin Earth? The initial supposition of the visitors would be that 'pencil' on Twin Earth had the same meaning as 'pencil' on Earth. However, that belief would be corrected as soon as it was verified that "pencils" on Twin Earth were not artifacts but organisms.
4. However, before that discovery, speakers of Twin Earth had not been aware of the real nature of pencils. Earthlings and Twin Earthlings had possessed the same beliefs regarding the word 'pencil', namely they had been in the same psychological state regarding that term.

[4] The organisms *à la* Albritton to which Putnam refers in the experiment quoted above are organisms which look exactly like pencils look, but which have nerves and which spawn, so that we can see the offspring grow into full-grown pencils.

5. Nevertheless, both before the discovery as well as afterwards, the reference of the word 'pencil' was to a set of artifacts on Earth and a set of organisms on Twin Earth.
6. Hence, despite the fact that the beliefs of the speakers of Earth and of the speakers of Twin Earth were the same, they were using 'pencil' with a different extension. [5]

Moreover, Putnam adds:

> Suppose now the situation to be as in Albritton's example both on Earth and on Twin Earth. Then we would say "pencils are organisms". Thus, whether the "pencil-organisms" on Twin Earth (or in other possible universe) are really pencils or not is a function of whether or not the *local* pencils are organisms or not (Putnam 1975, pp. 242-243).

A passage which can be paraphrased as follows:

7. What would happen if the pencil-organisms case was proven on Earth, namely if we found out that all the pencils which exist and have always existed in *our* world are actually organisms? We would say: "Pencils (*these* objects) have turned out to be organisms".
8. Therefore, it is the nature of our *local* pencils that actually establishes whether or not the extension associated with the term used in other possible words is

[5] This argument is meant to be formulated in the exact same way of the Twin Earth one described in the previous section. By both arguments Putnam tries to show that if we apply a descriptivist theory to our kind terms we reach an absurd conclusion. To see that, we shall reformulate the pencil-organisms thought experiment: 1. Let us suppose that Descriptivism is true of these words. 2. Earthlings and Twin Earthlings are in the same psychological state regarding the term 'pencil': they think it has the same meaning both on Earth and on Twin Earth (at least before the discovery). 3. Then, for a, b (both the descriptivist assumptions), 'pencil' has the same extension both on Earth and on Twin Earth. 4. But 3 is absurd! For we know that 'pencil' refers to a set of artifacts on Earth and to a set of organisms on Twin Earth. 5. Consequently, Descriptivism fails for artifactual kind words (for we cannot accept both its assumptions). Nonetheless, it has to be underlined that Putnam's theory about our kind terms rests on some further assumptions, like the empirical presupposition (cf. next section) for instance. This is why, in the end, this argument runs into a fallacy (see below, section 3).

correct. If pencils are artifacts (in our world), they are artifacts with metaphysical necessity.

So, basically, the pencil-organisms thought experiment aims to show that if pencils are what we think they are – namely, *artifacts* with the common acknowledged purpose of a writing instrument – then, if a possible world ever existed in which there were organisms with the appearance of pencils, that would not be a world in which *pencils*, those very objects, *were* organisms. *If* pencils are artifacts, they are artifacts *with metaphysical necessity*.[6] Once the "nature" of pencils is discovered, there is no possible world in which it can be said of something with a different "nature" that it is a pencil. Neither is there any possible world in which this fact occurs. Nevertheless, also in the case of artifacts, all speakers could always find that they have been massively mistaken about, or even completely ignorant of, what this "nature" actually is.[7,8]

Therefore, Putnam felt he had reached the following conclusion: "That *pencils are artifacts* is not epistemically necessary in the strongest sense and, *a fortiori*, is not analytic" (Putnam 1975, p. 242). Not only the natural kind terms' extension, but not even the reference of a word like 'pencil' and the like words is analytically determined by the definition linked to them. Semantic Internalism does not then work with this kind of terms either, since "when we use the word 'pencil', we intend to refer to whatever has the same *nature* as the normal examples of the local pencils in the actual world"

[6] Putnam borrows from Kripke the distinction between *epistemically necessary* assertions (namely, assertions rationally not revisable) and *metaphysically necessary* claims (namely, true in every possible worlds). 'Water is H_2O', as well as all the statements in which a natural kind term occurs, is a (metaphysically) necessary truth, namely true in all possible worlds, independently of the fact that we know or we will ever know that truth. Those kinds of statements – Kripke revolutionarily points out – are, if true, necessary, and yet *a posteriori* (their justification depends on experience and empirical knowledge) and *synthetic* (true by how their meaning relates to the world). Putnam generalize these considerations to artifactual (and social) words as well.

[7] Again, within limits. See Ch. 1, footnote 20.

[8] Our epistemic access to the referents of our kind words is limited; nevertheless, their extension is always determined by the actual nature of their specific referents. What we can do – Putnam suggests – is point out a standard, but it can always turn out that we were wrong about or even ignorant of the nature of that selected standard.

(Putnam 1975, p. 243).[9] This is another step towards the conclusion that Externalism holds true across the board.

I have so far described the lines of the argument by means of which Putnam maintains that he has proven that Externalism applies both to natural and to artifactual kind words. The aim of the next section is then to see whether or not he actually succeeds in doing so, concluding that he does not. I will further clarify how a valid argument in this respect should be formulated.

2. On Putnam's Pencil-Organisms Argument

The Twin Earth thought experiment seems pretty convincing. In order to persuade us of the wide scope of his semantic theory, Putnam tried to sketch out a supporting reasoning about artifactual kind terms along the same lines. Yet, did he really succeed in his intent? In this section I will show that, whereas the pencil-organisms thought experiment seems, at least *prima facie*, equally persuasive, it does not in fact attest that artifactual kind terms have externalist reference. The analogy between the two arguments holds only in a superficial way. Let us examine in detail why.

Putnam is convinced that – despite appearances – natural kind terms and words for artifacts are semantically on a par:

> It seems that there is a strong tendency for words which are introduced as "one-criterion" words to develop a "natural kind" sense, with all the concomitant rigidity and indexicality. In the case of artifact-names, this natural-kind sense seems to be the predominant one (Putnam 1975, p. 244).

In order to demonstrate his thesis, Putnam conceived the pencil-organisms thought experiment as working (supposedly) exactly as the Twin Earth one does. They are both meant to support the view that we could be ignorant of or in error about most of the things that surround us. Nevertheless, the meaning of the words which denote those objects never changes, since it is given by an objective relation with the world. What

[9] It is noteworthy that Putnam does not explicitly mention the distinction among categories of terms; instead, he only uses examples of single words.

changes is only our knowledge about the referents of the words. Although it is epistemically possible[10] that water could turn out to have a different microstructure, and that pencils could turn out to be organisms, given that water is H_2O, and that pencils are artifacts, these things must be so. They possess a given underlying structure with metaphysical necessity – it is a fact which entirely depends upon how the things actually are, while being absolutely independent of our epistemic understanding about it. Nevertheless, it is such a trait what fixes those terms' extensions. So, Putnam believes he has shown that the semantics of a word like 'pencil' does not work – as we might initially believe – like the semantics of one-criterion words (words whose meaning is determined by an analytic specification) does, but rather like the semantics of natural kind terms.

Points 7 and 8 of the argument are intended to highlight the similarity between the two kinds of words even more. The history of science counts many cases of discoveries about the underlying structure of natural substances and species, as the example of gold and iron pyrites or fool's gold proves (cf. Ch.1, § 1.3). Once we found out that fool's gold was chemically different from typical gold (i.e. once we found out that they do not share the same nature), we did not say: "Some instances of gold are fool's gold". We rather said: "*Gold* (this very substance) has turned out not to be the same as this other mineral". As mentioned at points 7 and 8 of the argument, an analogous situation would happen regarding pencils: if we found out someday that all our pencils were actually a species of organisms we would say: "*Pencils* have turned out to be organisms". Our intuitive reaction to this counterfactual discovery is supposed to illustrate that also the assertions about the "nature" of artifacts are all but irrevocable. If, on the contrary, the semantics of such words required a descriptivist-internalist account, we would be obliged to apply 'pencil' to whatever matches the description associated with it – in this case the definition is not revisable.

The argument appears quite persuasive – maybe this is why many philosophers took its effectiveness for granted. Yet, it fails in its goal. If the analogy between natural and artifactual kind terms proved to be correct, we should be able to draw up, for both sorts of words, the criteria which rule their application. For natural kind terms, such a

[10] Something is epistemically possible if it may be true, based on what we know.

criterion seems quite evident (and Putnam recalled it clearly).[11] It is the essence common to all the samples of a certain substance or to the members of a certain species, namely, having the formula H_2O for water, having a certain DNA for tigers, having a certain atomic number for chemical elements, etc. In other words, the criterion is given by what exactly determines the membership in a certain kind. But what about the criterion which governs the application of artifactual kind terms? Which one is it, if there is one? Putnam seems to be quite reticent in this respect. His pencil-organisms thought experiment only attests the fact that 'pencil' applies, loosely, to artifacts – i.e. to objects that we make, which do not exist in nature without our intervention. But this cannot be sufficient. Selecting any other criterion than artifactuality as the "nature" of pencils (as well as, in general, of the referents of all artifactual words) seems to provide too general and vague a criterion. In order to prove what it was meant to prove, the argument should have better shown that even words like 'pencil' meet the two fundamental requirements of Externalism (in the way that the Twin Earth thought experiment did with the term 'water'). However, on closer inspection, Putnam's argument fails to show that the term satisfies either of them. According to these requirements, a term has an externalist semantics if and only if:

a. it refers to things that have a nature (necessary features) possessed by all of the members of the term's extension, and only by them;

b. linguistically competent speakers can be ignorant of, or mistaken about such nature.

That is to say, a term 'A' has an externalist semantics if and only if the members of its extension do share the same hidden structure, which must be metaphysically necessary albeit epistemically contingent. The word's reference is governed by the possession of such features independently of whether the linguistic community (and its individual members) are aware of them or can accurately describe them.

[11] Putnam specifies that the meaning of a word like 'water' is given by an "ostensive definition" (something like: "Let's call this liquid such and such"), which respects an empirical presupposition: the body of the liquid pointed to must bear a certain sameness relation (the relation of being the $same_L$ as) to most of the stuff the speakers of the linguistic community have been calling 'water' (i.e. to a paradigmatic bit of water).

These requirements follow from the *empirical presupposition* of Putnam's semantic Externalism – deducible from Putnam's discourse about natural kind words – which can be summarized as follows:

> (**EP**) There is a real nature shared among all (or most of) samples of a substance or the members of a species denoted by some classes of terms, even though we do not have or will we ever have any epistemic access to it.

In other words – as all the real and counterfactual examples point out – Putnam's theory presupposes that, in order to have an externalist semantics, a given term must apply to something which has a certain internal structure shared among all its referents. However – Putnam underlines – the description of that nature is always susceptible to modifications, due to the evolution of the research; even its most recent specification is corrigible.[12]

Hence, if a theory means to prove that terms for artifactual kinds do refer externalistically, first of all it has to be established whether or not there is such an artifactual "nature" shared by the things designated by these terms and what it could be. However, to specify only a series of necessary conditions for membership in a kind is not enough; such a "nature" must obviously be metaphysically necessary even if epistemically contingent. Thus, Putnam should have tried to identify what such a "nature" might be. He seems instead to elude the problem, showing simply that the word 'pencil' could have turned out to be a natural kind word, since it could have turned out to denote organisms. Nevertheless, this fails to show that the term, *qua artifactual* (i.e. as a term which refers to artifacts), has an externalist semantics.

We might then think that Putnam's pencil-organisms thought experiment shows that 'pencil' has externalist reference if it is at least *originally believed* to denote an artifactual kind. But Putnam's argument actually does not confirm this either, since, on

[12] This scientific progression is what Putnam tries to account for with his semantic theory. Cf. above, footnote 5. As Hilary Kornblith points out about Putnam's view: "we can only make sense of the history of science by rejecting a description-based account of the reference of natural kind terms, because early on in the scientific understanding of various natural kinds, the descriptions associated with natural kind terms typically fail to be precisely true – in some cases they fail to be even roughly true – of the referents of those term" (Kornblith 2007, p. 139).

closer inspection, what it actually proves is just that, if we in fact discovered someday that all the referents of the word 'pencil' actually had a natural nature, then 'pencil' would have an externalist semantics. However, – as already noted – in this kind of situation we would not cast doubts on this statement, for the externalist presupposition (EP) would be met: the word would apply to all and only the things which share the same DNA of those organisms which serve as the kind's paradigms. Nonetheless, this cannot establish that the word 'pencil', as an intended *artifactual* kind term, does have the semantics in question.

In conclusion, Putnam's alleged externalist argument about artifactual kind words is simply an *ignoratio elenchi*: it does not prove what it was meant to prove, but it rather reaches a different conclusion. This is evident from the following summary of the argument:

1. 'pencil' has an externalist semantics iff:
   - it refers to things that share a common nature (artifactuality, in the supposed case), which determines their membership in the kind of pencils;
   - linguistically competent speakers can be mistaken about such nature.
2. It is imaginable, and thus epistemologically possible, that pencils turned out not be what we thought they were (artifacts), but rather a species of organisms.
3. Therefore, 'pencil' has an externalist semantics.

The argument proves, *strictu sensu*, that the term 'pencil' can turn out not to be artifactual (i.e. not to denote artifacts), whereas it fails to prove that 'pencil' and artifactual terms in general (*qua* artifactual) refer externalistically. To prove that, Putnam should have shown that it is epistemologically possible, although metaphysically impossible, that we can be mistaken about (or even ignorant of) the specific *artifactual* "nature" of pencils. Moreover – as seen above in this section –, his argument is not effective even if the word 'pencil' is only *originally regarded* as referring to an artifactual kind. For, if we take an artifactual term to be a term that is *intended* to refer to artifacts, Putnam's argument tells us that such intention might be frustrated, but it does not tell us what semantics the term has when it is not frustrated. Either way, Putnam's reasoning tells us nothing about the semantics of the word 'pencil' in those cases in which the term is *purported* to refer to artifacts

and it *does*. In sum, Putnam's claim tells us little about how our actual well functioning artifactual kind terms work.

3. Putnam on Social Kind Terms: the Argument and its Refutation

Putnam's line of reasoning about social kind terms (words like 'pediatricians', 'university'. 'money', etc.) sounds more like a suggestion rather than a real argument. He imagines: "Couldn't it turn out that pediatricians aren't doctors but actually Martian spies? Answer "yes", and you have abandoned the synonymy of "pediatrician" and "doctor specializing in the care of children"" (Putnam 1975, p. 244). The structure of the argument is supposed to be similar to the previous ones:

1. Let us suppose that Internalism is true of these words.

2. Before the discovery, all the speakers are in the same psychological state regarding the term 'pediatrician': they think it refers to doctors specialized in the care of children

3. Then, for a, b (both the internalist assumptions), 'pediatrician' should be synonymous with the definition connected to it.

4. But 3 is absurd! For we know that 'pediatrician' refers to different things from the ones specified by the definition – it refers to Martian spies.

5. Consequently, Internalism fails for social kind words (for we cannot accept both its assumptions).

As evident from the scheme, also this one was meant to be conducted on the same lines of the other two arguments about the other kinds of words. However, it has to be admitted that it sounds far less plausible and convincing. We are taught and inclined to think that what actually defines a pediatrician is that she went to medical school, she specialized in medicine for children, and so on. The fact the agent in question is also specialized in secretly trying to get information about a country, or a person, or an organization, does not prevent her from also being a pediatrician. Moreover, the fact the she might not be a human being does not prevent that either, as long as she is able to treat our kids. That pediatricians are actually humans is something which may be implicit in the definition, but it does not seem to be an *essential* requisite. Only if none of the entities we usually identify as pediatricians went to medical school, etc. we would describe the situation as one in which it turned out that we were wrong, since the

original associated criterion was not met.[13] In the case of social and artifactual kind terms the original specification seems to dominate the new empirical discoveries.[14] Or so one might argue.[15] Whether or not this is really so it something that I am not interested in proving. What I aim to point out is that also this suggestion does constitute an authentic externalist argument.

The previous section makes explicit which requirements must be in place for a term to have an externalist semantics. They imply that a compelling argument which purports to confirm that artifactual kind words refer externalistically must show:

a. that their referents share a common essence;
b. that those objective features which define the membership in the same extension are metaphysically necessary albeit epistemically corrigible. Arguments from ignorance and error ought to be outlined in order to demonstrate this truth.[16]

By the imaginary case described in this section, Putnam did not show that we could be ignorant of or wrong about the peculiar "nature" which may characterize social kinds. He has at best gave a hint that we could discover that the referents of 'pediatrician' are not humans but Martians. But this cannot be sufficient. Selecting humanity as the common "nature" characterizing pediatricians (as well as, in general,

[13] This observation is also due to Stephen Schwartz (1980 p. 193, 1983).

[14] This argument is provided by Schwartz (1980, p. 191 and 1983, pp. 477-478). As Marconi puts it: "the description persists as a criterion for the application of that word" (Marconi 2013, p. 498). It seems that in the case of artifacts and social kinds, we would still apply the terms that denote them to objects conforming to the original description connected to them, despite the possible empirical discoveries made about the paradigmatic examples of those kinds.

[15] One might object to this, reporting cases in which we made some empirical discoveries about some artifactual kinds, and still that was no reason to give up the term or to apply it to the objects that still satisfy the original description. An exemplar case in this respect is the one of cable wires: originally produced for power distribution and later found very effective in data transmission. A counter-objection to this thesis might be that in those cases there was no other object to be applied the initial word to, nor any reason to manufacture a separate one. Or one might reply back that this is so because the empirical discoveries depend on the natural kind out of which the cable wires are made (see Marconi 2013).

[16] Precisely, it has to be shown that possession of such a common essence determines the term's reference independently of whether or not the linguistic community is aware of that essence (ignorance) or able to describe it (error).

referents of all social kind terms), does not provide only a too broad and vague a criterion, but rather no criterion at all. Moreover, the only plausible discovery implied by this example seems to be that the referents of the term did not go to medical school – they were only pretending to be doctors. However, this fact gives no evidence of the fact that 'pediatrician' and the like words have externalist reference – at least not in Putnam's sense.

4. Conclusion

Above (section 3), I argued that Putnam's pencil-organisms thought experiment does not prove what it was meant to: namely, that Externalism applies not only to natural kind words but also to artifactual kind terms. His argument only proves that we could be wrong about whether a term is artifactual or a natural kind term. In the supposed case, we thought the term 'pencil' was artifactual, whereas it actually referred to biological things. But, if that was case, nobody would cast doubts on which semantics would apply to it: namely, an externalist one, since it would apply to all and only the things which share the nature of the organisms that we had paradigmatically designated by that term (if there is a species they belong to). However, Putnam's argument is hardly informative as to what our actual well functioning *artifactual* terms do refer externalistically. Putnam should not have eluded the question; at the very least he should have hypothesized what this artifactual "nature" could be and laid out arguments from ignorance and error about it.

A similar line of objections can be moved against Putnam's suggestion about terms for social kinds.

We might thus possibly infer that proving that artifactual (or social) kind terms have an externalist semantics is seriously threatened by the fact that determining such artifactual essence might be far from easy. It does seem difficult to believe that there is something like a hidden structure which determines the kind of which artifacts are members.[17] Furthermore, it seems even more difficult to prove that we can be ignorant

[17] Although there are studies which reach the conclusion that, especially when talking about children, the essentialist assumption is quite pervasive, in that "young children seem to carry with them an essentialist bias, a bias that is then extended in to adulthood" (S. A. Gelman, J. D. Coley, G. M. Gottfried 1994, p.

of this structure or that empirical research may reveal that we misunderstood it, since artifacts are, by definition, objects that we make.[18] Nevertheless – as I will rapidly show in the next chapter –, it has to be mentioned that many authors in the literature have been affirming that there is such a common essence of artifacts (or social kinds), and have tried to indicate what it could be.[19]

Either way, the argumentation outlined in these pages is meant to recommend that a genuine externalist argument about artifactual (or social) kind terms needs to be grounded on the two requirements mentioned many times already: a good candidate must be selected as the common "nature" of artifactual kinds, about which adequate Ignorance and Error Arguments should be provided.

---

360), this hypothesis is very controversial. As much as this is quite widely believed when it comes to the categorization of natural kinds (see for instance S. Carey 1985, S. A. Gelman & J. D. Coley 1991), this is not so regarding artifacts. Studies on developmental psychology (see for instance F.C. Keil 1989, 1994) claim that, although kindergarten children cannot fully differentiate between artifacts and natural kinds, as the children grow they tend to develop causal theories that guide their selection. According to those studies, they tend to take into account hidden structures and discoveries in the case of natural kinds, whereas privileging more upfront features in the case of artifacts. B. C. Malt and S. A. Sloman (2007a, 2007b) conclude that to think there is an essence that we seek for in classifying artifacts is foolish, for artifact categorization, both in children and adults, depends on specific contexts and purposes. Also, P. Bloom (2007a) has a similar theory, in claiming that children rather tend to create hybrid concepts, so that they think, for instance, that water is both a natural and an artifactual kind. On the contrary, earlier, Bloom (1996, 1998) held an essentialist theory in this respect, affirming that we categorize artifacts on the basis of the intentions of their creators. This is only a little sample of the debate on whether or not there is an essence that defines membership in artifactual kinds. Moreover, very little agreement is reached about what this essence actually is. To mention an example, D. Gentner's experiments (1978) show that children and adults classify artifacts based on their form rather than their function, whereas W. Ahn's results (1998) explain the conflict about categorization in terms of causal role.

[18] I will not press this point here, however it remains a deep problem that many authors have tried to argue about. See for instance: R. Hilpinen (1992, 1999) and A. L. Thomasson (2003, 2007).

[19] As it is easy to figure out, and as already stressed in footnote 17, a strong debate is still open on this problem, which it is not the aim of this work to examine in detail. For reference, see for instance – in addition to the aforementioned articles and books by Ahn, Bloom, Elder, Gentner, Kornblith, Hilpinen, Malt and Sloman, Marconi, Putman and Thomasson – also: R. G. Millikan (1993, 1999); L. R. Baker (2004); P. Bloom (2007b); R. E. Grandy (2007); D. Sperber (2007); M. Carrara and P. Vermaas (2009); D. Mingardo (2011); S. Evnine (2016).

Chapter 3

Brief Commentary on the Debate about the Semantics of Artifactual and Social Kind Terms

In the previous chapter I mentioned that Putnam's proposal of extending the scope of Externalism to other kinds of words besides the natural ones has open a so far still ongoing debate. Indeed, although it is (virtually) unanimously acknowledged that natural kind terms have an externalist semantics, a long discussion converges on whether such a semantics can be extended to artifactual and social kind terms as well. I will now turn to give a brief overview of the dynamic of the debate in light of the considerations drawn out so far. In particular, I will focus on the two relevant requirements that a convincing externalist argument ought to satisfy, according to which:

a. those terms refer to things that have a nature (necessary features) possessed by all of the members of the term's extension, and only by them;
b. linguistically competent speakers can be ignorant of, or mistaken about such nature.

I will start by sketching a succinct sample of the main difficulties that puzzle both the internalist-descriptivist and externalist theories (section 1), in order then to focus on portraying the main positions of the debate, bringing to light their main weak points (section 2).

1. Artifacts and Social Kinds: Ontological and Epistemological Complexities

On a strong *Realist Paradigm* view and its three main theses, natural kinds are commonly regarded as *real*, whereas artifactual and social kinds are not; the latter ones result to be ontologically inferior and secondary in respect to natural kinds – as Michael Devitt, among others, underlines (1991).[1] According to the realist ontological thesis this is so because natural kinds meet the *Independence Principle*: things of kind K are real if they exist independently of all mental states. They also meet the *Ignorance* and *Error Principles*, according to which the conditions that determine whether or not something is of a kind K (i.e. the conditions that determine the nature of a K) can be completely

[1] Among others, exponents who hold a stronger position, according to which artifacts and social kinds should not even be included in the "furniture of the world" are: D. Wiggins (1980, 2001); P. Van Inwagen (1990); J. Hoffman & G. Rosenkrantz (1997). According to this Aristotelian anti-realist position, artifacts and social kinds are not real entities, since they are not associated with clear identity criteria which can individuate their essence.

unknown (Ignorance) to anyone or can turn out to be massively wrong (Error). In other words, the boundaries of natural kinds, as much as their existence, are utterly independent of humans' intentions and actions. According to the realist epistemological thesis, only the natural kinds' essence is determinable independently of anyone's concepts or knowledge. Potentially substantive discoveries subject to possibility of confirmation and error do not regard the other two kinds, about which we seem to have (at least some) epistemological privileges. Similarly, on the realist semantic perspective, the pre-existing boundaries characterizing natural kinds are what determine the reference of the terms designating them, independently of anyone's concepts or knowledge regarding the kind. However, because of the metaphysical and epistemological differences of their *denotata*, this phenomenon does not concern artifactual and social kind terms.

Putnam seems to implicitly challenge – quite surprisingly – this famous traditional paradigm, insinuating that also artifactual and social kinds perfectly fit it. As we shall see in the next chapter, this move is not unexpected at all – it is perfectly in accordance with his approach to modality. Whether or not he is right, it is something that I leave aside. What is certain, nonetheless, is that such a proposal has been very debated, especially if one wants to grant him the fact that it is supported by intuitions. Let us now rapidly sketch the main complexities that may arise around his suggestion.

Given the strict interrelation between the ontological, epistemological and semantic questions, one crucial point is represented by individuating, from those standpoints, the main differences and similarities among all the kinds at stake: natural, artifactual and social. Arguments in favor or against Putnam's implicit proposal usually focus on advocating for establishing similarities rather than discrepancies, or vice versa. From the Realist Paradigm described above results a background division between the *natural* and the *cultural realm* (which includes both artifactual and social kinds) frequently taken for granted, since it seems to provide a quite intuitive portrayal of what separates the kinds of objects at stake. However, as we shall see, this is far from obvious or unanimously shared.

Admittedly, the fact that social kinds constitutively depend on the collective acceptance of certain norms or constitutive rules, both for their origin and persistence, weights on the side of the realist. By contrast, mental states seem to play a different role

in the creation, and especially in the persistence, of artifacts and artifactual kinds.[2] A microwave would still exist even if all humanity would disappear, whereas in the same situation a dollar bill would not count anymore as such. Moreover, the microwave would not only still exist, it would also potentially work as such (unless, of course, it was irremediably damaged).[3] Even so, artifacts and naturally occurring entities have been usually classified differently under all the relevant aspects, since specifying the identity conditions of the latter appears a real challenge. Due to the possibility of disassembly, radical reassembly, and reparability such identity conditions seems to be very arbitrary. Another problem is represented by the virtual impossibility of identifying those kinds' individual essence with the class essence, which converge in the associated problem of their classification. Nonetheless, despite their undeniable historical dependence on mental states (after all, artifacts are the product of our deliberate actions), many authors have claimed that they are as much ontologically respectable as natural kinds, grounding their thesis on individuating a common essence which determines membership in each artifactual kind.[4] The most adopted choice seem to opt for their function(s), intended as proper functions, intended functions, actual functions, causal functions or as a combination of these properties.[5] Artifact function(s) is considered the relationship that must exist among different artifacts (meant as single items) so they can be grouped under the same kind. Consider the case of radios: radios can be shaped very differently (e.g. some resemble Coke cans); nevertheless, we categorize them in base of their possibility of receiving radio waves. Function(s) seem to be the most plausible and objective candidate to represent the "nature" of artifactual kinds. They provide necessary and sufficient conditions for the existence, persistence and extinction of artifacts and artifactual kinds. Moreover, functions can be object of

[2] J. Lowe (2014), for instance, claims that artifactual kinds are mind-dependent only from a *causal*, but not from a constitutive point of view. Therefore, they are just as real as natural kinds (cf. p. 20).

[3] This example is due to Marconi (2016).

[4] For references see the ones mentioned in footnotes 17 and 19 of Chapter 2.

[5] On the discussion of these accounts see for instance: R. Cummins (1975); R. G. Millikan (1984, 1993, 1999); F. B. Preston (1998); F. Keil (1989); J. L. Rips (1989); K. Neander (1991); A. Matan and S. Carey (2001); P. McLaughlin (2001), A. L. Thomasson (2003, 2007a, 2014); L.R. Baker (2007); C. Elder (2007); R. Grandy (2007); M. Carrara and P. Vermaas (2009); W. Houkes & P. Vermaas (2010); M. Carrara & D. Mingardo (2013); J. Vega-Encabo & D. Lawler (2014); M. Franssen *et al.* (2014);.

empirical discoveries. Take for example the teslascope: such instrument, invented by Nikola Tesla, was initially thought to receive signals from extraterrestrial beings; it was later discovered that it actually functioned only as a wave receiver (of signals from Jupiter's magnetosphere). Functions could represent the underlying trait which determines membership in artifactual kinds – a further reason for treating those kinds on a par with natural kinds. It appears to be the most viable candidate for meeting also the aforementioned requirement (b).

However, such choice has not been undisputed. All the functionalist theories clash with some problems: there are still artifacts which those candidates are not able to individuate and categorize. Moreover, many authors have cast doubts on such most shared insight about the essentiality of functions: "functionality is not as important for describing artefacts as it is often taken to be. This calls for a change of focus: for properly understanding technical artefacts, philosophers, but also engineers, should consider the intentional actions that involve these artefacts instead of merely regarding them as functional objects" (Houkes and Vermaas 2010, p. 4).[6] In place of function(s), several other hypothesis of essence of artifactual kinds have been put forth: e.g. structure, form, intentions understood in various ways, or a combination of such properties.[7] Nonetheless, little or no agreement has been reached so far.

From the realist paradigm follows that artifactual and social kinds are set apart from natural kinds also when it comes to epistemology, for it is commonly believed that the crucial dependence of the former on mental states and human intentions prevent us from (most kinds of) ignorance and error about their "nature". However, also on the epistemological point of view, the distinction between artifactual and social kinds is not as sharp as it is generally conceived, nor is it the division between the natural and the cultural realm. As remarked above when talking about functions, artifacts – but not social kinds – manifest a feature partially shared with natural kinds: the possibility of making discoveries on their essential properties.[8] This possibility, however, may be

[6] On the critique of *functionalist* accounts for artifacts and on the proposal of an *intentionalist* view, see also, for instance: A. L. Thomasson (2007a, 2014).

[7] For references see the ones mentioned in footnote 19 of Chapter 2.

[8] *Discoverability* is intended in the realist sense. If it is discovered at the time *t* that p, then: (i) it is known at *t* that p, (ii) until *t*, it was not known that p, (iii) at least for some time before *t*, it was (already) the case that p. Thus, for instance, if we *discovered* that water is H_2O, then water *is* H_2O, and *was* H_2O even

thought to partially be a consequence of the fact that concrete artifacts are made by assembling and modifying natural materials in various ways. Aspirin, for instance, was invented with the purpose of being a pain-killer, whereas later it was discovered to function very well also as a blood-clotting preventer. The additional discovered function depends on the unpredictable causal properties resulting by the assemblage and modifications of the materials used in producing aspirin.[9] However, on closer inspection, this is not an exclusive characteristic of artifactual kinds. Since social objects sometimes need a physical support, this entails that (some sorts of) empirical discoveries can be made also about them: a driver's license can reveal the property of turning blank after some years (say, because it has deteriorated). As in the case of artifacts, the event would obviously bear on the license's function, not just on its material constitution and form: the driver will not be allowed to exercise his rights unless he substitutes it with a legible one.

In this sense, in spite of their being mind-dependent entities intentionally created by humans, both artifacts and social kinds have mind-independent *properties*. It seems, then, that artifactual as much as social kinds can be considered *real* as much as natural kinds, and not only because of their (possible) dependence on physical or natural kinds.[10] Nonetheless, artifacts display a crucial difference in respect to social kinds: artifacts' essential properties all ultimately depend on their natural basis; whereas social entities' essential characteristics depend on *revocable* human intentions and norms. As mentioned above, a microwave could still perform the function of heating foods and beverages even if it was decided to stop using it to do so, whereas if some laws were abolished they would immediately cease to be applicable and to have their normative

---

before our discovery. Nonetheless, if water was H_2O even before our coming to know that it is, then its being H_2O does not depend on our knowledge of that fact. Moreover, we can discover that p even in the absence of any prior belief that p, so that its being the case that p turns out to be independent of our beliefs as well. There is no epistemological privilege regarding that fact.

[9] See on this Marconi (2013).

[10] Some authors claims, on the contrary, that cultural kinds should be part of our ontology, exactly because of this dependence on natural kinds. Among others, D. Putman (1982); C. Elder (1989, 2007); L. R. Baker (2004, 2007) affirm this position. In Baker's own words: "It is tendentious to claim that the existence of artifacts depends not on nature, but on us. Of course, the existence of artifacts depends on us: but we are part of nature. It would be true to say that the existence of artifacts depends not on nature-if-we-did-not-exist, but on nature-with-us.-in-it" (Baker 2007, p. 107).

function – as (fortunately) the Nuremberg Laws testify.[11] It seems true then, that the essential dependence of social kinds on human intentions and norms set those kinds apart, basically ruling out the possibility of singling out a "nature" thereof, at least in the relevant sense here in object, consequently making any attempt to advocate for their designators to have externalist reference fail.

In sum, it result that the realist is right in claiming that there is an essential distinction between natural and social kinds from both the metaphysical and the epistemological points of view. Nonetheless, such a distinction appear much more controversial when opposing natural to artifactual kinds. The latter ones may be claimed to be as real (in the strong sense) as the former from all the relevant aspects. The boundaries drawn by the realist between the two kinds seem to collapse, or at least to be much less defined as they are commonsensically thought to be. Just think, for instance, to GMOs, or anthropogenic hybrids, such as square watermelons[12] or seedless grapes[13], or patented genes[14]. Those boundaries seem to becoming increasingly blurry, also due to the evolution of science and technology. Contrary to what Schwartz claims (next section), we might someday even be able to create "genuine" cats, hence making his distinction fail (cf. Schwartz 1983, p. 477).[15] Paul Bloom holds that even water can be considered a hybrid to a certain extent: it is a natural kind, for it "fall from the sky, after all, and is found in oceans, rivers, and lakes. But there are also good reasons to take it as an artifact kind", for it comes from bottles, is filtered, purified, etc. (Bloom 2007, p. 155).[16] The underlined similarities between the interested kinds have indeed brought

[11] For a detailed and exhaustive treatment of these topics see Marconi 2016.

[12] This example is due to F.C. Keil, M. L. Greif, R. S. Kerner (2007).

[13] This example is due to D. Sperber (2007).

[14] See on this, D. J. Kevles "The Gene You Can't Patent" (2013).

[15] Schwartz is referring to the phenomenon of "creating counterexamples". The sharp metaphysical distinction between natural and artifactual kinds is buttressed by the fact that only in the latter case we can create counterexamples to whatever statements in which an artifactual kind term occurs. We could create "genuine" artifactual pencils, but not "genuine" animal cats. This is why the terms referring to the respective kinds cannot be regarded as behaving in the same way.

[16] To mention a different interesting example, also *iron* could be considered a sort of hybrid: some of its physical and chemical properties (such as its melting point or solidity) do not depend on our purposes and desires. On the other hand, pure Fe does not exist in nature: it something created. Some of its properties (such as its shape and mass) have been chosen by us. Iron depends, to some extent, on human intentions.

many authors to hold that "there is a continuity between natural kinds and artifacts, then it doesn't seem that there should be a metaphysical distinction" (Grandy 2007, p. 24), or an epistemological one.[17],[18]

In light of the considerations drawn out in the present section, let us now focus on the debate arose around the main questions of the present work.

2. The Debate on the Semantics of Kind Terms: Externalism *vs* Internalism[19]

For the purposes of the present analysis, I will outline the dynamic of this debate the in light of what specific claim(s) of Putnam's they stress in order to prove Externalism true or false of the other kind words. In particular, I will show that, even though internalist-descriptivist theories seems to meet general intuitions about artifactual and social kinds, they are not free from criticisms. None of the externalist accounts, on the other hand, represents a complete, strong justification of Externalism applied to the terms designating those kinds, for it does not meet any of the fundamental requirements highlighted in the previous chapter (cf. Ch. 2, § 3). Among the externalist positions, I will analyze only the arguments about artifactual words, since – as stressed above – the particular metaphysical and epistemological status of social kinds bears also on the possibility (if any) of widening the scope of the externalist theory to their designators.

[17] Among others, representatives of these positions are: D. Putman (1982); C. Elder (1989, 2007); L. R. Baker (2004, 2007); P. Bloom (2007a); R. Grandy (2007); D. Sperber (2007); M. Carrara and P. Vermaas (2009); M. A. Khalidi (2015).

[18] The last considerations may be thought to lead towards an internalist (cognitivist) standpoint, according to which we categorize things according to a anti-essentialist perspective. The way to categorize the objects is not intrinsically discrete (by necessary and sufficient conditions) – but rather continuous. However, - as implied in what I will argue at a great in the next chapter – this cannot not be what Putnam had in mind. This is why – on the basis of the framework in which the present discussion is inscribed – I will leave aside a detail discussion of these positions.

[19] For an introduction to the topic, see Ortega Cano (2013). For other references on the develop of the debate, see Marconi (2013) and. M. Carrara & D. Mingardo (2013).

2.1. A Critique to Putnam: No Common Essence

Stephen Schwartz (1978, 1980, 1983) has been the first one to put Putnam's proposal into question: "Terms for kinds of artifacts [and nominal kind terms in general] do not even start out as indexical" (Schwartz 1978, p. 572).[20] His main argument consists in insisting that "what makes something a pencil are superficial characteristics such as a certain form and function. There is nothing underlying about these features. They are analytically associated with the term 'pencil', not disclosed by scientific investigations" (Schwartz 1978, p. 571). If Putnam was right, we would apply 'pencil' to whatever has a certain inner essence, and we would even extend (or withdraw) the term to things that do not superficially resembles the paradigmatic pencils, as long as they do (or do not) share the same hidden structure. However – Schwartz claims – the difference between natural kinds and nominal kinds (such as artifactual and social) is that the extension of the latter ones is not gathered by any underlying trait. Schwartz holds that, contrary to the case of natural kinds, membership in nominal kinds comes with satisfying the description. That this is so results even more evident by the fact that "in the case of nominal kind terms the original specification "dominates" [the new empirical discoveries]" (Schwartz 1983, p. 477). If we found out – as in Putnam's examples – that pencils are organisms or that pediatricians are Martian spies, nothing prevent us from making artifactual pencils or someone from going to medical school and specializing in the care of children. "We would all think that now we have [real pencils]" (p. 477).[21] By contrast, it seems that, if we were to discover that all cats are robots (cf. Putnam 1970, 1975, pp. 243-244), we could hardly make a "genuine" cat.[22] Whereas it is an important

[20] Barbara Abbott agrees with this distinction and general considerations, holding that: "Artifacts are typically made by humans and are categorized according to their purposes, so we know how they are shaped and what they are used for. When it comes time to name them we have the reference-determining properties there at hand, we know what we are talking about. It is only in the case of nature's species that we have observable kinds whose real essence is mysterious, and so only in that case must we leave the reference-determining properties open" (Abbott 1989, p. 281).

[21] Schwartz is actually here referring to the case put forth by James Nelson, according to which we could find out someday that all our pencils are in fact devices by which Martians spy and control our activities (cf. § 3.3). In that case, we would probably start manufacturing "real" pencils.

[22] On this observation and its consequences see Marconi (2013, pp. 498-499).

feature of a natural kind term that in determining its extension the underlying nature dominates the associated description, with nominal kind terms the situation is just the reverse: the associated description is what plays the semantic role in determining their reference. "The correct theory of natural language will recognize that natural kind terms are indexical, while at the same time recognizing that nominal kind terms are not" (Schwartz 1978, p. 574).

This view, however, faces two main problems. On the one hand, it neglects the difference between artifactual and social kinds;[23] whereas on the other hand, the solution stated seems to be too simplistic, taking for granted a distinction between the natural and the cultural realm. that He holds that each single competent speaker have virtually incorrigible beliefs on the objects of the cultural realm. However, – as the previous section shows – this is much more disputable than it commonsensically considered to be. From a strict semantic point of view, however, Schwartz's account has a related further problem: according to his account every competent speaker can provide, simply because she is a competent speaker of a specific language, an individuating definition in terms of form and function of the referents of the nominal kind terms she uses. This thesis may be convincing when thinking about terms for everyday life objects such as 'chair', 'table', 'pencil', etc.; yet, it is far less obvious when it comes to more specialized terms such as 'pantograph', 'franchising', 'radiator', etc. Hilary Kornblith points out and tries to overcome those difficulties, although – I shall point out – also her semantic approach presents some weakness.

2.2. An Externalist Defense: Deference to Experts

In his "How to refer to artifacts" (2007), Kornblith holds – in contrast to Stephen Schwartz (1978, 1980, 1983) – that "while there are, beyond doubt, important metaphysical differences between artifacts and natural kinds, the mechanisms of reference are insensitive to these differences" (Kornblith 2007, p. 241). To both natural

[23] Such a distinction has actually been neglected in both directions: on the one hand, by authors – like Schwartz – who endorsed that artifactual and social kinds are on a par and should be distanced from natural kinds because of their mind-dependent essence and properties; on the other hand, by authors who tried to bring to light what similarities those kinds have in common with natural kinds. (e.g. C. Elder 1989, 2007).

and artifactual kind words an externalist semantics applies, for they are both subject to the *division of linguistic labor* (cf. Ch. 1, § 2.2).[24] "The traits which determine artifactual kind membership need not be known by a speaker for reference to succeed, for artifactual kind terms are just as susceptible to the phenomenon of the division of linguistic labor as are natural kind terms" (Kornblith 2007, p. 148). He underscores that, despite many kinds of artifacts are ones which competent speakers can actually characterize by a description (or a cluster of descriptions), this is not true of every kind (e.g. 'rheostat', 'Chippendale furniture', etc.).[25] Yet, "the fact that a speaker does not know a uniquely individuating description does not seem to prevent the speaker from referring" (Kornblith 2007, p. 139). This suffices to defeat the application of semantic Descriptivism to this kind of words.

Nevertheless, it is very crucial to note that the *social component* of the meaning is only a part of what gives our terms' extension.[26] As I have stressed many times in the previous chapter, what is really decisive in determining a word's reference is the contribution of the real world (the indexical component), and this can be totally obscure to *all* speakers. As a consequence, in Putnam's view, both competent speakers and expert speakers can be utterly ignorant of or wrong about the nature of the things to which they refer, and consequently about the actual extension of the terms they use. According to Putnam, the ultimate criterion to refer to in determining a word's referent is the *indexical component*: the reference of a term is governed by the real nature of the things paradigmatically indicated by the ostensive definition. The term refers to whatever has that nature. By ruling out the metaphysical component in the explanation

[24] In Kornblith's words: "What is clear, however, is that the division of linguistic labor is just as much a part of the world of artifacts as it is a part of the world of natural kinds and individuals. And this point by itself is sufficient to ground Arguments from Ignorance and Error in the case of artifacts which are exactly parallel to those which underwrite the new theory of reference for names and natural kind terms" (Kornblith 2007, p. 144).

[25] Kornblith advances this point also against Amie Thomasson's thesis of *epistemological privilege* regarding artifacts (see 2007, p. 55 and pp. 63-64).

[26] This is confirmed by Putnam's own words: "We have now seen that the extension of a term is not fixed by a concept that the individual speaker has in his head, and this is true both because extension is, in general, determined *socially* – there is a division of linguistic labor as much as of "real" labor – and because extension is, in part, determined *indexically*. The extension of our terms depends upon the actual nature of the particular things that serves as paradigms" (Putnam 1975, p. 245).

of the semantics of artifactual kind words, Kornblith has neglected the very core of Externalism. Needless to say, his attempt to justify the extension of such semantics to artifactual words was doomed from the start.

Such an account actually represents a form of "masked" Internalism, for the reference of our kind terms is fixed by the experts of the linguistic community in which they are used, so that they are available, through them, to all the individual speakers of that community.[27] It then overcome Schwartz's criticism in this respect. Nonetheless, it does not take into account the important influence that metaphysical and epistemological differences among all the kinds at stake has on the semantics of their designators. From this point of view, Thomasson's response to Putnam's proposal represents a much viable account.

2.3. A Critique to Putnam: Communitarian Internalism[28]

Amie Thomasson (2003, 2007a), insists – with Schwartz – on the fundamental ontological and metaphysical differences between natural kinds, on the one hand, and artifactual and social kinds, on the other hand. She argues that "not only are artifacts produced by humans, but also that, in virtue of the different ways in which our artifactual terms and concepts function (ways that are established definitively by the speakers' and thinkers' intentions), that specific natures of artifactual kinds are determined (often gradually and collectively) by makers' [substantive] concepts about what features are relevant to kind membership" (Thomasson 2007, p. 73). This is valid also for social kinds – the only difference is that those kinds depend on humans' intentions even more, for their existence and essential features depend on the collective acceptance of constitutive rules which outline sufficient conditions for membership in those kinds. Such difference bears, consequently, also on the epistemology and the semantics of those kinds, putting us in a certain epistemic privileged relation regarding those kinds, and entailing that the reference of those kind terms is determined, at some

[27] I will examine this account and its criticisms in detail in the appendix.

[28] The view proposed here is only the embryonic version of the hybrid approach she will defend later on (2007b) and that I will discuss in detail and endorse in the next chapter.

level, by our concepts.[29] All the principles of the Realist Paradigm (cf. § 1), insofar as they are so suitable for natural kinds, fail for artifactual and social kinds. The "nature" of the latter kinds is, at least in part, stipulatively established, as much as, consequently, are their boundaries and the reference of their designators. A parallel with natural kinds is not even remotely possible. In particular, regarding reference, Thomasson concludes: "not *all* cases of reference to these kinds may proceed along a causal model" (Thomasson 2003, p. 604). Insofar as the main condition for membership in a kind are established or stipulated, very little room for revision, ignorance or massive error is left.[30]

Like Schwartz, Thomasson holds that "we may require very different theories to handle independent parts and aspects of the world, and those that are in part our own construction (Thomasson 2003, p. 607). She proposes to ground reference of the other kinds of terms by appealing not to a common internal essence which determines membership in a (natural) kind, as it happens when grounding reference of natural kind words, but rather to common largely successful *intentions* of creating something of that kind. The grounder of those terms (typically, the maker(s)) will know, first of all, what sorts of features are relevant to ground an artifactual rather then a natural kind word, and, second, what specific traits are relevant to membership in that kind. However, unlike Schwartz's, this is not an internalist view in the traditional sense, for "epistemological realism and causal theories of reference may apply locally (relative to some groups G), but not universally, to kinds that lack natural boundaries" (Thomasson 2003, p. 584). In this sense, it could be better defined as "Communitarian

[29] In Thomasson's own words: "As a consequence, both our epistemological and semantic theories must differ substantially from those appropriate for the presumed independent kinds of nature, for certain groups must have substantive concepts of the nature of the kind, where those concepts play a crucial role in determining the extension of the relevant kind terms, and are protected from certain forms of ignorance and error (Thomasson 2003, p. 607).

[30] In Thomasson's own words: "where K is an essentially artifactual concept, if a K *does* exist, it follows from this that someone (namely, at least the maker) has a substantive concept of the nature of Ks that is not subject to massive error, whereas members of natural kinds may well exist without anyone having any concept whatsoever of that kind or its nature" (Thomasson 2007, p. 61).

Internalism".[31] The maker(s) of the kind function as the expert(s) to whom defer our possible lack of competence.

Nonetheless, it is not so evident – as showed by the discussion of the previous section (§ 1) – that what matters are the maker(s)'s, rather than the user(s) concepts', or that instead what really defines the membership on a artifactual kind is its function(s). Think about the vacuum tubes. This type of components that, thanks to an external power source, provide an amplified power signal, have been created to be used in electronic equipments such as radio receivers and transmitters, televisions, and in general in all types of electrical signal amplifiers. By chance, it was discovered that the waves emitted can be used to heat or bake food, so that vacuum tubes have become the essential component of microwave ovens. The view here advocated by Thomasson fails in taking into accounts those changes in use or discoveries: we could not call those components 'vacuum tubes', for they do not fit their maker(s)'s original concept.[32]

Some of the externalist positions try to overcome this problem. Let us now turn to examine them in detail.

2.4. An Externalist Defense: Artifacts do have a Common Essence

Requirement (a) entails that a good externalist argument for artifactual words ought to prove that artifacts do share a common "nature". The best theory to this effect was the one advanced by James Nelson in his "Referring to Artifacts" (1982). With his thought experiment of pencil-control devices (that we shall see below), he aims to suggest that this essence consists of "a certain combination of structural and functional features" (Nelson 1982, p. 362). By putting forward this hypothesis, he has already made progress compared to Putnam. Nelson indeed at least tries to identify what the "nature" of artifacts can be, not confining it only to a general identification with a class of objects. His argument from error (next paragraph) adds further proof thereof. Nonetheless, Nelson's theory becomes untenable when he specifies that the set of properties connected to those kinds are regarded as "surface properties" (Nelson 1982, p. 363). One could mention many examples of structural and functional features which

[31] See on this Marconi 2013, p. 506.

[32] A similar critique is articulated by Marconi (2013, pp. 506-507).

are far from being superficial.[33] However, what is noteworthy here is that Nelson further weakens his case when he claims that "these "superficial properties" may well serve as the "nature" of those things which lack a unified "hidden structure", as do artifacts, or as does water in the latest counterfactual [case]" (Nelson 1982, p. 363). Contrary to what Nelson declares, this is exactly what Putnam denies. If – as in the counterfactual situation mentioned – "it turned out that the bits of liquid we call 'water' had *no* important common physical characteristics *except* the superficial ones" (Putnam 1975, p. 244), 'water' would refer descriptively. Nelson's argument, although initially promising, cannot show that we should not have reservations about the scope of Externalism.

2.5. An Externalist Defense: a Case of Ignorance and a Case of Error

Kornblith's (1980) and Nelson's (1982) thought experiments represent emblematic attempts at describing, respectively, the two cases. Kornblith asks us to imagine what we would say of a Martian anthropologist who found a doorstop and baptized it and the class to which it belongs 'glug'. "It seems that in spite of the Martian's ignorance of the function of doorstops, he has succeeded in using the term 'glug' to refer to doorstops" (Kornblith 1980, p. 114). He adds also that, despite the fact that this correct use of 'glug' to denote doorstops might seem parasitic upon the existence of human doorstop experts (who can perfectly pick out the term's extension descriptively), this is not the case. For "the Martian anthropologist might try to construct a theory about the nature and function of glugs" (Kornblith 1980, p. 114). But could he really do that? What if the Martian finds another doorstop, of a different shape and material? It seems that he would have no reason to call it 'glug', since there would be no evident clues to the correct application of the term. The only way the Martian could be considered to be referring to doorstops with 'glug' is if we attribute to him at least the intention of baptizing generically *the kind to which that thing belongs* – which happens to be the kind of doorstops. Or if he intended to baptize the class of which that object was

[33] This is only the tip of the iceberg of the huge discussion about the metaphysics of artifactual kinds. However, it is not the interest of this work to examine it in detail. My aim is to show the main problems of the arguments principally from a semantic point of view.

originally intended to be a member. Still, he would not use 'glug' competently, since he would know nothing about doorstops. Nor would this be a genuine externalist argument, since he would be only deferring his judgment to humans.[34,35] Nelson imagines instead that we might have been fooled by malevolent aliens who made us believe that pencils are what we think they are, i.e. tools with the purpose of writing, marking and drawing, whereas they are actually devices by means of which aliens manipulate our activity. Were the same situation to involve Twin Earth – Nelson adds – and were we to visit the planet, we would say: "The things on Twin Earth that pass for pencils are really devices of control and manipulation". "We would not admit that there are any pencils on Twin Earth" (Nelson, 1982, p. 362). Nelson concludes, thus, that 'pencil' *is indexical* – for it is "associated with a set of properties in a way that is both epistemologically contingent and metaphysically necessary" (Nelson 1982, p. 362). However, despite the obvious problem of it being a skeptical argument, Nelson's defense is subject to a deeper objection: strictly speaking, it does not prove that pencils' properties are metaphysically necessary albeit epistemically corrigible. If we did not accept the argument of massive illusion, the device's function of manipulating humans would just be additional to the usual writing one. On the contrary, it seems that with natural kinds things go differently: once we acquire further information about their essence, such knowledge substitutes the previous one. The parallelism does not stand. Indeed, Nelson might even be further challenged. As long as we stick with the alien-induced delusion, once the real function of pencils is discovered we might want to produce instead "real pencils" (the ones that fit the description associated with the term), displaying a willingness to consider the definition predominant.[36] Last but not least, Nelson's is not a genuine argument from

[34] For a detailed discussion of objections to Kornblith's argument from ignorance see Marconi (2013, p. 506-507).

[35] The last objection can be traced back in Thomasson's critique of Nelson's argument (Thomasson 2003, p. 603). The objection can be generally formulated as follows: as long as the maker(s) of artifacts exist (or existed), it cannot be said that there is no expert(s) to defer to. As specified above about Kornblith's argument of deference, for an externalist argument to work, ignorance and error need to involve *all* speakers.

[36] This last objection, defined "persistence of description", is endorsed by Marconi (2013, p. 507).

error, since there is a group, the aliens, who in fact produce our pencil-control devices and that therefore know what they are for.[37]

Although Kornblith and Nelson seemed, *prima facie*, to be on the right track, their arguments turn out not even to satisfy requirement (b).

2.6. An Externalist Defense: Empirical Discoveries

In order to demonstrate what is asked for by requirement (b) one could emphasize the fact that arguments from ignorance and error only need to show that we could make empirical discoveries on the objects in question. This is Daniel Putman's strategy. In his "Natural Kinds and Human Artifacts" (1982), he states that, imagining that an alien anthropologist came to Earth to study our species, he would determine the extension of our artifactual kind terms "by a similarity relation pegged to a paradigm and underlying traits would become the subject of scientific investigation" (Putman 1982, p. 419). Consequently, the alien's statements on such traits, if true, would be necessarily true even though epistemically contingent, as they would be attempts at describing objective but easily mistakable features of *Homo Sapiens*'s patterns. Needless to say, this cannot count as discovering the objective properties which define the essence of an artifactual kind. Nor does it count as providing arguments from ignorance and error, since Putman gives no example in this sense.

Putman insists that, from the alien's perspective, "terms like 'joke' or 'game' or even 'neighborhood' and artifacts terms for tools, types of food, articles of clothing, etc. could either be natural kind terms themselves or be essential properties of the natural kind *Homo Sapiens*. Their status would in all cases be empirically determined" (p. 419). Still, Putnam's argument fails in its purpose. Showing only that we could make empirical discoveries about those terms, is not showing that their reference is governed by an underlying trait their *denotata* shares with the members of the same kind. Nor is it

[37] This is Thomasson's objection to Nelson's argument mentioned in footnote 25. In Thomasson's own words: "this is not a case in which makers are massively mistaken about the nature of an artifactual kind they create, but in which certain people mistakenly believe they are makers and thus have mistaken beliefs about which artifactual kind certain objects belong to (Thomasson 2003, p. 603).

showing that those referents possess such alleged essence with metaphysical necessity. At best, the argument shows that there are some objective patterns shared among our species, that the alien could be ignorant of or in error about.[38]

3. Conclusion

Above I have drawn attention to the fact that, in sum, all the authors who engaged themselves in defending Externalism about artifactual kind terms have unwarrantedly chosen, among those of Putnam's theory, one criterion over another, thus not providing a tenable account.[39,40] On the other hand, the arguments described from the descriptivist standpoint fails in providing real suitable alternatives.

[38] Crawford Elder (1989, 2007) argues for a quite similar position, to which similar objections can be moved. This is why I do not discuss his account in detail.

[39] Many further objections and observations could be addressed to all the arguments mentioned in this section, based on the metaphysical similarities or differences between the two kinds at stake. However, I would rather leave this discussion for another work.

[40] Marconi (2013) actually formulates a persuasive externalist argument about the kind of words being discussed. However, a very limited part of artifactual terms are concerned – defined "names of quasi-natural artifactual kinds" (p. 509). As evident from the definition, there is a very thin line between their extension and natural kinds. That is why I do not consider the argument in the present analysis.

Chapter 4

Towards a Hybrid Approach about Kind Terms

So far, I have pointed out that Putnam's argument about both artifactual and social kind terms cannot prove what it was purported to prove, because it fails to meet the two requirements a convincing externalist argument ought to fulfill. I have shown that meeting those requirements is all but a walk in the fields, and that none of the defenses of Putnam's suggestion have actually succeed in doing so. I will not try to meet this challenge myself. I think it is more interesting instead to consider the debate from a new angle, trying to identify the source of those problems to verify whether they can be avoided. In this chapter I will be concerned with showing that Putnam's account of reference clashes, in general – even when limited to natural kind words –, with several problems. In what follows I shall show that, conversely to what it is widely believed, also many other prima facie steady cases bring about rivalry intuitions. I will then advocate for alternatives approaches, both to modality and to the reference determination.

1. Conflicting Intuitions

Putnam's account of reference has become very popular and quite widely adopted, since it seems to handle very well a series of cases, as the examples of 'water', 'gold', 'tiger', 'elm', ' beech', etc. show. It accounts for some of ours intuitions, according to which there is a "mysterious" component in fixing the reference natural kind words, as well as for the fact that, despite our lack of competence in picking out the exact reference of the terms we use, we succeed in referring. As previously argued at a great length, Putnam suggests that also artifactual and social kind terms, proposal from which a flow of opposing theories and conflicting intuitions results.

Usually the controversy about Putnam's semantic theory is thought to originate just from such a proposal. However – as we shall see below – many other *prima facie* steady cases bring about rivalry intuitions.

'Cat' vs 'Key sparrow'. Putnam (1970, 1975) holds that, if we were to discover that all the cats on Earth are and have always been robots (remotely controlled from Mars), we would claim: "*Cats* have turned out to *be* robots". Accordingly, we would keep on applying the term 'cat' to all the objects to which we have always applied it, provided

that they share the same nature as the paradigmatic robotic cats.[1] So far so good, one might think.[2] But would we actually react in the way described by Putnam? Our intuitions might go completely towards the opposite direction – as the 'Key sparrow' hypothetical case exemplifies. "If ornithologists coin the term 'Key sparrow' to name a new race of sparrows apparently discovered to be living in the Florida Keys, only later to find that all of the supposed exemplars observed were sophisticated animatronics planted by a glory-hungry birdwatcher, it seems we *would* say there are no Key sparrows (since the things observed were no birds at all), not that it turns out that Key sparrows are little robots" (Thomasson 2007, p. 49).

'Okapis' vs 'red panda'. Let us also consider some real cases. Okapis are a species of large African ungulates. At first, when it was discovered, the species was named '*Equus johnstoni*'. However, right after it was found out that those animals were not related to horses but to giraffes, they were renamed '*Okapia johnstoni*'. This shows, against Putnam, that, after the discovery, we did not keep the name by which those animals were originally denoted, for we partly believed that the content connected to the term had to be privileged.[3] Conversely, in the case of 'red panda', the term was kept despite the fact that its referents do not display the properties originally connected to it: red pandas belong in fact to a peculiar family; however, they have more features in common with felines than with bears. This second case supports Putnam's theory instead.

Despite its initial persuasiveness, Putnam's Externalism seems to be threatened not only by examples in which artifactual or social kind words occur, but also by other cases like the ones mentioned, especially if – as it seems – he aims to convince us by means of intuitions. In fact, very controversial intuitions apparently arise from the hypothetical situations he reported. These is also evident if we consider claims in which empty names occur, like the following:

---

[1] This example might not be sound to many people. One might think to this as a circumstance in which we would rather say: "There are no cats", retaining the word's previous inferential content.

[2] The friend of Externalism might back up all these cases, arguing that Putnam's theory works somewhat always for words for natural substances and species, as long as they are at least originally *believed* to designate a natural kind (see on this Marconi 2009, p. 132). However, this interpretation runs into problems – as made evident by the case of 'ether' (see on this Marconi 2009, p. 234).

[3] This example is mentioned by Marconi (2009, p. 130).

'Witch'/'ether'. If Putnam was entirely right, it seems that we could not have found out that there are no witches. On Putnam's account, 'witch' would apply to whatever shares the same nature as the paradigmatic witches. But what is the nature of witches? On the contrary of what Putnam claims, the term seems to apply (if it applies at all) to whatever fits the definition connected to it. Since there is nothing, or better, no one who happens to fulfill the description associated with the word 'witch', we reached the conclusion that there are no witches. Also the name 'ether' was abandoned. Once it was discovered that there was *no* elastic, isotropic, homogeneous medium that fills the upper region of space, the name 'ether' was forsaken. Why was this so? Because there was no substance, or better, no unified cause of the observed phenomena. Both cases support the idea that, in order to make sense of claims like: "Ether does not exist" or "We believed that there was such thing as ether, whereas there is not", there must be some descriptive elements connected to the terms that we are privileging in grounding their extension. Otherwise, if – as Putnam claims – those extensions were entirely externally determined, we could not account for the inferential content still attached to those words. It seems that, following Putnam's theory, we should say that there is always something to refer to and that *that* something exclusively determines our words' reference. But, again, our intuitions conflict here.

If these examples are sound, we are left with a number of cases that Putnam's theory cannot explain or, at the least, which raise several conflicts. In what follows I aim to show that this is a consequence of the fact that Putnam implicitly relied on a hyperrealist view on modality (section 2). This view notoriously encounters several problems, that I will briefly sketch out (section 3). In order to avoid those problems, Amie L. Thomasson proposes an alternative approach to modality, Modal Normativism (section 4). I will highlight that adopting a normativist approach leads us to abandon Putnam's Externalism and to embrace a hybrid theory of reference (section 5). I will then conclude by showing that this new account of reference combined with Modal Normativism not only overcomes the semantic problems of Putnam's Externalism, but it also represents a more tenable and attractive theory of the semantics of our kind terms (section 6).

2. Putnam's Externalism and Heavyweight Modal Realism

As recalled in Chapter 1, § 2, in order to demonstrate his semantic view, Putnam resorts to some science fiction examples, meant to prove that some of our modal claims, such as "Necessarily, water is H_2O", or "Necessarily, pencils are artifacts", or "Necessarily, pediatricians are humans specialized in the care of children", and so on, are metaphysically necessary albeit epistemically contingent. This means that, if they are true, they are necessarily true, but they are always epistemically corrigible. In what follows I aim to show that, on closer inspection, these examples and the semantic theory Putnam derived from them give evidence that he implicitly relied on a hyperrealist view of the metaphysics of modality, the one defined by Thomasson as "heavyweight modal realism" (*forth.*). According to this robust metaphysical approach to modality, there are modal facts out there in the world that are genuinely discoverable through substantive investigation. "There are distinctively modal properties or facts that *explain* what it is that makes some modal statements true" (Thomasson, *forth.*). This traditional account is a descriptive one, according to which our claims of metaphysical necessity are attempts to describe modal features of the world. Let us collect proofs thereof.

Putnam grounds his semantic theory on such a heavyweight account of modality. He aims to convince us that, first of all, the referents of our kind terms possess their nature essentially. Given that water's chemical formula (i.e. its nature) is H_2O, nothing with a different nature counts as water.[4] If we agree on that, then – Putnam underlines – we would be willing to apply the term 'water' just to whatever has such a nature. His famous Twin Earth thought experiment was conceived for this purpose – to drive our intuition in this direction. He underlines that we would not say that the same word refers to both H_2O and XYZ, because we would rather apply 'water' only to whatever has the same essential features (i.e. the same nature) of what we had paradigmatically indicated by that term (namely, the substance which has chemical formula H_2O). This proves that specifying the hidden structure of a certain set of referents never corresponds to giving a description that ultimately and uniquely pick out the word's extension. On this view, it is a metaphysical fact that water, or any other natural substance, has the nature it has: a

[4] In Putnam's own words: "Once we have discovered that water (in the actual world) is H_2O, *nothing counts as a possible world in which [the substance called] water isn't H_2O*" (Putnam 1975, p. 233).

truth which entirely depends upon how the things in the real world actually are. Putnam's theory implies that that water is H_2O is a modal fact that exist out there, waiting to be discovered. This modal facts is what explains the truth of "Necessarily, water is H_2O". The same reasoning is valid of real cases such as 'gold', 'elm', 'tiger', and so on. That very phenomenon applies also to artifactual and social kind terms – Putnam argues. "The use of […kind] words reflect an important fact about our relation to the world: we know that there are kinds of things with common hidden structure, but we don't yet have the knowledge to describe all those hidden structures" (Putnam 1975, p. 244). As mentioned in the introduction, Putnam holds that if we found out that all pencils are actually organisms, we would say that *pencils*, those very objects, turned out to be organisms. Similarly, in the case of pediatricians turning out to be Martian spies. He insinuates that, in both circumstances, we would have eventually discovered the very nature of the kinds denoted by those terms.

In a nutshell, what Putnam meant to prove is that although it is imaginable, and thus epistemologically possible, that water, or pencils, or gold, and so on, turned out to have a different essence from the one they actually have, since water is H_2O, pencils are artifacts, gold has atomic number 79, and so on, this must be so: this *is* the essence of the relevant kinds.

Putnam explains that this is so, because the reference of our kind words is fixed by an ostensive definition (e.g. "Call this stuff 'water'"), which respects an *empirical presupposition*[5] (cf. Ch. 2, fn 15). Paraphrasing it, we can derive the following principle:

(**EP**) For every x (an object), 'P' (a term) applies correctly to x if and only if it shares the same nature of most of the things that are paradigmatically called 'P'.[6]

[5] Quoting Putnam's own words regarding the Twin Earth thought experiment: "My "ostensive definition" of water has the following empirical presupposition: that the body of liquid I am pointing to bears a certain sameness relation (say, x *is the same liquid as* y, or x *is the same*$_L$ *as* y) to most of the stuff I and other speakers in my linguistic community have on other occasions called 'water'" (Putnam 1975, p. 225).

[6] It is noteworthy that Putnam's Externalism is not a is not a mere instance of a paradigmatic case argument, according to which: for every x, to be paradigmatically called 'P' suffices to be correctly called 'P'. Fulfilling the empirical presupposition is a further, and yet fundamental, requirement. To share the

In other words, the extension of our kind words is fixed by an ostensive definition pegged to a paradigm. Sharing the same nature with the paradigmatic referents of a certain kind word is indispensable for that word to be correctly applied and re-applied to an object. The reference of a kind term is governed by the *real nature* of the things paradigmatically indicated by the ostensive definition.[7] The term refers to whatever has that nature.

Thus, as highlighted many times already, a given kind word has an externalist semantics if and only if:

a. it refers to things that have a nature (necessary features) possessed by all of the members of the term's extension, and only by them. The essence of its referents is what actually determines the word's extension.

At this point, one can legitimately object that, on the base of the considerations drawn so far, Putnam is not committed to any specific view on modality. Showing that the extensions of our kind terms are singled out by their essences is not peculiar only to a hyperrealist approach.[8] However, we shall remember that, according to Putnam's Externalism, there is another crucial requirement that must be fulfilled:

b. linguistically competent speakers can be ignorant of, or mistaken about such nature.[9]

---

same nature as most objects that are ordinarily called 'P' is necessary to be a P. See on this Marconi (2009b pp. 125-128).

[7] In Putnam own words: "the ostensive definition conveys what might be called a defeasible necessary and sufficient condition: the necessary and sufficient condition for being water is bearing the relation $same_L$ to the stuff in the glass; but this is the necessary and sufficient condition only if the empirical presupposition is satisfied" (Putnam 1975, p. 226). If the stuff in the glass is not water, the ostensive definition ought to be discarded.

[8] Also forms of conventionalism are compatible with the theses outlined. Alan Sidelle (1989), for instance, defends this view, claiming himself that: "nothing in any possible situation will count as water if it does not have that microstructure" (Sidelle 1989, p. 48). However, Sidelle's view does not allow for the total ignorance or massive error implied in Putnam's theory.

[9] This requirement follows from the notion of *theoretical relation* stressed by Putnam: "Whether something is or is not the same liquid as this may take an indeterminate amount of scientific investigation to determine. Moreover, even if a "definite" answer has been obtained either through scientific

Although our epistemic knowledge about our kind terms' referents increases during time, our knowledge is never final. The nature of those referents can be totally obscure to the speakers. *All* the speakers (including *expert speakers*) can be wrong about or even utterly ignorant of the nature of the things to which they refer, and consequently about the actual extension of the words they use. Nevertheless, it is that nature of the objects denoted by those terms that actually determines the term's extension. This second requirement gives evidence that Putnam's semantic theory of our kind words is grounded on a robust metaphysical realism about essences. If it was not so, Putnam would agree that specifying the hidden structure equals to give a description that ultimately pick out the word's reference. Or at least he would agree that the reference is given by our up-to-date definitions. On the contrary, – he stresses – "it is only by confusing *metaphysical* necessity with epistemological necessity that one can conclude that, if the (metaphysically necessary) truth-condition for being water is being H_2O, then 'water' must be synonymous with H_2O" (Putnam 1975, p. 240).

This commitment to the possibility of ignorance is precisely what shows that Putnam cannot be assuming anything like a conventionalist view. He is not indeed simply showing that the referents of our kind terms necessarily have an essence, as a conventionalist can perfectly accept or even claim.[10] Putnam endorses a much stronger view: that this essence could be totally unknown to the speakers, even though it is that essence that determines the extension of our kind words.[11] Our epistemic access to the

---

investigation or through the application of some "common sense" test, the answer is defeasible: future investigation might reverse even the most "certain" example" (Putnam 1975, p. 225).

[10] As mentioned in the previous footnote, Sidelle (1989), for instance, is a defender of this view. However, according to his view "there is no [real] necessity 'out there'" (p. xi). As pointed out in this section, this has nothing to do with Putnam's approach. Sidelle's view could not be more distant, insofar as in holding that "the only empirical fact that needs to be 'plugged in' to the sort of convention he [the conventionalist] proposes is that enough of the instances to which a term is applied in fact share a common feature (for example, a peculiar microstructure)". "…but this fact is posterior to, and explained by, our conventions (plus nonmodally packed independent facts" (p. 48). More specifically, Sidelle precises: "The task, then, for the conventionalist is to say how this actual application [of a term] is determined in some way that does not require antecedently given, modally extended item" (p. 47).

[11] For further proof, consider the following quote: "Suppose, now, that I have not yet discovered what the important physical properties of water are (in the actual world) – i.e., I don't yet know that water is H_2O. […] the operational definition, like the ostensive one, is simply a way of pointing out a standard –

referents of our kind words is limited, nevertheless their extension is always determined by the actual nature of their specific referents. What we can do – Putnam suggests – is point out a standard, but it can always turn out that we were radically wrong about the nature of that selected standard – as in particular the examples of pencil-organisms and robotic-cats are supposed to show. "Human intuition has no privileged access to metaphysical necessity" (Putnam 1975, p. 233).

The two requirements are summed up in Putnam's explanation of the notion of meaning, which gives further evidence of the robust realist view of modality on which Externalism is based. Putnam highlights that all kind words have an unnoticed indexical component, which is also the crucial one in determining their reference.

In sum, Putnam seems to hold that the world is essentially full of modal facts, our task is to discover them, where that is possible.[12] This is the core of Putnam's heavyweight modal realism. There are plenty of modal facts out there in the world, we just do not have epistemological access to all of them, or perhaps to none of them. Nonetheless, those modal facts are what actually make our modal statements true and determine the extensions of our kind terms. This derives, once again, from the weight he put on the indexical component of the meaning. Once a word has been indexically introduced, what counts as "the same kind as" the paradigm will be a matter of empirical research and scientific theorizing.

3. The problems of Heavyweight Modal Realism

The heavyweight approach to modality described in the previous section, which Putnam relied on for his semantic Externalism, notoriously encounters several problems. I will here briefly sketch out those problems, in order to then come to portray the approach proposed to deflate them (section 3).

---

pointing out the stuff *in the actual world* such that for x to be water, *in any world*, is for x to bear the relation $same_L$ to the *normal* members of the class of local entities that satisfy the operational definition" (Putnam 1975, p. 232).

[12] As Thomasson puts it: "[the metaphysician] care[s] about the real modal properties of things in the world […] the metaphysician sets out to discover these category of entity and to determine the modal features of entities of different categories" (Thomasson 2013, p. 144).

If we think, like Putnam, that there are discoveries to make about what really exists and that there are modal facts out there that *explain* the truth of our modal statements, we are discouragingly left with a number of problems. Most familiar are the epistemic problems of how we are supposed to have access to those relevant ontological facts about modality. It seems that modal facts cannot be perceived or empirically detected – as David Hume already observed long ago.[13] How do we discover what modal properties an object has or what modal facts obtain?[14] How do we discover that water is necessarily H_2O or that pencils are necessarily artifacts (assuming we allow that this could be a discovery)? In short, the hyperrealist view notoriously has difficulty with saying how we can have epistemic access to those modal facts.[15]

A plausible immediate objection might be that Putnam could simply reply by highlighting that the scientific theories tell us that the essence of water is H_2O, that tigers have a certain DNA, and so on; or that human sciences tell us what are the essences of artifacts or of social and institutional kinds.[16] However, this response is exactly what Putnam tried to avoid. Or better, he enhanced that, even so, the statements expressing truths on the nature of those objects are never necessary in the epistemic sense. Putnam meant indeed to defeat the traditional idea, according to which all kind words refer to "whatever satisfies the *contemporary* [italics mine] 'operational

[13] D. Hume (1777/1977) argued that we cannot be thought to know that any matter of facts holds necessarily on the basis of the experience of the world, for however well a statement may be conformed through experience, that only shows that it does (so far) hold, not that it *must* hold (cf. pp. 39-53, and Thomasson *forth.*, Ch.1).

[14] For a discussion of this problem and relative references see Thomasson (*forth.*, Ch. 1).

[15] Actually, Barbara Vetter (2015) and Tim Williamson (2008) have tried to give a plausible story about how we can come to know modal facts. Nonetheless, their solutions do not improve the discussion here at stake, since they do not show how we can know the distinctively *metaphysical* necessities of interest in the present debate. Vetter anchors the possibility of acquiring modal knowledge to the modal properties of individual objects such as potentialities or dispositions (i.e. the fact that a glass is fragile, that paper can burn, etc.). However, she does not talk about how we can acquire knowledge on essences (about the fact that water is H_2O or about whether the pencil-organisms *are* pencils). Williamson's work is concerned with the truth or falsity of counterfactuals – he aims to avoid the epistemological problem by claiming that we only need the capacity of evaluating those counterfactual conditionals. But, again, this does not show anything about essences.

[16] This is the strategy adopted, for instance, by Crawford Elder (1989, 2007) and Daniel Putman (1982) in order to claim that semantic Externalism applies also to the mentioned kind of words.

definition'" connected to them (Putnam 1975, p. 235), for it "does not provide a necessary and sufficient condition for the application of any such word[s]" (p. 238). What matters for the determination of our kind words' reference is how the things actually are, not our epistemic knowledge about them. As mentioned in the previous section, one of the crucial notions Putnam highlights in defending his theory is the one of *theoretical relation*, according to which even the most certain answer "obtained through scientific investigation" is defeasible. Our method of identifying those natures has grown incredibly sophisticated, so that we have been enabled to accurately update our definitions according to the empirical discoveries. Nonetheless, the extension of the words we denote has not changed accordingly.[17] Putnam would not accept and say that those natures of the objects we refer to are given "intra-theoretically".[18] Those descriptions empirically given within the contemporary scientific theory provide only the stereotype[19] associated with the word that denote those object, but it can never serve to determine the word's extension.[20] We would not say – Putnam would insist – that all tigers are striped, or that a three-leg cat is not a cat, only because the operational definition, as much as accurate it can be, does not include exceptions. What expressed by those statements is never epistemologically necessary. Therefore, all the epistemological difficulties mentioned above still stand.

[17] This is why – Putnam stressed – we ascribe to the word 'gold' the same extension both before and after 1750, i.e. both before and after the discovery of chemistry. Even though before 1750 we did not have any clue as to distinguish gold from iron pyrites, we would not say that 'gold' referred to gold and fool's gold, whereas now 'gold' refers to gold or whatever has atomic number 79.

[18] A significant quote in this respect is the following: "It is beyond question that scientists use terms as if the associated criteria were not *necessary* and *sufficient conditions*, but rather *approximately* correct characterizations of some world of theory-independent entities, and that they talk as if later theories in a mature sense were, in general, *better* descriptions of the same entities that earlier theories referred to. […] My point is that if we are to use the notions of truth and extension in an extra-theoretic way (i.e., to regard those notions as defined for statements couched in the language of theories other than our own), then we should accept the realist perspective to which those notions belong" (Putnam 1975, p. 237).

[19] In Putnam own words: "a standardized description of features of the kind that are typical, or "normal", or at any rate stereotypical. The central features of the stereotype generally are criteria – features which in normal situations constitute ways of recognizing if a thing belongs to the kind, or, at least, necessary conditions (or probabilistic necessary conditions) for membership in the kind" (Putnam 1975, p. 230).

[20] As Putnam highlights: "We may give an 'operational definition', or a cluster of properties, or whatever, but the intention is never 'to make the name *synonymous* with the description'" (Putnam 1975, p. 238).

The traditional descriptive approach also faces what Huw Price has called the "placement problem" (2011), the problem of whether we should accept that there are modal facts, or properties, or essences and how they are supposed to fit in the natural world. Modal facts and properties do not seem to be in any way like those facts and physical properties studied by empirical sciences. It is difficult then trying to figure out what those modal facts and properties should be and how they are related to the non-modal facts and properties investigated by the natural science.[21] This is also a version of the so called "grounding problem".[22] Modal properties seem not to supervene on non-modal properties: for example, a statue and a lump composed all of the same parts and material at the same time have all of their microstructural and relational properties in common; yet, they are said to differ in their modal properties: the statue cannot survive being squashed, whereas the lump can. So, what can possibly ground their alleged modal differences given that they are otherwise so alike? As Karen Bennett points out, the heavyweight modal approach seems to have a hard time answering this question. [23]

In sum, the descriptivist account of modality leads to a mysterious epistemology and ontology.[24] It seems then that taking claims of metaphysical necessity as attempts to describe certain properties or facts of this or other possible worlds is the source of deep problems with very little agreement about possible solutions. This is what motivated some to seek a different approach, Modal Normativism, which is the focus of the next section.

---

[21] For a discussion of this problem and relative references see Thomasson (*forth.*, Ch. 1).

[22] M. B. Burke (1992), K. Bennet (2004), K. Fine (2008) and L. deRosset (2011) provide excellent discussion of it.

[23] Karen Bennett (2004) examines this problem in detail, concluding that: "nothing purely non sortalish determine which sortalish properties [persistent conditions, sortal properties, any properties dependent on the first two kinds] either of them [the statue and the lump] has. […] the non sortalish properties that constitute their subvenience base cannot be held up as an answer to the grounding problem" (Bennet 2004, p. 344). If this is so, the heavyweight modal realist is left with further epistemological difficulties.

[24] The descriptivist account leaves us also with a crisis in methodology. Since, following heavyweight modal realism, we cannot provide a good explanation about how we can acquire modal knowledge, this leaves us without a good story about what methods we can apply in order to investigate the questions in metaphysics (cf. Thomasson *forth.*, Ch. 1, 7). I will not analyze further these problems, since they do not affect Putnam's theory.

4. An Alternative Approach to Modality

In order to bypass the criticisms mentioned in the previous section, Amie Thomasson proposes an alternative account of modality: Modal Normativism (2013, *forth.*). This view denies that modal statements are descriptive at all.[25] The basic function of talk about what is *metaphysically* necessary is absolutely not to try to describe modal features of the world. It is, rather, to provide a very useful way of expressing constitutive semantic and conceptual rules[26] in the object language.

In Thomasson's own words:

"...properly understood, modal claims are not attempted descriptions of special modal features of reality at all. [...] modal statements of necessity are (attempted) object-language expressions of constitutive semantic rules" (Thomasson 2013, p. 145).

This new view, in opposition to a standard descriptive account, does not posit modal facts as truth-makers of our modal propositions. On the normativist view, modal terms of metaphysical modal claims have the function of expressing rules or norms in the indicative form. Statements apparently about modal properties or facts have a completely different function from the one assigned to them by the traditional approach, namely to track what really is out there in the world. The most basic function of what pass as basic claims of metaphysical necessity is rather to *instruct* and *correct* those who are unsteady in using the terms. However, those "basic modal claims are expressions in the object language, and thus are world-oriented; they are not

[25] Thomasson backs up her non-descripritivist approach to modality in the work of the early 20th century analytic philosophers such as Schlick, Wittgenstein, Carnap, Ayer, Sellars and Ryle. She argues that this approach was mistakenly abandoned for the wrong reasons. Thomasson resuscitates it, making it a view that avoids the most important criticisms which those accounts were commonly thought to run into (see Thomasson *forth.*).

[26] Constitutive semantic rules are rules the subjection to which is constitutive of using the term (as opposed to a different homophonic term) at all. They are on a par of the constitutive rules of, for instance, social institutions or games (see on this Searle 1995). "Just as constitutive rules of games are rules that must be in force for that game to be played at all, so must constitutive semantic rules be in force for that very term (or concept) to be used at all" (Thomasson 2013, p. 147).

descriptions that aim to describe our language, concepts, or rules. […] they are not descriptions at all" (Thomasson *forth.*, Ch. 2).

Taking the normativist approach brings several advantages, one above all is demystifying the epistemology of modality. Modal Normativism indeed considers the move from using language to knowing basic modal facts as a matter of moving from "*mastering* the rules for properly applying and refusing expressions (as a competent speaker), to being able to explicitly *convey* these constitutive rules in the object language and indicative mood" (Thomasson *forth.*, Ch. 6), though it is not required to any speaker to be able to recite those rules even if she mastered them (as happens with grammatical rules).[27] According to this approach, claims of the form "Necessarily p" are used "to convey constitutive semantic rules in the object-language indicatives and may be used to *correct* or *condemn* uses that violate the rule, but again without *reporting on* the linguistic rules or on the statement *p*" (Thomasson 2013, p. 147). "Necessarily, all bachelors are men", for instance, may be used to correct someone who believed that her aunt is a bachelor.

On this account, we are licensed to add 'Necessarily' not only to any analytic claim (like the previous one), but also to any object-language expression of an ostensive definition, such as "Call stuff of this chemical kind 'water'" and get "Necessarily, stuff of this chemical kind=water".[28] No special features of the world are then involved in order to make true those metaphysical claims; there is no need of those special features or facts that serve as corresponding to those ones supposedly tracked by modal discourse. Once the descriptivist assumption is given up, "we also lose the assumption that there must be a causal relation to the relevant facts or features to legitimate the claim to have knowledge of them" (Thomasson *forth.*, Ch. 6).

In order to acquire modal knowledge we can consider the multiplicity of actual or imagined cases in which the terms would properly be applied or refused, so to acquire a more explicit grasp of their underlying constitutive semantic rules. "So, for example, we can engage in semantic descent to move from "'Bachelor' must only be applied where

[27] Those constitutive semantic rules of using terms are conditions for their application and co-application. Application conditions establish the circumstances in which a term is to be applied or refused and the co-application conditions settle down when a term may be applied to one and the same entity (see on this Thomasson 2007, Ch. 2). I will also come back to this in the next section.

[28] On the function and meaning of modal discourse see Thomasson 2013, pp. 146-152 and *forth.*, Ch. 3.

'man' applies" to "Bachelors must be men", or (more clearly avoiding scope ambiguities) "Necessarily, whatever is a bachelor is a man". We would then also be able to instruct and correct those who are unsteady in wielding the terms. By using this sort of thought experiments we can gain a more systematic and explicit grasp of what actually are those constitutive rules. In Thomasson's own words:

> …it seems that considering 'what we would say' in case of surprising empirical facts tells us not that we should reject the idea that there are rules, but rather that we should understand them more subtly, as often involving conditionalization on various empirically discoverable features of the world. Rules of all kinds can be conditionalized on empirical conditions [Thomasson *forth.*, Ch. 4].

Those rules can be combined either with each other or filled in by empirical facts and discoveries. Conditionalizing the rules then will provide the needed disambiguation, while still leaving it open what to do should the relevant empirical conditions not turn out as expected. Focusing more precisely on the claim 'Water is necessarily H_2O', the constitutive semantic rule governing the term 'water' – according to Thomasson's theory – should have the following conditional form: *if the relevant empirical application conditions (for a chemical compound term) are fulfilled*, the term has to refer to a certain chemical compound (with a certain microstructure). From this follows that the modal fact will be explained as: *if water exists, and is a chemical compound with a certain microstructure*, then water necessarily has that microstructure.[29] Therefore, we can conclude that Thomasson's normativist approach with its conditional rules leave open the possibility that the relevant empirical conditions turn out not to be

[29] Actually this rule is not explicitly stated, but may be reconstructed from the material given in the chapter. The scheme could be the following:

- Schematic rule: 'Water' is to be applied only to whatever has [whatever microstructure the originally baptized stuff has];
- Empirical fact: The originally baptized stuff has the microstructure H_2O;
- Filled-in rule: 'Water' is to be applied only to whatever has the microstructure H_2O;
- Object-language expression: Water necessarily has the microstructure H_2O.

as expected and so that, for instance, water turned out to be XYZ (taken as en epistemic possibility).[30]

This gives a precious contribution to the knowledge of *derivative* modal facts. For example, it can be a constitutive rules of the term 'statue' that its referent cannot survive being squashed; but a derivative, empirically discoverable, consequence thereof that its referent cannot survive a certain pressure.[31]

Following this, the advantage over Putnam's approach is immediately clear. Both Thomasson and Putnam would use actual and imaginative cases in order to work out the rules of the application of our kind terms. However, the difference in their underlying reasoning is substantial. Putnam tried to ground the fact that, for instance, we would not be willing to apply the term 'water' to a liquid that does not have chemical formula H_2O by appealing to the modal fact that exists out there that necessarily water has *that* hidden structure. However – as pointed out in the previous sections – this move is very problematic.[32] By contrast, on the normativist approach, those intuitions that Putnam is calling on would play a major role in making explicit which semantic rules we follow when we use a term such as 'water'. Hence – as pointed out above –, a claim such as "Necessarily water is H_2O" may be used to correct or condemn uses that deny its content, uses that deny that water is H_2O. Such a modal claim can be used for instance to correct a speaker who arrived on Twin Earth and mistook the liquid that constitutes its seas, rivers, lakes, etc. for water. Nevertheless, the relevant constitutive rules can be world-deferential, in the sense that, for example, 'water' is to refer to "whatever has the same chemical structure as this (the paradigmatic sample)", leaving for empirical discovery what that structure is. This approach, hence, can still very well account for those intuitions which have favored the proliferation of Externalism.

[30] This normativist view can then perfectly account for *de re* and *a posteriori* necessities, which usually challenge deflationists approaches.

[31] For a detailed discussion of those advantages see Thomasson (*forth.*, Ch. 1, 6).

[32] This solution brings straight to the epistemic problems of such descriptive account. It does not seem that we should let those thought experiments to play a major role in explaining how we can acquire modal knowledge. Thought experiments are certainly not used in deciding among scientific theories. If we think, like Putnam, that there are facts to be discovered, how can we let our intuitions or abstract cases decide? For a deeper discussion of this problem and relative references see Thomasson 2017.

We can therefore see how the normativist can better handle the epistemological problem, giving a credible account of how we come to know modal truths, while also preserving our relevant intuitions. "What it is for a claim of metaphysical necessity to be true is for it to give an object-language expression of an actual semantic rule (or consequences thereof). And we can come to know the truth of basic claims of metaphysical necessity by starting with semantic competence and coming to gain explicit knowledge of the actual semantic rules involved, and/or by reasoning through their consequences" (Thomasson *forth.*, Ch. 6).[33]

Adopting modal normativism also brings ontological advantages, as on this view modal claims do not need truth-makers at all. "By contrast [with the traditional modal realism], on the normativist view, modal facts and properties are not 'posited' to 'explain' what makes our modal statements true" [Thomasson, *forth.*, Ch. 5]. The entitlement to add 'necessary' to a basic claim arises from its giving object-language expression of a semantic rule. "We do not need to appeal to a property of necessity in the world to explain its truth. On the normativist approach, talk of modal properties or facts is explanatorily downstream from modal truth" [Thomasson, *forth.*, Ch. 5]. What we gain is not an *explanatory* realism (as is the one of Putnam), but rather a *simple* realism. Accordingly, we are entitled to say that those modal facts and properties called on in the modal discourse do exist and we can even *describe* them (depending on the function of the modal term involved)[34], but this does not mean that we appeal to them or that they are required to *explain* the truth of our modal statements. The questions provoked by the "placement problem" or the "grounding problem" mentioned in the previous section are simply out of place here, "for though 'modal fact' and 'modal property' are noun terms, they enter our language with an entirely different function than tracking some wordly features such that we might want to investigate what it is really like or how it can arise out of lower-level wordly features" (Thomasson *forth.* Ch. 6). However – as specified above – this a view in which there are modal facts and in

[33] The normativist account provides then a good solution to what Christopher Peacocke (1999) has called the "integration challenge": reconciling a plausible account of what it is for a statement of a given kind to be true with a convincing account of how we can know them. (Cf. Thomasson *forth.*, Ch. 6).

[34] On this normativist approach, we distinguish between indicatives that function to *describe*, *track* and *co-vary with* elements of our environment, and indicatives that serve other functions such as *conveying rules* or *norms* (cf. Thomasson *forth.*, Ch. 5).

which the relevant modal facts are mind-independent and objective. Their truth does not depend on the existence of our language, conventions and minds.[35]

Above I have sketched out the alternative approach to modality and how it brings with it ontological as well as epistemological advantages, of simplifying and demystifying various issues, like the ones mentioned in the previous section.[36] If we accept this view, holding that it is more tenable and also more attractive than the traditional heavyweight modal realism, this leads to the consequent dismissal of semantic Externalism and to the acceptance of a hybrid theory of reference. I will now turn to show such connection.

5. Towards a Hybrid Theory of Reference

In the previous section we have seen how on a modal normativist account basic modal claims are object-language expressions of constitutive rules of use of our terms. This presupposes that there are rules of use which are meaning constituting. If we embrace this view, we are then committed to saying that Putnam's Pure Externalism is false, for we cannot accept that the meaning of our kind terms is entirely externally determined. Moreover, embracing this new approach to modality enables us to adopt a hybrid semantic theory, which – as we shall see – can better handle the controversial cases raised within Externalism.

As highlighted in the previous section, on a modal normativist account, "the basic function of talk about what is 'metaphysically necessary' is […] to provide a particularly useful way of expressing constitutive semantic and conceptual rules in the object language (Thomasson 2013, p. 145). Those rules (which involve some conceptual content) can take the form of application and co-application conditions that:

[35] For a detail discussion of the response to the ontological problems and also to the criticism of conventionalism see Thomasson *forth.*, Ch. 3, 5.

[36] Additionally, this normativist account of modality helps in clarifying the methodology of metaphysics. A form of conceptual analysis must play the primary role in answering metaphysical modal questions. In this sense, intuitions have a justified role in making explicit the semantic rules we follow, and thereby in discerning specifically modal facts that corresponds to those semantic rules. As mentioned in footnote 29, I will not press this point here, since Putnam's heavyweight modal realism does not directly run into it.

– establish under what circumstances the attempted grounding would or would not be successful in fixing reference (*application conditions*);

– specify under what circumstances the term would be applied again to one and the same entity (they provide some identity conditions for the thing(s) (if any) denoted (*co-application conditions*)).[37]

The normativist approach is the modal framework that commits us to the adoption of a hybrid rather than an externalist theory of reference. According to a hybrid theory of reference there must be some implicit but crucial descriptivist elements in grounding a term's reference (at least the category of objects the word refers to is determined by us as a linguistic community).[38] The core of the theory can be summarized like this: "…reference is only unambiguously established to the extent that our nominative terms are associated with a high-level conceptual content establishing what category of entity is to be referred to by the term, if it refers at all" (Thomasson 2007, p. 38). The connection with the normativist view of modality is quite clear, as well as the distance from the traditional descriptive approach. If we espouse Thomasson's Modal Normativism, we are committed to the thesis that there are rules of use of our (kind) terms that are meaning constituting. This is in utter contradiction with the basic idea of Putnam's semantic Externalism, according to which the meaning of our kind terms is determined by the objective relation between the world and the speakers. Such a rivalry between the two theories is even more evident by recalling the two fundamental requirements of Externalism (cf. section 2). The modal normativist cannot accept the externalist thesis according to which we can always be utterly ignorant or mistaken about what our kind words actually refer to – about the nature of all the things that surround us and that we denote.[39] If we are convinced by the theoretical virtues of the

[37] As Thomasson puts it: "Working out the application conditions for our sortal terms gives us the means for knowing the most basic (conceptually but not empirically relevant) existence conditions for members of the kind. Discerning the co-application conditions for our sortal terms gives us the means for reading off the most basic identity and persistence conditions of members of the kind" (Thomasson *forth.*, Ch. 6).

[38] Such a theory governs the application also of proper names. I will here be concerned only with kind terms.

[39] This can also be seen as the *problem of the opacity of the semantics*. One might object that, if we adopt Pure Externalism, the semantics of every given term will always be opaque, for we could never have any a priori criterion to decide which semantics applies to it. If we completely espouse this theory, it seems

normativist view, we ought to reject semantic Externalism together with the heavyweight modal realism which grounds it. On the other hand, we are inclined towards a hybrid theory of reference, a theory already proposed in order to avoid the problems of Pure Externalism, which nonetheless preserves most of the intuitions that motivated the wide adoption of Putnam's account.

Such a hybrid theory of reference, put forth by Devitt and Sterelny (1999) and Thomasson (2007b), arose from the idea that, in grounding a kind term's reference, there must be some peculiar features of the set of objects to which that term is supposed to refer that are to be considered more relevant than others. It seems that even when we use our kind terms (natural, artifactual or social) in ostensive contact with the sample of the appropriate kind, we already have in mind which relevant features we need to look for in order to pick out a certain kind term's extension rather than an other's.[40] If the reference grounding is to succeed, some conceptual concept must play a role in picking out the extension of, say, a natural kind word *qua* member of a natural kind (and most of the times, even *qua* member of a particular natural kind). For example, if we are to coin a new term for certain things right in front of us, we need some guidance at least regarding which of the various different kinds it belongs to (natural, artifactual or social). Most of the times we even need some further conceptual content to tell us which are the relevant properties to seek if we are to refer, for instance, to a mammal rather than simply to a vertebrate.[41] On the contrary, if we follow Pure Externalism, it will be

---

that we would be never able to say which semantics a given term has, for we may never find out what it actually designates. We would always need some empirical research to tell us what every term actually denotes. Think about the example of pediatrician-Martian spies or to the term 'ether': we do not have any clue to their semantics until we actually conduct empirical research on most of their referents. Therefore, it seems that we may never be able to tell what semantics applies to a given word, unless we make empirical research about all the referents the word is supposed to refer. However, this may appear pretty counterintuitive. This objection was inspired by some of the considerations on Externalism pointed out by Marconi (2009b).

[40] In Devitt and Sterelny's own words: "whatever relevance ostensive contact does have for these other terms [artifactual and social], it cannot have the same relevance as it does in the grounding of a natural kind term [Devitt and Sterelny 1999, p. 73].

[41] As Devitt and Sterelny put it: "In sum, the grounder of a natural kind term associates, consciously or unconsciously, with that term, first some description that in effect classifies the term as a natural kind

radically indeterminate what the term refers to, since we do not have any way to disambiguate to which entity it is to refer to, if it refers at all. This is commonly known as the "*qua*-problem".[42] It seems that, contrary to what Putnam holds, to determine what a kind term refers to is not something that we can do by deferring exclusively to the world. We need to accept that our words have at least certain kind of conceptual content to disambiguate whether or not a name is successfully grounded (and comes to refer) and distinguish which entity (of variously empirically indistinguishable entities) it is to refer to, should it succeed in referring. This would also avoid what has been labelled as the "problem of handling non existence claims", of which Putnam's theory is accused.[43] According to his Pure Externalism, indeed, it would not only be indeterminate to which entities a given word is to refer, but also whether or not it successfully refers. As pointed out in § 1 (cases 3), he leaves us with no explanation of the semantics of those empty names such as 'ether' or 'witch', since there is no concrete paradigm to which their reference can be pegged, nor anyway to disambiguate what we ought to regard as the relevant features of the *denotata*. This problem vanishes when following a hybrid semantic theory, for at the basic level, it is the *conceptual content* that establishes what basic conditions are and are not relevant for the term to be applied and reapplied to one and the same object.

Nonetheless, all the stress put on the conceptual side does not implicate that there is no room for ignorance and error. Such conditions may also be highly *incomplete* – they certainly do not provide necessary and sufficient conditions for a term to be applied and reapplied. The constitutive rules that govern our kind terms' application are "*revisable* in light of both empirical discoveries and our purposes in using the terms" (Thomasson 2007, p. 52). Constitutive rules involve those conditions that are conceptually relevant to whether or not the reference is established, but not all the conditions that may be empirically discovered as relevant (the world-deferential ones). If we are to coin a new name, 'Orky', as the name for an animal (swimming in the environs), the attempt to

---

term; second, some descriptions that determine which nature of the sample is relevant to the reference of the term" (Devitt and Sterelny 1999, p. 74).

[42] For a detailed discussion of the *qua*-problem see Devitt and Sterelny (1999, Ch. 5) and Thomasson (2007b, Ch. 2).

[43] For a detailed discussion of the *problem of handling non existence claims* see Devitt and Sterelny (1999, Ch. 5) and Thomasson (2007b, Ch. 2).

ground its reference may fail if there was no animal nearby but only a patch of seaweed.[44] Further conditions relevant to establishing which exact class the object is part of may defer to the real world and empirical investigations. Those rules can indeed be *conditionalized on empirical conditions*, so that, for example, the name 'Key sparrow' applies if the relevant empirical facts are as expected, namely if all the referents which we usually denote by that term are animals (cf. case 1). The co-application conditions for the same term can also be revisable and made more subtle in light of further discoveries. In order to know if an artifactual term, say, a name for an artifact used in some ancient Indian tribe, applies to the same sorts of objects, might require some discoveries about those objects as well as about the customs of that tribe.

This goes hand in hand with the normativist approach to modality, which accounts for *de re* and *a posteriori* necessities by showing that the form that our constitutive rules can take can vary[45]:

- they can be *intra-language* or *language-language* rules: rules governing when a kind word is to be applied, stated in terms of when another kind term is to be applied. E.g. "All bachelor are men" gives us the rule the govern the application of a word like 'bachelor'.
- They can be *world-language* rules: rules governing when a kind term is to be applied, set up by ostensive definitions. These rules are to be completed and filled in by empirical information, which might be unknown to competent speakers. E.g. "Call *this* chemical kind 'water'".

It also noteworthy to observe that, even though the combination of application and co-application conditions provides the relevant disambiguation, hence grounding the circumstances under what something is correctly called 'P', this does not imply that competent speakers must be able to verbally articulate what those frame-level conditions for the relevant categorical term are – any more than it was requested with

[44] This example is due to Thomasson (2007, p. 39).

[45] According to this view, *de re* and *a posteriori* necessities can be seen as reflections of semantic rules combined with empirical "filler facts". Those constitutive rules may be schematic and world-deferential, waiting to be completed and filled in by empirical information. Such empirical "filler facts" may be unknown to competent speakers.

the constitutive rules involved in the modal talk (cf. section 3).[46] Competent speakers are only requested to associate a given kind word with the relevant category of objects, and so to understand what it takes to be a member of that category. This involves the ability "to apply and reapply terms of that sort properly, to evaluate whether or not they would be properly applied or reapplied in various situations, to rebuke others for misuse, and so on – not the ability to explicitly recite the relevant conditions" (Thomasson 2007, p. 44).

All the considerations laid out so far should show the close connection between Modal Normativism and a hybrid theory of reference, as well as the consequent refutation of Putnam's semantic Externalism based on a hyperrealist view of modality. Both at the level of the semantics and the level of modality, the constitutive rules serve the primary function of instructing and correcting those who are unsteady in employing certain terms. Room for ignorance and error is still left, but in establishing what basic conditions are relevant for those terms to be applied and reapplied a major role is played by the conceptual content. In the next section I will turn to show the several advantages that this new perspective has and how it better handle the controversial cases mentioned in the introduction.

6. Advantages

In the previous section we have seen how a hybrid account of reference can overcome some important criticisms that Putnam's Pure Externalism confronts – criticism such as the *qua*-problem, the problem of handling non-existence claims, and the problem of the opacity of the semantics. It seems that, contrary to what Putnam claims, we should not hold that the world alone determines what a kind term refers to. There must be something going on in the ostensive contact that makes some features more relevant than others in picking out the term's reference, *qua* reference of that word. Frame-level application and co-application conditions outlined by the alternative

[46] As Thomasson puts it: "…to say that language has constitutive semantic rules is not to say that these rules may all be stated in a metalanguage, nor that competent speakers must be capable of reciting these rules – instead, they must simply be masters at *following* them. […] it can't be the case that the semantic rules of use for all of our terms can be (informatively) stated verbally" (Thomasson *forth.*, Ch. 2).

theory of reference are fundamental in determining a kind word's extension. Espousing such a theory has many advantages, since it can not only account for the cases that played a major role in supporting Putnam's semantic approach, but it can also better accommodate all the conflicting intuitions that some of them raise – as we shall see below.

'Water', 'gold', 'lemon', 'swan' and the like terms.. A hybrid theory of reference can explain the semantics of those terms even better than Putnam's Externalism. It can indeed account for the fact that in grounding those terms' reference we know, first of all, that we aim to ground the extension of natural rather than artifactual or social kind terms; and it consequently tell us which specific sorts of features are relevant in each and every case. It tells us that we need to look for a certain chemical formula in the case of 'water', for the atomic number in the case of 'gold', for a certain biological structure in the case of 'lemon' as well as in the case of 'swan'. Nonetheless, all the rules that govern the application and co-application of such words are open to revision due to empirical discoveries, so that a hybrid theory of reference can still account for those intuitions that are externally-driven in this sense. Such a hybrid theory can perfectly explain why we would not say that a liquid with a different microstructure *is* water, or that iron pyrites *is* gold, or that fruits biologically unrelated to lemons *are* lemons, or why we would instead agreed that black swans *are* swans. All the empirical discoveries can be fully addressed by conditionalizing the constitutive semantic rules on empirical facts. Furthermore, when the frame-level conditions are not met, whether to keep or forsake the term, as in general what to do pragmatically speaking, is a matter of decision.

'Elm'/'beech'. Also such examples can be better handled by a hybrid semantic approach, since it can provide a good explanation of speakers' competence, while still avoiding the *qua*-problem. In order for a speaker to be considered competent and not only referring parasitically, she must grasp the frame-level application and co-application conditions. Yet, she is not required to be able to recite those conditions as well as to know *all* the relevant properties for membership in a certain kind (e.g. the distinguished features of elms as opposed to those of beeches). She can defer possible lacunas to the grounders of the reference of such words, who play the role of *expert*

speakers.[47] Still, those constitutive rules governing our kind terms' application can be filled in by empirical facts, but not all the burden is shifted on the world side, hence guarantying the advantages provided by such an account of reference.

'Pencil'. A hybrid theory of reference can definitely better explain the semantics of artifactual words. In such cases – as it was with the natural kind words examples mentioned above – we would know that we aim to ground an artifactual kind term and we would then seek the relevant characteristics in this respect. If – as Schwartz argues – we count among those properties only the superficial ones, so that we would apply the term 'pencil' indifferently both to some lookalike pencils which are a species of organisms and the usual artifacts made out of wood and graphite, this, again, would be a matter of decision based on several factors (empirical and conceptual). Moreover, a hybrid semantic theory can perfectly respond to those who might insist that empirical discoveries can be made also about artifactual kinds – as mentioned in footnote 5 with the case of cable wires. The cable-wires example proves indeed that, if – as it happened – an artifact created to serve a certain function turns out to additionally perform a different one, we can update our beliefs, the definition connected to its designating term, and we can decide whether to keep or abandon that very term according both to the empirical discovery and our practices and purposes.[48] As argued at length in the

[47] This phenomenon, called the "division of linguistic labor" was already theorized by Putnam (1975) in accounting for the fact that common speakers have to acquire words like 'gold', 'tiger', 'water', etc., but not the method of recognizing that the things to which they apply these terms are really gold, tigers, water, etc. They can rely on a special subclass of speakers: the *experts*, whose judgment they can trust in case of doubt. So, ultimately, this knowledge (the criteria for recognizing a term's extension) is possessed by the collective linguistic body, even though it is not possessed by each individual member of the body. However – as highlighted many times – the ultimate criterion to refer to in grounding a term's extension is the indexical component, i.e. the real nature of the things that serve as paradigms for that kind.

[48] Another interesting example in this respect is the one of vacuum tubes. This type of components that, thanks to an external power source, provide an amplified power signal, have been used in electronic equipments such as radio receivers and transmitters, televisions, and in general in all types of electrical signal amplifiers. Once it became clear that the waves emitted can be used to heat or bake food, it has become the essential component of microwave ovens. The new function dominates the one for which the object was originally created. Nonetheless, we still use the term 'vacuum tube' to refer to that object. There are other cases in which, on the contrary, the term was forsaken in favor of one which better fitted the inferential content initially connected to the term.

previous section, the application and co-application conditions can vary according to new empirical information, as well as to other relevant conceptual elements.

'Pediatrician'. The same holds in the case of social kind terms. This seems an example explainable in terms of intra-language rules (cf. section 4). One might think that such alternative account meets our intuitions better than the one theorized by Putnam, especially in these situations. It seems that we would not agree with Putnam in saying that we could discovered the very "nature" of pediatricians, since the word's application seems to be governed by totally different criteria than the ones which govern the application of 'water' for instance.

'Cat'/'Key sparrow'. As conflicting intuitions show in respect to these situations, "the right lesson seems to be not that reference survives any failure of associated basic application conditions (so that even the most basic application conditions must be empirically discovered) [– as Putnam suggests –], but rather that where such conditions fail, we have to make a *decision* about what to do with the term based on our surrounding practices and concerns" (Thomasson 2007, pp. 49-50). We might retain the term's original meaning but accept that the term fail to refer; or, we might continue to use the term despite the fact that the most basic, frame-level application conditions were not met. The very fact that cases could go either way suggests that this is a matter of semantic decision based on our interest and purposes, rather than discovery of modal facts out there in the world.

'Okapis'/'red panda'. The previous example is a hypothetical case of what actually happened with the names 'Okapis' and 'red panda'. In the former case, the name was forsaken. We privileged the inferential associations in governing its extension: that is why the term was dropped. In the latter case, on the other hand, the name was kept even though the relevant properties originally associated with the term were no longer fulfilled. Again, a hybrid account of reference can perfectly account for both situations, without running into any contradiction: the decision of keeping or abandoning those very terms has not depended on how the modal facts in the world are, but rather it has been modulated both on empirical discoveries and practical interests.

'Witch'/'ether'. A hybrid account of reference can easily solve the challenge represented by the claims in which such words occur – a criticism of which Pure Externalism was accused (cf. section 4). Claims like "witches do not exist", or "ether

was a name for such and such, but it was found out it does not exist" are "true just in case the history of those uses does not lead back to a grounding in which the relevant application conditions are met" (Thomasson 2007, p. 48). If there is no relevant nature ("the one that is, as a matter of fact, responsible for the properties picked out by the descriptions associated with the term in grounding) […] the groundings will fail" (Devitt and Sterelny 1999, p. 74). These cases can be perfectly explained, for, within such a semantic view, what matters is that what it takes to be the relevant nature is something decided *a priori*, and not entirely deferred to the world.

As the examples illustrate, the main advantage of this theory is that the constitutive rules are "open-textured, often simply leaving unspecified what to do in cases in which the condition fails" (Thomasson *forth.*, Ch.4). Consequently, the rules that govern a term's reference can always be revised given our interests or our empirical discoveries about the world. Such a view can then also perfectly accommodate also the empirical discoveries that we can make not only about natural, but also about artifactual or social kinds. We can use the constitutive semantic rules of those kind terms as a guide for their application, while still reserving some room to make decisions about all the uncertain examples. We will not need to establish to which semantics all our kind terms are subject, for "likely enough, different options are appropriate for different kinds of terms" (Devitt and Sterelny 1999, p. 79).We may then see most of the prominent discussions about how we should apply our kind terms (e.g. discussions about whether 'cat' should still be applied to a robot, or 'pencil' to an organism), or should they designate at all, as cases of what Timothy Sundell and David Plunkett define "metalinguistic negotiations" (2013, 2015), namely as pragmatic disputes conducted in the object language about how the terms we use to denote those objects *ought to be applied.* We are not merely talking past one another about which modal facts in the world should explain the truth of modal claims like "Cats are necessarily animals" or "Pencils are necessarily artifacts", or about which intuitions should drive our assent to Putnam's claim that "Cats have turned out to be robots remotely controlled from Mars" is surely non deviant, whereas "There are no cats in the world" is highly deviant (cf. Putnam 1975, p. 247). Instead, we will see these disputes as cases for which "we have to make a *decision* about what to do based on our surroundings practices and concerns" (Thomasson 2007, p. 50).

In sum, according to a hybrid theory of reference, in order to be grasped and successfully applied, our kind terms need, first of all, to be recognized as terms of the relevant category. Secondly, they have to be applied and reapplied according to the relevant public norms of use involved in a certain linguistic community. The frame-level application and co-application conditions are the result of a negotiation within a linguistic community, according to various factors that are considered to be relevant by that community. Our kind terms have some receptive and normative features, involving how they need to be regarded, used, or applied. Stressing these characteristics would definitely loosen the conditions of application, hence better account for all those cases in which a certain tension between the intentions of the grounder and the intentions of the competent speakers might arise. [49]

In conclusion, the hybrid theory of reference endorsed in this work is superior in many ways to the previous semantic views that have been advanced. It not only overcomes the famous problems of which Pure Externalism has been accused, but it can also entirely account for all fuzzy cases, while preserving the intuitions that motivated Externalism.

7. Conclusion

The most famous and generally defended unified account of the semantics of our kind words is Putnam's Externalism. This theory, nevertheless, confronts many objections, most of which are believed to arise from Putnam's proposal of extending the scope of his semantic theory to artifactual and social kind terms. I have pointed out that, on closer inspection, such semantic theory seems to be unsteady also in explaining the semantics of terms for natural substances and species. I argued that the source of its problems can be found in the fact that Putnam's semantic account implicitly relies on a hyperrealist view on modality. I have laid out how this descriptive modal approach faces ontological, epistemic and methodological problems and how Modal Normativism attempts to avoid those problems. I have then argued that such an alternative modal account commits us to abandoning Putnam's Externalism and to embracing a hybrid

[49] This formulation was inspired by the considerations about the "nature" of artifacts held by Thomasson in her "Public Artifacts, Intentions, and Norms" (2014).

account of reference. Lastly, I defended that such hybrid semantic theory based on a normativist approach to modality handles very well (or at least better than Externalism) the conflicting intuitions that some of Putnam's examples may raise. It maintains all the advantages of the externalist account of reference, since it is not at all unresponsive to empirical discoveries, while it overcomes the main criticisms thereof by loosening the constitutive rules according also to our purposes and interests. It represents, in the end, a more tenable but also attractive account over the main pre-existing theories.

Conclusion

To briefly recap, in the present work I have shown that in order to formulate a genuine externalist argument, two requirements ought to be satisfied: a) establishing whether or not there is a "nature" that determines membership in each kind; b) formulating arguments from ignorance and error about such "nature". I have proven that, on closer inspection, Putnam's arguments about artifactual (and social) kind terms do not meet any of them, and also that none of the authors who tried to ground the fact that those very terms do refer externalistically actually succeeded either. I have then described how the source of Putnam's Externalism problems can be tracked down in its relying on a heavyweight modal realism and how espousing Modal Normativism can lead to the dismissal of the externalist theory in favor of a hybrid account of reference, which seems to be overall more tenable and attractive.

I aim now to conclude by showing further payoffs of this approach. Resorting to Plunkett (2013) and Thomasson's (2016, 2017) notion of metalinguistic negotiation, I aim to show how, even in the deflationary framework the account I here advance is inscribed, those disputes can be seen as important.

Deflating the Debate

> For how we use words *matters*, given their relation to other aspects of our conceptual scheme, and to our non-verbal behavior (Thomasson 2017, p. 11).

The hybrid semantic approach based on a normativist account of modality I here propose leads to a natural deflation of the debate here at stake – the debate about the semantics of artifactual and social kind terms.

As argued at a great length (cf. Ch. 4), if we take such a deflationary approach, we no longer see the cases usually discussed within the debate as presenting instances in which there are modal features of the world that outstrip our conceptual resources and ability to know them. Moreover, appealing to the notion of metalinguistic negotiation – as implied in the view I advance – carries a further great advantage of seeing the disputes which have arisen around the mentioned controversial cases not as merely verbal or nonsensical – as they would be seen in traditional deflationist approaches – but rather as

prescriptive discussions about how the terms involved *should* be used. "And such disputes may be very much worth having" (Thomasson 2016, p. 9).

Metalinguistic negotiations are more common in our conversations than we may think. Music fans discussing for a surprisingly long time whether or not rock music should include some bands, or kids contesting the way their parents use the category of vegetables ("why fries are not in it?") are examples. Imagine a 7 year old girl in front of a dish of boiled spinach and French fries. When asked to have her vegetables, she eats her fries and pushes the spinach aside. Her mother repeats the request, explaining she was referring to the spinach (for fries are not vegetables), but the girl is clear: she's already had vegetables. A metalinguistic negotiation has begun together with the nutritional one: no matter how hard the mother tries to explain the differences between healthy greens and fried food, the girl persists in advocating that fries *are* vegetables. Nor is she interested in the fact other people agree with her parents on how the term is correctly applied. Lastly, she would not give up the discussion even if presented with evidences, say, from books or dictionaries.

These and similar disputes are characterized by some hallmarks that, according to David Plunkett and Tim Sundell (2013), separate metalinguistic negotiations from other verbal disputes. In Thomasson's interpretation, those hallmarks are: a) the disputants may share the same knowledge or beliefs on the matter, and agree about the fact any further discoveries might resolve the debate; b) even if the disputants agree about the term is commonly used, the dispute doesn't end; c) the dispute may persist even if the disputants acknowledge the fact that they are using the term differently.

In Thomasson's example of a debate on the nature of figure skating these hallmarks are quite evident: even if disputants agreed about the athletic performances of skaters and the amount of technical and physical training required, as well as on the different, respective uses of the word 'sport', they would not give up the dispute. "For regarding of whether or not we see the speakers as literally asserting conflicting propositions, their real disputes arises at the level of what it is pragmatically communicated" (2017, p. 9). In uttering: "Figure skating is/is not a sport" each discussant is pragmatically advocating for a how the term 'sport' *should* be applied, given its relation to other concepts and pragmatic consequences. This is why these disputes do not easily

eliminate, but they rather matter. That figure skating is/is not regarded as a sport bears indeed on its receiving awards, sponsorships, taking part in the Olympics, etc.[1]

For what I am here concerned, those fuzzy cases discussed in this work, as the imaginative examples of robotic cats, pencils-organisms, spying-device-pencils, and so on, may be seen not as disagreements on the possible discovered fact (that all the things usually called 'cats' are in fact robots, that all pencils-like objects are spying devices, and so on), or on the ways the term in question is actually appropriately used, or on the different ways the two discussants are using it. Rather, we may see the two sides as engaging in negotiations on how the term should be applied. Recalling Nelson's case (cf. Ch. 3, § 2.4), in claiming: "pencil-control devices *are* pencils", the Martian would not be giving a competing description of a modal fact out there, but she would rather be *pragmatically* advocating for a certain application of that very term. She would be negotiating a revision on its use, based on practical purposes and considerations. After all, her interest is keeping on spying on us through those devices. The opposite position can maintain that 'pencils' should not be applied to those objects, exactly for preserving the Earthlings' rights, privacy, and so on. Another practical consequence thereof would be that – as Schwartz point out – we would start fabricating "real" pencils.

We may even see the negotiation taking place not only on how the word should be applied but even on whether it should be applied at all. The example of 'witch' may be seen as a case in which the adversary of Putnam's account on the matter is advocating for a forsake of the use of 'witch', on the basis of the negative conveyed meaning it has. Think about the time in which the term was actually use to discriminate and burn the women label with it. Utterances such as: 'Witches do not exist" may be used to alter the norm for using the term, in force of its pragmatic impact on way of life. In this sense, utterances of trivial truths or falsehood like these ("Witches do not exists", "robotic cats are not animals", and so on) may have "a deeper pragmatic point: to press for ways the very linguistic/conceptual scheme used in expressing the obvious truth or falsehood is to be used" (Thomasson 2017, p. 12).

[1] I myself witnessed – years before my interest in the matter – to a heated discussion among friends exactly on the presence of artistic gymnastics, synchronized swimming and similar should be competing in the Olympic games.

Those disputes cannot be dismissed as merely verbal and are very much worth having. Deflating the semantic debate object of the present work maybe thought as a move that does not take into account the discussion still open or as a way of dismissing even the deep disputes about the metaphysics of modality at their background. On the contrary, a the examples show, seeing the rival positions that take part in the debate as engaging in metalinguistic negotiations enables to preserve the important payoff of demystifying the epistemology of modality that grounds those semantic approaches, while also giving and improved account of why those disputes may be important.

Chapter 5 - Appendix

On Further Externalist Defenses

As recalled in Chapter 3 (§ 2), many routes have been tried in order to defend Externalism *tout court*, namely as explaining the semantics of all our kind terms. I have shown that none of them succeeded in doing so. However, one may be still convinced there are different possible ways of preserving such an account of reference. I will here theorize and analyze them, then bringing to light their difficulties. There are three main further ways to support a thoroughly Externalism. One is to insist on the fact that, in order to have externalist reference, a kind term need only to be originally thought to refer to something that purportedly share a common structure (section 1). The second consists in underlining the importance of the linguistic norm over the indexicality (section 2). The third is to maintain that reference has to be secured to a paradigm and not necessarily to a nature in the sense used so far (section 3). I will conclude by illustrating how the hybrid account of reference seems still to be preferable (section 4).

1. Defending Externalism. Further options

1.1. An extreme defense of Putnam's theory: Super-Pure Externalism

From Chapter 2 we can derive the following lesson – an artifactual kind term can be defined as:

a. one which does *refer* to artifacts;

b. one *originally regarded as* referring to members of an artifactual kind.

If a., Putnam's argument is not effective, since it doesn't tell anything about the semantics of our well-functioning artifactual kind terms – as concluded in Chapter 1 already. If b., I also reached the conclusion that Putnam's argument does not prove what it was meant to prove, since it does not tell us what semantics a term has when is *intended* to refer to artifacts and such an intention is not frustrated. In both cases, Putnam's argument shows, at the very best, that we could be mistaken regarding the category a given term is supposed to belong to.

Nevertheless, one might object that we can loosen the two externalist requirements, as far as for them to entail that for any kind word to have externalist reference is sufficient that:

a) it refers to things that we *think/believe* to refer to something of a certain kind, *supposedly* identified by a certain nature;

b) linguistically competent speakers can be ignorant of, or mistaken about either the fact that it actually refers to that certain kind or to that alleged nature.[1]

According to this loose interpretation, which can be labeled as *Super-Pure Externalism*, any of our terms, including artifactual and social kind terms, can turn out to refer to very different kinds of objects than what we think – exactly as Putnam seemed to imply. We only need to *presuppose* that the extensions of our kind terms belong to the kind we initially associate them with. What matters it is just that we could discover someday that those kind terms do not refer to what we initially thought them to designate. Think about the robotic cats case: from this example, one might conclude that what matters it is just that 'cat', a word purported to designate a natural kind, turned out to refer to an artifactual kind instead. Following this extreme defense of Putnam's theory, one can argue that Putnam is only committed to show that the word 'pencil' or the word 'pediatrician', as well as their like terms, do not refer to what we *originally thought* them to refer – as he in fact suggested by his thought experiments. Thus – following this line of reasoning –, Putnam does not really need to outline arguments from ignorance and error about a supposed artifactual or social essence. He only has the burden to prove that we could be in error about most of the referents of the terms we normally use. Therefore, Putnam's arguments – the same objector can conclude – actually proved that Externalism, in this Super-Pure version, applies, at least, to all our kind words, and allegedly to other terms as well, as long as they fulfill its requirements.[2]

Such an extreme version of Externalism, can be grounded on a semantic corollary of Externalism that Putnam enhances: the idea of "semantic markers". He grants that we intuitively tend to associate certain words with certain categories of objects, as well as that there are some inferential associations that seems to be more central and irrevocable

[1] The difference between Pure Externalism and Super-Pure Externalism is, principally, the following: only according the latter, more extreme, version, terms have externalist reference if it is discovered that the shared hidden structure among their referents is not the one originally purported to be (and this reach the extreme level for which the kind of kind to refer to is never taken into consideration as an important vector to determining the extension).

[2] Recalling Putnam's claim: "…the points we have made apply to many other kinds of words as well. They apply to the great majority of all nouns, and to other parts of speech as well" (Putnam 1975, p. 242).

than others. “It seems to me reasonable that, just as in syntax we use such markers as “noun”, “adjective”, and, more narrowly, “concrete noun”, “verb taking a person as subject and a abstract object”, so in semantics these category indicators [such as “natural kind”, “artifactual kind”, and, more narrowly, “animal”, “utensil”, …] should be used as markers” (Putnam 1975, p. 268). However, he specifies that those semantic markers never provide necessary and sufficient conditions for membership in a kind, nor do they provide any analytic specification of the meaning of the terms they are associated with. Although – Putnam admits – it is hard to imagine, it can always turn out that we were wrong, and that cats *are* actually robots or that pencils *are* actually organisms. The fact that we could discover that their designators do not refer to the category of objects they were initially purported to refer does not invalidate their having an externalist semantics, but it rather confirms it. Putnam seems to suggest that those described would be situations in which we would find out that the properties usually regarded as central of cats or pencils are others instead – as it would happen in the counterfactual case of reptile-tigers (cf. Ch.1, § 1.3). The linguistic community would change the definitions connected to their designators in accordance with the discovery, for he reference of our kind terms is fixed by an ostensive definition pegged to a paradigm, whatever the paradigm turns out to be.

Therefore, Super-Pure Externalism seems to be more promising for the friend of Externalism applied to all kind words. Such an extreme version of the theory not only seems to provide a more convincing explanation of Putnam’s argument about artifactual and social kind terms, but it apparently also to give a good account of some controversial cases which seemed to lead us farther from it.[3]

[3] Super-Pure Externalism eludes even the problem of the opacity of the semantics, since the semantic properties of a kind term are made to depend, not on its actually referring to something which share a common nature, but on its *purporting* to designate that something. The linguistic community is always supposedly aware of the semantics of kind terms: it is always allegedly externalist.

1.2. Replying to Super-Pure Externalism

If one is convinced by the account just described and does not see anything jarring about it, it is hard to be convincing of the contrary. However, for all others who are not fully persuaded by all the theses and implications of such a super-pure theory, let us make some considerations.

First of all, it has to be noted that – although *prima facie* promising – this extreme version of Externalism is grounded on a quite implausible interpretation of Putnam's claims and examples. It seems indeed that the robotic cats example may be more coherently read as implying that 'cat' would still refer to something that has a nature which determines membership in the kind, the difference is that that "nature" would be artifactual rather than biological. The fact might rather be that we don't know what it takes to be that artifactual essence. As Putnam claims: "we know that there are kinds of things with common hidden structure, but we don't yet have the knowledge to describe all those hidden structures" (Putnam 1975, p. 244). This can be confirmed by the fact that in the case of water, gold, tigers, etc. what is taken to play the semantic role in determining the extension of their designators is not the fact that those things generally belong to natural kinds, but rather their specific inner structures. Putnam's Externalism seems to tacitly presuppose that what fixes the extension of different kinds of terms is still the appropriate hidden structure of the proper referents, rather than *any* nature (actual or alleged).

Furthermore, it is to remember that Putnam's thought experiment of robot-cats, pencil-organisms, and pediatrician-Martian spies are meant to claim that if cats, pencils, and pediatricians *are*, respectively, animals, artifacts, and human pediatricians they are so *with metaphysical necessity*. The sense in which we could discover someday that things are otherwise is only epistemological – on the contrary of what Super-Pure Externalism seems to entail. This is also confirmed by Putnam's notion of "conceivability". He holds that it is perfectly *conceivable* that cats, to take one of the examples, turned out not to be animals.[4] However, it is not *logically possible*.[5]

[4] In this, Putnam's theory differentiate itself, once again, from Kripke's. According to Kripke, what we are really conceiving when we entertain the mental state that cats could turn out not to be animals is a situation in which not *cats*, but something that causes the same epistemic evidence of cats (i.e. an

Conceivability is a matter of speculation on epistemological hypotheses about how the things might have been, whereas logical possibility is a matter of modal facts (which are out there, waiting to be discovered). "What is conceivable is a matter of epistemology […]. [Whereas] When terms are used rigidly, logical possibility becomes dependent on empirical facts" [Putnam 1990, p. 62]. Although it is perfectly conceivable that cats are not (or turn out not to be) animals, once we find out that they are actually animals, there is no possible world in which it is not so[6]: cats not being animals is a metaphysical impossibility.[7] This fact does not depend on our knowledge or possibility to describe it, but rather on how the things actually are.[8]

"epistemic counterpart" of cats), is a robot. In other words, we can conceive of being in a "qualitatively identical epistemic situation" [Kripke 1980, p. 104], except from the fact that the evidence we have regarding *cats* is generated by *catslike* objects which are not animals. The difference relies on the fact that even if Putnam borrows the notion of "epistemic counterpart", he use it to clarify *why* it is conceivable that cats are robots, whereas Kripke adopts the same notion to explain *how it may seem* that it is conceivable that *cats* are robots when actually it is not. See on this also Marconi (2009a).

[5] In Putnam's own words [with due substitutions]: "We can perfectly well imagine having experiences that would convince us (and that would make it rational to believe that) water isn't H_2O. In that sense, it is conceivable that water isn't H_2O. It is conceivable but isn't logically possible! Conceivability is no proof of logical possibility" [Putnam 1975, p. 233]. A further proof of that can be found in the following quote about the possibility of tigers not being animals: "…it is curiously hard to think of the case to begin with, which is why it is easy to make the mistake of thinking that it is "logically impossible" for a tiger not to be an animal. On the other hand, there is no difficulty in imagining an individual tiger that is not striped; it might be an albino" (Putnam 1975, p. 267). Although is harder to imagine tigers not being animals (but robots for instance) than it is to imagine albino tigers (probably because we actually found out that there are albino tigers), this does not rule out the possibility of conceiving that situation happening. However, conceivability is a matter of epistemology and no proof of logical possibility. It is the metaphysics of the modal facts out there which establish what is or is not logically possible.

[6] Of course, one might object that, to the contrary, the fact that pencils are artifacts or that pediatricians are doctors specialized in the care of children, does not exactly seem something we discover. I will come back to this issue in section 4.

[7] As Putnam puts it [with due substitutions]: "…we will now be unwilling to say that any logically possible situation is "one in which water is XYZ". And that means that it is *impossible – metaphysically impossible* – for water to be XYZ" (Putnam 1990, p.61).

[8] As evident from the considerations outlined in this paragraph, also this argument for the invalidity of the inference from conceivability to possibility can give further evidence of the epistemological problems of

In conclusion, it seems that Super-Pure Externalism is not really preferable to simple Pure Externalism. Super-Pure Externalism runs into many contradictions, also with most of Putnam's claims, leading us even to doubting that it is actually a theory endorsing the indexicality of kind terms. The core of Putnam's Externalism is, ultimately, subsumable in the following claim mentioned many times already: "The extension of our terms depends upon the *actual* [italics mine] nature of the particular things that serves as paradigms, and this actual nature is not, in general, fully known to the speakers" (Putnam 1975, p. 245). What matters is the *actual* nature of the things that serve as paradigms, not *any alleged* nature. In this respect, Super-Pure Externalism, would be grounded on a probably even more extreme version of heavyweight modal realism, for on such an account, we can only aspire to vaguely suppose to which category of entities a given term refer, and secondly to give an accurate description of the alleged nature of its extension. Nonetheless, we will, in principle, never be able to reach that achievement. As pointed out (cf. Ch. 4, § 1), also Pure Externalism relies on a heavyweight view of modality (this is why the word 'actual' is so crucial), but taking such modal view to this extreme level, as Super-Externalism does, leads to undermining the semantic theory which derive from it till radicalizing it into something very implausible. However, it was because of its plausibility, although probably only *prima facie*, that Externalism has been so quite widely adopted.

Putnam's Externalism can be at best weakened as requiring that a given term has externally reference iff:

a) it refers to things that we *think/believe* to have a nature (necessary features) possessed by all of the members of the term's extension;

b) linguistically competent speakers can be ignorant of, or mistaken about such alleged nature.

In this respect, a loose adaptation of the empirical presupposition of Putnam's Externalism (cf. Ch. 2, § 2) would be met.[9] But this weakened interpretation runs into the difficulties I pointed out in Chapter 2 (§ 2, 4). It does not tell us what semantics a

---

Putnam's heavyweight view of modality. Cf. what already remarked in Chapter 4, § 2. It is hard to see how we can explain the epistemology of modality within such an approach.

[9] The empirical presupposition can be loosen and becomes: (**LEP**) *Supposedly*, there is a real nature shared among all (or most of) the members of a kind denoted by some classes of terms, even though we do not have or will we ever have any epistemic access to it.

kind term has when is *intended* to refer to a given kind and such an intention is not frustrated.

Either way, my objections to Putnam's argument still stand. Even these radical versions of Externalism do not represent a real alternative. If one is convinced by what I showed at a great length in Chapter 4, a hybrid theory of reference constitutes still a more tenable and overall desirable account of the semantics of our kind terms – as I shall point out in section 4.

2. Defending Externalism. Further options

2.1. A weak defense of Externalism: Social Externalism[10]

Another way of defending Putnam's Externalism *tout court* is by underestimating the indexical component in favor of giving importance only to the social component of the meaning. This is the strategy adopted by Kornblith, for instance (cf. Ch. 1, § 3.2). We already seen in Chapter 2 (§ 4.1) that such an account is not a good defense of Putnam's Externalism. However, since most authors take for granted that this is the way Externalism should be read, I considered it may warrant more discussion.

Kornblith states that "the division of linguistic labor is just as much a part of the world of artifacts as it is a part of the world of natural kinds and individuals" (Kornblith

[10] By "Social Externalism" I intend the semantic account argued for by commonly attributed to Tyler Burge and his "Individualism and the mental" (1999). On such a view, in case of error or ignorance common speakers of a certain linguistic community are always willing to accept the experts' corrections or directions, and to revise their previous beliefs in accordance with those corrections and directions. This deference attitude towards the expert speakers is what makes them part of *that* linguistic community. Burge's most famous example is the one involving the term 'arthritis': if, in a counterfactual situation, the linguistic norm of the counterfactual linguistic community established that the term refers to a disease causing painful inflammation and stiffness of the muscles, a common speakers who ever expressed to his doctor his beliefs of having developed arthritis to a joint, would be promptly corrected and accept the correction. I argued that Kornblith seems to extend this form of Externalism to artifactual (and social) kind terms. I will limit my analysis to this semantic theory understood as such.

2007, p. 144).[11] Following Putnam's considerations, he affirms that what fixes those kinds terms' reference is not a uniquely individuating description a single speaker has in mind, but rather whatever in a given linguistic community embodies the semantic norm. The semantic norm established in a linguistic community secures the extension of our terms, for it is capable of discriminating between the correct and improper use of those words. A defender of Externalism applied across the board may try to argue that this is exactly what Putnam really intends, settling such a conclusion in the fact that he has substantially stated several times that the reference of our terms is socially determined. Putnam has indeed meant to prove principally that "the extension of our terms is not fixed by a concept that the individual speaker has in his head" (Putnam 1975, p. 245). The several counterfactual and real cases he brought about were to show exactly that. Some of them (in particular the aluminum/molybdenum example)[12] were precisely meant to show that "it is only the sociolinguistic state of the collective linguistic body to which the speaker belongs that fixes the extension" (Putnam 1975, p. 229). Consequently, one may reasonably think that this is what Putnam ultimately endorses. Consider again the 'elm' and 'beech' example. Putnam comments that example, saying that "the reason my individual "grasp" of 'elm tree' does not fix the extension of elm […] is rather that the extension of 'elm tree' in my dialect is not fixed by what the average speaker "grasps" or doesn't "grasps" at all; it is fixed by the community, including the experts, through a complex cooperative process" (Putnam 1975, p. 263). The example seems to show that what matters for a word to have externalist reference is only the objective relationship between the term and the social environment in which it is uttered. The experts play a crucial and reference-determining role. They are to establish the correct application of a term, and therefore to possibly correct or rebuke others for misuse, on the basis of extensions they themselves individuate. Or so it seems to emerge from Putnam's discourse about the division of linguistic labor.[13]

[11] One of the main criticisms of Kornblith's theory derives exactly from his making Kripke's causal-intentional theory of proper names collapse on the externalist account of reference of terms for natural substances and species.

[12] Cf. Ch. 1, § 2.2.

[13] Quoting Putnam: "in many cases, extension is determined socially and not individually, owing to the division of linguistic labor" (Putnam 1975, p. 226).

Such a reading seems very promising if one aims to ground an externalist semantics about all kind words. As Kornblith underscores, there is plenty of cases in which we don't know what description(s) uniquely pick(s) out the referents of a given artifactual term; yet we succeed in referring. Think about complex artifacts such as high-tech equipments or artifacts specific to a certain field or discipline. If I am not a chef I may not know any description that serves in distinguishing a blast chiller from a regular freezer. And I may not be able to tell the difference even in presence of the two objects – as it happens in the case of elms and beeches. I can still use those terms and succeed in referring, for I can rely on the experts of my linguistic community. In the case of social kind terms this phenomenon of the division of linguistic labor may appear even more pervasive, if anything, at least for the fact that it is really the norm widespread in a linguistic community what actually gives those words' reference. What ultimately determines the meaning and the reference of our words is the semantic norm adopted by the linguistic community in which they are used.

2.2. Replying to Social Externalism, or, rather, "Masked" Internalism

Although this account appears very promising, I aim to prove that it is not only independent from, but even incompatible with, the form of "metaphysical" Externalism object of the present work. As anticipated and underlined many times already, the indexical component of the meaning cannot be so easily neglected – at least if one means to account for Putnam's semantic theory. What described above is not his view at all. The so-called "Social Externalism" is rather a form of "masked" Internalism, for what matters for the determination of our kind words' extension is exclusively the linguistic norm established in a given linguistic community.

According to Social Externalism, every linguistic community adopts a certain semantic norm. Speakers belonging to a certain linguistic community acquire the terms according to the semantic norm conventional to that linguistic community. In other words, they acquire the terms in the way they are used in their social environment and apply those terms to whatever the members of that linguistic community apply them. This means that the same word could be applied to a different kind of entities in different linguistic communities, if the semantic norms respectively adopted by those linguistic communities tell so.

But this fails to be what Putnam's implies, even implicitly. Putnam claims that what matters in the determination of the reference of a given word is the *indexical* component. The nature of the paradigm sample governs the extension of the term, independently of anyone's concepts or beliefs. Even expert speakers may be massively mistaken about, if not even utterly ignorant of, what this nature could be – as real and counterfactual cases such as the one of water, gold, pencils, and so on, show. The fact that we could be ignorant or massively wrong about even all the properties associated to a certain term's extension proves that – contrary to what Kornblith affirms – the problem of ignorance and error do not arise solely because of the division of linguistic labor.[14] They arise because the "metaphysical" externalist view always presupposes that we might never come to know what the individuating properties of the referents of our terms are. As Putnam puts it: "if there is a hidden structure, then generally it determines what it is to be a member of the natural kind, not only in the actual world, but in all possible worlds. Put another way, it determines what we can and cannot counterfactually suppose about the natural kind" (Putnam 1975, p. 241). This is exactly what he means when he talks about rigid designations. It is because we refer to the nature of water, to its chemical formula, which is the reference-determining property of water, that we can imagine and describe counterfactual cases about *water*.

What Putnam is endorsing is not a form of cluster-descriptivism (cf. Ch. 1, fn 4), for which we may be ignorant of or wrong about *some* of the properties associated to a given term, but rather that all those properties are attached to its extension completely *independently* of anyone's beliefs and knowledge about it, so that we could be massively mistaken about (virtually) *all* of them. The fact that the success of referring of a random competent speaker is not prevented by her incapacity of telling an elm from a beech, it is certainly due to the cooperation among speakers of the linguistic community, but it is especially due to the world-deferentiality of the reference-determining properties of those words. The role of the expert is only *pragmatic* and never semantic. Indeed, if we found out that the properties typically associated with a given extension are wrong we will be immediately willing to give them up in favor of

[14] In Kornblith's own words: "And this point by itself is sufficient to ground Arguments from Ignorance and Error in the case of artifacts which are exactly parallel to those which underwrite the new theory of reference for names and natural kind terms (Kornblith 2007, p. 144).

the discovered ones – as the fictional case of reptile-tigers shows. By contrast, on a social externalist view, this could not happen. If the semantic norm of a given linguistic community established that reptiles are, as the definition connected to the term specifies, "air-breathing animals having a backbone, a heart with three chambers, a completely bony skeleton, and a covering of dry scales or horny plates", the peculiar looking reptiles of the example (the ones looking like tigers) would not *be* reptiles. They would not be classified as such, for they do not fulfill the description the experts of the linguistic community connect to the term 'reptile'. The role of the expert speakers in always *semantic*, for the reference of our terms never changes due to empirical discoveries about the nature of their referents, but it is rather governed by the linguistic norm adopted by that community. The role of the experts [15] is *prescriptive* rather than purely descriptive. On a metaphysical Externalism, to the contrary, the experts govern the *use* of our terms, but never their reference. On Putnam's view, semantics and pragmatics are very distinct: he admits, for instance, that the nature of the required minimum competence depends heavily upon the culture and the topic; however, this has nothing to do with the way our terms' extension is determined. Grounding a term's reference is one thing, and the other are the practical purposes of communications.

This rule out also the possibility of claiming that Putnam's Externalism entails a form of Social Externalism, if this was interpreted as taking the semantic norm only as a *formal* norm, namely as merely prescribing to use the terms as they are used in the linguistic community in which they are uttered. Consider the case of gold and fool's gold. Before 1750, nobody, not even the expert speakers of any linguistic community, could distinguish between the two substances. *All* the speakers applied 'gold' indifferently to both substances. Nonetheless, after 1750, even the formal norm changed, due to the discovery. On a social externalism account, the experts, to which speakers defer by the formal norm, represents the ones who can ultimately discriminate between correct and improper use of words, whereas on Putnam's view, this role is played by the hidden structure of the referents, to which, most of the times, not even experts speakers have access.

[15] Note that by "expert speakers" is intended whoever or whatever can embody the semantic norm: lexicographers, experts in the fields, dictionaries, etc.

On Putnam's Externalism the indexical component "dominates" the social one in determining the reference. This is why his account is not only independent from, but even incompatible with, the other *alleged* form of Externalism. On Kornblith's theory, as much as on any Social Externalist theory whatsoever, such indexical component is not even mentioned; rather it is utterly disregarded. It is true that the meaning and the reference of our words depends on the an objective relation between them and the world, but "world" is intended only as the "social environment" in which they are uttered, and never as the "physical world".

In sum, Kornblith's argument according to which "there are interesting metaphysical differences between artifactual and natural kinds, but these metaphysical differences play no role in the semantics for terms which refer to these kinds" (Kornblith 2007, p. 149) is misleading, at least if it is meant to extend the scope of Putnam's very semantic theory. A hybrid semantic view is neither unresponsive to those metaphysical differences, nor to empirical discoveries. Nonetheless, it does not undermine the importance of the role played by our concepts, consequently still providing a preferable account of the reference of all our kind terms – as I shall remind in section 5.

3. Defending Externalism. Further options

3.1. A moderate defense of Externalism: "Externalism of the Paradigms"

"Essentially contested terms and natural kind terms both, then, apply to something now just in case it has a certain relation to an original sample or an historical exemplar" (Evnine 2014, p. 128). We may construe this defense of Externalism as one which takes as central the fact that the word bears the same objective relation to a certain group of objects which serve as the *paradigm* of the kind that word is meant to refer.[16] According to such a theory "current uses of the term continue to be connected to the original sample […] through the (causal) historical connection between the sample baptism and current usage" (Evnine 2014, pp. 126-127). On this view, what matters in grounding the

[16] The idea for this defense of Externalism is inspired by Simon Evnine's "Essentially Contested Concepts and Semantic Externalism" (2014). However, even if I quote his article to make my point, the one reported is not faithfully the theory he is arguing for.

reference of a term is that "it bears a certain kind of relation to samples or exemplars that have played an historical role in the use of the term" (Evnine 2104, p. 127). This is to reject the idea that the meaning of essentially contested terms (or in general social kind terms) is given by some descriptive content given as their definition. As an example we may consider the term 'Christianity': if its meaning was given by some descriptions constituting its definition there could be no genuine dispute between what may be called the Christian Right and the Christian Left. To the contrary, the dispute is legit and it is still going on because there is an historical phenomenon, an exemplar, "and each of the contesting parties claims for itself a certain relation to that historical phenomenon, a relation such that if it obtains in one case, it cannot obtain in the other" (Evnine 2014, p. 124).[17] For the parallelism with natural kind terms to hold, such a defense of Externalism needs to affirm that a given word has an externalist semantics iff:

a) it refers to things that are in the same objective relation to a certain exemplar which serve as their paradigm;

b) linguistically competent speakers may be ignorant of or mistaken about what constitute such a paradigm.

On such a view, even though there are obvious differences over the nature of both exemplars (it is natural for natural kinds, cultural for essentially contested kinds), the comparison between the semantics of their designating terms still stand. "Natural kind terms and essentially contested terms are two species of the same semantic genus" (Evnine 2014, p. 119), for their meaning is given by an objective relation occurring between them and the world (natural or cultural).

3.2. Replying to the Externalism of the Paradigms

Properly adapted, the defense of Externalism just described seems to have the main advantage of being capable of securing that all our kind terms refer externalistically.

[17] Specifically, the objective relation in this case is "being the heir of" that phenomenon. Leaving aside this notion, which may be problematic to define, we may simply construe such a theory as one according which essentially contested term refers to whatever bears the same objective relation to a paradigmatic exemplar.

There is a paradigmatic exemplar to which refer when grounding a kind term's reference, which all the referents of that designating term are in an objective relation with. On such a view, natural kind words have externalist reference for they share a certain sameness relation with the inner structure of things which serve as paradigms, whereas artifactual and social kind terms have an externalist semantics for they share an objective relation with some things or phenomena regarded as paradigmatic for those kinds. Thus construed, such a semantic view may be seen as a theory in between the strong and the weak defense of Externalism. We may infer that there is distinction between natural and cultural kind terms, so that the category of entities to which the terms refer is internally fixed. Yet, the paradigmatic exemplar or the nature of the objective relation at stake cannot be specified by the descriptions associated with the relevant words, which never play a crucial semantic role. Rather, what it takes to belong to the same kind as the paradigmatic exemplars is left for empirical discovery.

But what it takes to be a paradigmatic exemplar for artifactual and social kind terms? For those kinds the paradigmatic exemplars seem to be represented by a historical phenomenon or by a set of objects to which the referents have some historical connection.[18] However, only claiming that those referents are "part of" the same kind as the relevant paradigmatic exemplars seems to provide too vague a criterion. No "nature" in the sense intended so far is specified, nor is it allowed that such paradigmatic exemplars are constituted by some semantic rules, albeit revisable.

Related to this, such a semantic account, has also the problem of not being a good defense of Externalism. According to this externalist-friendly account, indeed, the operative relation is "belonging to the same kind as". But – as already pointed out – such a notion does not bring about any particular sense in which it should be intended, rather, it can be inferred that the same entity may belong to different kinds (cf. Evnine 2014, p. 129). "There is no reason why there could not be a semantic phenomenon in which we introduce a term with reference to a set of samples and the principle only that the term refers to things that are of the same kind (*however* that is determined) as the members of that set" (Evnine 2014, p. 130). This fits the regular notion of indexicality, according to which an indexical expression is a word or phrase that is associated with different meanings or referents on different occasions or contexts. In this sense, it can be

[18] Such a notion resembles the conception of the "nature" of culturally generated kinds by Elder (2007).

considered a form of "communitarian Internalism" (cf. Ch.1, § 3,6). However – as argued throughout the present work –, Putnam's Externalism interprets the notion of indexicality in a, different, peculiar way: an indexical term refers to things which share the nature of those objects that serve as paradigms of the designated kind. And once such a nature is found out, the term refers to *that* set of objects, which share *that* nature, in all possible worlds in which it refers. Therefore, for something to be a paradigm in the relevant sense, it must have a nature, an essence in the same sense of the one shared among natural substances and species, one which determines membership in the kind.

Thus, Externalism of the Paradigms inherits the difficulties faced by its stronger and weaker alternatives. Disregarding Putnam's notion of indexicality, it proves to be a form of "masked" Internalism; not specifying a clear criterion to fix the reference of our kind terms, it clashes into the same problems of Super-Pure Externalism.

The moral we can gain from Externalism of the Paradigms leads in favor of a hybrid semantic theory: in grounding a term's extension, we ought principally to disambiguate what are the relevant properties to be privileged. The term would refer to those object which display those relevant properties (empirically and conceptually defined). The road towards a hybrid approach to reference seems more and more paved – as we shall see in the following section.

4. Why still favor a Hybrid Account of Reference

Above, I outlined further possibilities of defending Putnam's Externalism thoroughly applied, whether are they stronger or weaker. I have concluded that none of them represents a good argument in this respect. Super-Pure Externalism has mainly two problems: on the one hand, it seems to misinterpret, to some extent, some of Putnam's claims, making us doubt that it is still an indexicalist theory. On the other hand, it faces even deeper criticisms than Pure-Externalism, for it radicalizes the robust metaphysics of modality on which they both rely, making such an account even less plausible than its alternative. Social Externalism turned out to be a form of Internalism, incompatible with Externalism *tout court*. On a social externalist view, the determination of our kind terms' reference depends on what the semantic norm of a certain linguistic community prescribes, norm which most of the times can be verbalized as descriptions connected to

the words. Our terms' extension is always fixed regardless of any reference to the nature of their referents. By contrast, on Putnam's account that indexical component always prevails over the social one. Lastly, Externalism of the Paradigms provides too vague a criterion in the case of artifactual and social kind terms, besides taking as crucial a notion of indexicality which is not Putnam's notion. In conclusion, none of this further options actually represents neither a good defense of Putnam's theory, nor a more tenable account of the semantics of our kind terms.

One may persevere in holding that the "masked" Internalisms are nevertheless preferable to their rival views, for they account for the cases of ignorance and error when it comes to artifactual and social kinds. The fact that we formulate hypothesis about the function or use of certain artifactual objects, or about the meaning of certain rituals or practices, belonging to some distant or dead societies can be perfectly explained within such "communitarian" Internalisms. Ignorance and Error concern the essence of the referents of artifactual and social kind terms belonging to different communities from the one from which standpoint they are considered. Or so they can argue. Putnam's insight that the extensions of those words is never determined by what a single speaker has in mind would be preserved. These theories may serve as explaining the semantics of social kind terms, granting the fact that their essential properties depends on collective intentionality and public norms. Specifically, if we regarded artifacts and social kinds as something intended to be subject to certain norms (especially as something to be recognized as to be treated, used, regarded, etc. in certain ways) – extending Thomasson's definition (2014) – then, we may have paradigmatic examples with which the *denotata* of those terms share a certain objective relation (for those norms are certainly public for the reference to succeed). However, inasmuch as forms of Internalism, those semantic views cannot account for the fact that we can make empirical discoveries about those kinds even within the same community, as the discoveries about the functional or structural features of cable wires or vacuum tubes show (cf. Ch. 2, fn 15; Ch 4, fn 50). On Social Externalism, this is because the extension of those words is given exclusively by the definitions associated with the term by the relevant linguistic community, definitions not revisable. On Externalism of the Paradigm, it is because the criterion to disambiguate what it takes to belong to the same kind as the paradigmatic exemplars is too vague. It may be thought that the solution is

to conditionalize such a criterion on empirical facts or decisions, but, thus construed, it would be a form of a hybrid theory (provided we allow the constitutive semantic rules to play a major role in determining the reference).

In sum, a hybrid approach to reference is still preferable. As argued at a great length in Chapter 4, such an account grounded on a normativist view of modality avoids the main objections of which Putnam's Externalism has been accused, while still preserving the intuitions which have led to its wide adoption. It is hard to see what entitles us to say that, when we use a certain word, we are referring exactly to its proper extension, without letting some conceptual content to play a role in fixing some of the features that are to be privileged. Consider again the Kornblith's case about 'glug' (cf. Ch. 3, § 2.5). When the Martian anthropologist uses 'glug', he cannot be referring properly to doorstops, unless he knew its relevant application conditions. Moreover, in order to correctly reapply the word, he should know some central features of glugs which make them the members of the category "doorstops". In other words, in order to be consider a competent speaker, he should know at least the stereotype (using Putnam's terminology) associated to the word 'glug' (i.e. to the word 'doorstop'), which include some characteristics its referents are believed to peculiarly possess – and this is something that even Putnam grants.[19] The knowledge of the application and co-application conditions of 'doorstop' are a requisite of the Martian's competence in using the term 'glug', unless he wants to be considered as referring only parasitically. This can be extended to any kind word, whether is it natural, artifactual, or social. On the other hand, room for ignorance and error is preserved. Constitutive semantics rules are worldly-deferential and open to revision based on our interests and purposes. If someday the same Martian anthropologist, or any Earthling speaker, found out that the referents of 'doorstop' have an additional function or important feature, he is allowed to accommodating the discovery. Additionally, a hybrid approach provides for later

[19] In Putnam's own words: "Significant communication requires that people know something [at least the stereotype] of what they are talking about. […] What I contend is that speakers are *required* to know something about (stereotypical) tigers in order to count as having acquired the word 'tiger' "; although "[t]he nature of the required minimum kevel of competence depends heavily upon both the culture and the topic" (Putnam 1975, p. 248). "In the extreme case, the stereotype may be just the marker [the category-indicator)" (p. 230).

decisions of giving up the term or of radically changing the defining descriptions associated with it, according to practices and purposes of the interested users.

5. Conclusion

From all the consideration outlined so far, it results that an artifactual kind word can be intended as:

a. one which *does* refer to artifacts kinds;

b. one *thought/believed* to refer to artifacts, to members of an artifactual kind, but which can turn out to refer to something else (e.g. a natural kind).

c. one *constitutively intended* to refer to artifacts, to members of an artifactual kind.

It seems that if we choose c. we gain the advantages bring about by the hybrid theory of reference, of accounting for all the controversial cases which led farther from the externalist approach, while still leaving room for ignorance and error which have made its case. Moreover, we will gain a unified account which can accommodate the semantics of all our kind terms, whether are they natural, artifactual, or social.

References

Abbott B., "Nondescriptionality and Natural Kind Terms", *Linguistics and Philosophy* 12 (1989), pp. 269-291.

Ahn W., "Why are different features central for natural kinds and artifacts?: the role of causal status in determining feature centrality", *Cognition* 69 (1998), pp. 135-178.

Baker L. R., "The ontology of artifacts", *Philosophical Explorations* 7 (2004), pp. 99-112.

Baker L. R., *The Metaphysics of Everyday Life: An Essay in Practical Realism*, Cambridge University Press, Cambridge 2007.

Bennett K., "Spatio-Temporal Coincidence and the Grounding Problem", *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 118 (2004), pp. 339-371.

Bloom P., "Intention, History, and Artifact Concepts", *Cognition* 60 (1996), pp. 1-29.

Bloom P., "Theories of artifact categorization", *Cognition* 66 (1998), pp. 87-93.

Bloom P., "Water as an Artifact Kind", in E. Margolis and S. Laurence (eds.), *Creations of the Mind: Essays on Artifacts and their Representation*, Oxford University Press, New York 2007a, pp. 151-156.

Bloom P., "More Than Words: A Reply to Malt and Sloman", *Cognition* 105 (2007b), pp. 649-655.

Burge T., "Individualism and the Mental", *Midwest Studies In Philosophy* 4 (1999), pp. 73-121.

Burke M. B., "Copper Statues and Pieces of Copper: A Challenge to the Standard Account", *Analysis* 52 (1992), pp. 12-17.

Carey S., *Conceptual Change in Childhood*, MIT Press, Cambridge 1985.

Carrara M. and D. Mingardo, "Artifact Categorization. Trends and Problems", *Review of Philosophy and Psychology* 4 (2013), pp. 351-373.

Carrara M. and P. Vermaas, "The fine-grained metaphysics of artifactual and biological functional kinds", *Synthese* 169 (2009), pp. 125-143.

Cummins R., "Functional Analysis", *The Journal of Philosophy* 72 (1975), pp. 741-765.

deRosset L., "What is the Grounding Problem?", *Philosophical Studies* 156 (2011), pp. 173-197.

Devitt M., *Realism and Truth* (2nd ed.), Princeton University Press, Princeton 1991.

Devitt M. and K. Sterelny, *Language and Reality: An Introduction to the Philosophy of Language*, Blackwell, Oxford 1999.

Elder C., "Realism, Naturalism, and Culturally Generated Kinds", *The Philosophical Quarterly* 39 (1989), pp. 425-444.

Elder C., "On the Place of Artifacts in Ontology", in E. Margolis and S. Laurence (eds.), *Creations of the Mind: Essays on Artifacts and their Representation*, Oxford University Press, New York 2007, pp. 33-51.

Evnine S., "Essentially Contested Concepts and Semantic Externalism", *Journal of the Philosophy of History* 8 (2014), pp. 118-140.

Evnine S., *Making Objects and Events*, Oxford University Press, Oxford 2016.

Ferraris M. (ed.), *Storia dell'ontologia*, Bompiani, Milano 2008.

Ferraris M., *Documentalità*, Laterza, Roma 2009.

Fine K., "The impossibility of vagueness", *Philosophical Perspectives* 22 (2008), pp. 111-136.

Franssen M., P. Kroes, T. Reydon, P. E. Vermaas (eds.), *Artefact Kinds. Ontology and The Human-Made World*, Springer, Dordrecht 2014.

Frascolla P., D. Marconi, A. Voltolini (eds.), *Wittgenstein: Mind, Meaning and Metaphilosophy*, Palgrave MacMillan, London 2009.

Frege G., "Über Sinn und Bedeutung", *Zeitschrift für Philosophie und philosophische Kritik* 100 (1892), pp 25-50.

Gelman S. A. and J. D. Coley, "Language and categorization: The acquisition of natural kind terms", in S. A. Gelman and J. P. Byrnes (eds.), *Perspectives on language and thought; Interrelations in development*, Cambridge University Press, Cambridge 1991, pp. 146-196.

Gelman S. A., J. D. Coley, and G. M. Gottfried, "Essentialist beliefs in children: The acquisition of concepts and theories", in L. A. Hirschfeld and S. A. Gelman (eds.),

Mapping the Mind: Domain Specificity in Cognition and Culture, Cambridge University Press, Cambridge 1994, pp. 341-365.

Gentner D., "A study of early word meaning using artificial objects: what looks like a jiggy but acts like a zimbo?", *Papers and Reports on Child Language Development* 15 (1978), pp. 1-6.

Grandy R. E., "Artifacts: Parts and Principles", in E. Margolis and S. Laurence (eds.), *Creations of the Mind: Essays on Artifacts and their Representation*, Oxford University Press, New York 2007, pp. 18-32.

Hilpinen R., "On artifacts and works of art", *Theoria* 58 (1992), pp. 58-82.

Hilpinen R., "Artifact", *The Stanford Encyclopedia of Philosophy*, Stanford University Press, Stanford 1999.

Hoffman J. and G. Rosenkrantz, *Substance: Its Nature and Existence*, Routledge, London 1997.

Houkes W. and P. E. Vermaas, *Technical Functions: On the Use and Design of Artefacts*, Springer, Dordrecht 2010.

Hume D., *Enquiry Concerning Human Understanding* (1777), Clarendon Press, Oxford 1977.

Keil F. C., *Concepts, Kinds, and Cognitive Development*, MIT Press, Cambridge 1989.

Keil F. C., "The birth and nurturance of concepts by domains: The origins of concepts of living things", in L. A. Hirschfeld and S. A. Gelman (eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture*, Cambridge University Press, Cambridge 1994, pp. 234-254.

Keil F. C., M. Greif, and R. S. Kerner, "A world apart: How concepts of the constructed world are different in representation and in development", in E. Margolis and S. Laurence (eds.), *Creations of the Mind: Essays on Artifacts and their Representation*, Oxford University Press, Oxford 2007, pp. 231-245.

Kevles D. J., "The Genes You Can't Patent", *The New York Review of Books* 60 (2013).

Khalidi M. A., *Natural Categories and Human Kinds*, Cambridge University Press, Cambridge 2015.

Kornblith H., "Referring to Artifacts", *Philosophical Review* 89 (1980), pp. 109-114.

Kornblith H., "How to Refer to Artifacts", in E. Margolis and S. Laurence (eds.), *Creations of the Mind: Theories of Artifacts and Their Representation*, Oxford University Press, New York 2007, pp. 138-149.

Kripke S., *Naming and Necessity*, Blackwell, Oxford 1980.

Lowe J., "How Real Are Artefacts and Artefact Kinds?", in M. Franssen et al. (eds.), *Artefact Kinds. Ontology and The Human-Made World*, Springer, Dordrecht 2014, pp. 17-26.

Malt B. C. and S. A. Sloman, "Category essence or essentially pragmatic? Creator's intention in naming and what's really what", *Cognition* 105 (2007a), pp. 615-648.

Malt B. C. and S. A. Sloman, "Artifact Categorization: The Good, the Bad, and the Ugly", in E. Margolis and S. Laurence (eds.), *Creations of the Mind: Theories of Artifacts and Their Representation*, Oxford University Press, New York 2007b, pp. 85-123.

Marconi D., "Wittgenstein and Necessary Facts", in P. Frascolla, D. Marconi, and A. Voltolini (eds.), *Wittgenstein: Mind, Meaning and Metaphilosophy*, Palgrave Macmillan, London 2009a pp. 140-166.

Marconi D., "Being and being called", *The Journal of Philosophy* 106 (2009b), pp. 113-136.

Marconi D., "Pencils Have a Point: Against General Externalism About Artifactual Words", *Review of Philosophy and Psychology* 4 (2013), pp. 497-513.

Marconi D., "Artefatti e oggetti sociali", in T. Andina and C. Barbero (ed.), *Ermeneutica, estetica, ontologia*, Il Mulino, Bologna 2016, chapter XVII.

Matan A. and S. Carey, "Developmental Changes within the Core of Artifact Concepts", *Cognition* 78 (2001), pp. 1-26.

McLaughlin P., *What Functions Explain*, Cambridge University Press, Cambridge 2001.

Millikan R. G., "In Defense of Proper Function", in *White Queen Psychology and Other Essays for Alice*, MIT Press, 1993, pp. 13-29.

Millikan R. G., "Wings, Spoons, Pills, and Quills: A Pluralist Theory of Function", *The Journal of Philosophy* 99 (1999), pp. 191-206.

Mingardo D., “Making (another) sense of a debate on artifact categorization”, in *Atti Convegno 2011 dell’Associazione Italiana Scienze Cognitive*, University of Trento, Trento 2011, pp. 77-81.

Neander K., “The teleological notion of ‘function’”, *Australasian Journal of Philosophy* 69 (1991), pp. 454-468.

Nelson J., “Schwartz on Reference”, *The Southern Journal of Philosophy* 20 (1982), pp. 359-365.

Peacocke C., *Being Known,* Clarendon Press, Oxford 1999.

Plunkett D., “Which Concepts Should We Use?: Metalinguistic Negotiations and The Methodology of Philosophy”, *Inquiry* 58 (2015), pp. 828-874.

Plunkett D. and T. Sundell, “Disagreement and the Semantics of Normative and Evaluative Terms”, *Philosophers Imprint* 13, n.23 (2013), pp. 1-37.

Preston B., “Why is a Wing Like a Spoon; A Pluralist Theory of Function”, *The Journal of Philosophy* 95 (1998), pp. 215-254.

Price H., *Naturalism without Mirrors*, Oxford University Press, New York 2011.

Putman D., “Natural kinds and human artifacts”, *Mind* 91 (1982), pp. 418-419.

Putnam H., “Meaning and Reference”, *The Journal of Philosophy* 70 (1973), pp. 699-711.

Putnam H., “The Meaning of ‘Meaning’”, in *Mind, Language and Reality*, *Philosophical Papers* vol. 2*,* Cambridge University Press, Cambridge 1975, pp. 215-271.

Rips L. J., “Similarity, typicality and categorization”, in S. Vosniadou and A. Ortony (ed.), *Similarity and Analogical Reasoning*, Cambridge University Press, Cambridge 1989, pp. 21-59.

Russell B., “On Denoting”, *Mind* 56 (1905), pp. 479-493.

Schwartz S., “Putnam on Artifacts”, *Philosophical Review* 87 (1978), pp. 566-574.

Schwartz S., “Natural Kinds and Nominal Kinds”, *Mind* 89 (1980), pp. 182-195.

Schwartz S., "Reply to Kornblith and Nelson", *The Southern Journal of Philosophy* 21 (1983), pp. 475-479.

Searle J., "Proper Names", *Mind* 67 (1958), pp. 166-73.

Searle J., *The Construction of Social Reality*, Simon and Schuster, New York 1995.

Sidelle A., *Necessity, Essence, and Individuation: A Defense of Conventionalism*, Cornell University Press, Ithaca 1989.

Soames S., *Philosophical Analysis in the Twentieth Century, Vol.2: The Age of Meaning*, Princeton University Press, Princeton 2003.

Sperber D., "Seedless Grapes: Nature and Culture", in E. Margolis and S. Laurence (eds.), *Creations of the Mind*: *Theories of Artifacts and Their Representation,* Oxford University Press, New York 2007, pp. 124-137.

Strawson P. F., *Individuals: An essay on descriptive metaphysics*, Methuen, London 1959.

Thomasson A. L., "Realism and Human Kinds", *Philosophy and Phenomenological Research* 67 (2003), pp. 580-609.

Thomasson A. L., "Artifacts and Human Concepts", in E. Margolis and S. Laurence (eds.), *Creations of the Mind*: *Theories of Artifacts and Their Representation*, Oxford University Press, New York 2007a, pp. 52-73.

Thomasson A. L., *Ordinary Objects*, Oxford University Press, New York 2007b.

Thomasson A. L., "Norms and Necessity", *The Southern Journal of Philosophy* 51 (2013), pp. 143-160.

Thomasson A. L., "Public Artifacts, Intentions, and Norms", in M. Franssen et al., *Artefact Kinds. Ontology and the Human-Made World*, Springer, Dordrecht 2014, pp. 45-62.

Thomasson A. L., "Metaphysical Disputes and Metalinguistic Negotiation", in *Analytic Philosophy* 58, 2016, pp. 1-28.

Thomasson A. L., "What can we do, when we do metaphysics?", in G. D'Oro and S. Overgaard (eds.), Cambridge University Press, Cambridge 2017, pp. 77-100.

Thomasson A. L., *Norms and Necessity*, *forthcoming*.

Van Inwagen P., *Material Beings*, Cornell University Press, Ithaca 1990.

Vega Encabo J. and D. Lawler, "Creating Artifactual Kinds", in M. Franssen et al. (eds.), *Artefact Kinds. Ontology and The Human-Made World*, Springer, Dordrecht 2014, pp. 105-124.

Vetter B., *Potentiality. From Dispositions to Modality*, Oxford University Press, Oxford 2015.

Wiggins D., *Sameness and Substance*, Blackwell, Oxford 1980.

Wiggins D., *Sameness and Substance Renewed*, Cambridge University Press, Cambridge 2001.

Williamson T., *The Philosophy of Philosophy*, Blackwell, Oxford 2008.

Wittgenstein L., *Philosophical Investigations*, Blackwell, Oxford 1953.