Organon F 25 (4) 2018: 539-563

Theories of Reference and Linguistic Relativity

ANTONIO BLANCO SALGUEIRO[1]

ABSTRACT: The challenge to traditional theories of reference posed by experimental philosophers puts the focus on the question of diversity, cultural and linguistic, on the one hand, and cognitive (on intuitions), on the other. This allows for a connection between the problem of reference and the language-thought relation debate, and the linguistic relativity hypothesis conceived as the idea that linguistic diversity causes a correlative cognitive diversity. It is argued that the Kripkean view on proper names and natural kind terms is probably universal and that this empirical fact has plausible consequences for the universality of certain forms of human thought, but that there are nontrivial differences in the details of the workings of these expressions in different languages and that those differences influence the ways of thinking of speakers about individuals and kinds.

KEYWORDS: Language – linguistic diversity – linguistic relativity – names – reference – thought.

1. Introduction

The motivation underlying this work is the challenge posed by experimental philosophy to the theories of reference in Machery, Mallon, Nichols

[1] Received: 20 April 2018 / Accepted: 14 September 2018

✉ Antonio Blanco Salgueiro

Department of Logic and Theoretical Philosophy, Faculty of Philosophy
Complutense University of Madrid, 28040 Madrid, Spain

e-mail: ablancos@filos.ucm.es

& Stich (2004) and a long series of further articles.[2] They argue that the intuitions invoked by philosophers (according to those authors, as the sole evidence for their views) aren't universal, but vary across cultures and, in particular, differ substantially when comparing people from the West with people from the Far East.[3] Their experiments appear to show that when presented with stories such as Kripke's Gödel-case, Chinese participants tend to have descriptivist intuitions, while Americans tend to have Kripkean intuitions (Mallon *et al.* 2009, 34). The data also appear to show that the diversity is not only cross-cultural but also intra-cultural: 45% of Americans gave descriptivist answers (as did most Chinese), and 30% of Chinese gave Kripkean answers (as did most Americans). The subsequent work of this group of philosophers respond to a huge critical literature. The controversy doesn't only reach philosophy of language. The point is metaphilosophical, on the role of intuitions in the philosophical endeavor at large, and on the very idea of *intuition*.[4] Their aim is to question the project of constructing a theory of reference or any other philosophical theory (in fields like ethics, epistemology, etc.) taking as evidential ground the intuitions of lay people (or of experts). Some of the criticisms to Machery *et al.* seem fair to me, and I shall not enter into some of the more heated debates. In my view, the crucial point in their challenge is to place philosophy of language's focus on the problem of diversity. Firstly, on linguistic diversity, and secondly, on the possibility of a correlative cognitive diversity (after all, intuitions are mental states of speakers). This is what allows us to connect the controversy about experimental philosophy with the question of linguistic relativity, conceived as the idea that linguistic diversity (differences in the ways of speaking) brings along a correlative cognitive

---

[2] Some works of experimental philosophers are Weinberg *et al.* (2001), Machery *et al.* (2004), Mallon *et al.* (2009), Machery *et al.* (2009), Machery *et al.* (2010); Machery (2012); Machery *et al.* (2013), Machery *et al.* (2015), Machery, Sytsma & Deutsch (2015), Nichols *et al.* (2016), Stich & Tobia (2016).

[3] Mallon *et al*. (2004) rest on the ideas of cultural psychology (Nisbett 2003; Nisbett *et al.* 2003).

[4] This is just one sample of this critical literature, Liao (2008); Deutsch (2009); Jylkkä *et al.* (2009), Martí (2009); Lam (2010); Ludwig (2010); Ichikawa *et al.* (2011); Devitt (2011, 2012); Knobe *et al.* (2012); Vaesen *et al.* (2013); Sytsma *et al.* (2015); Nado & Johnson (2016); Heck (2017); Hannon (2017).

diversity (differences in the ways of thinking). The main idea that I intend to explore is that the differences in the mechanisms for the reference of certain types of expressions create differences in the ways in which humans mentally refer to individuals and kinds.

In Section 2, I highlight the empirical character of the claim that the mechanisms for reference are universal. Then, I present two languages in which names function in a different way from English (Section 3). The claim is not that they work in a non-Kripkean way, but that there are differences in the details of the mechanisms involved. In Section 4, I present the argument for linguistic relativity. Based on its first premise, in Section 5, I argue for a link between the universality of the Kripkean character of names and the universality of counterfactual thought, which amounts to a semiotic effect of *any* language on human cognition. In Section 6, I distinguish different kinds of linguistic diversity, and in Section 7 I connect them with possible differences in the referential mechanisms that could in turn cause differences in referential thought. Some implications are finally drawn from this hypothetical impact of language diversity on thought diversity.

2. Kripke and the universality of the mechanisms of reference

Does Kripke offer an account of proper names which claims universal validity, i.e, that intends to explain how proper names work in all human languages, real or possible, past, present or future? Of course, the same question may be addressed to an advocate of any other theory of names, in particular to any type of descriptivist, or to hybrid views. Most of Kripke's readers will answer affirmatively and assume that Kripke and they themselves aren't just studying how proper names work in English, although he uses only English examples and is not very explicit on this point.

Many linguists that have renewed the interest in linguistic diversity (Crystal 2000; Evans & Levinson 2009) highlight that throughout the decades of full Chomskyan hegemony, it was normal for someone to have a successful career in linguistics without studying or being fluent in any language except English, let alone in (what for us is) an exotic language. This holds still more accurately in the field of philosophy of language, where anything that aspires to have some relevance should be written in English.

But are we sure that we are doing philosophy of language and not philosophy of English?

One option would be to ground the answer in human biology. Are we all born with a disposition to use names in a causal-historical way? This nativist Kripkeanism could apply the "argument from the poverty of stimulus" and argue that without enough evidence children would begin to use proper names in a Kripkean way, not in a descriptivist way. This could be true, although I think it is not so.[5] Still, we would need an evolutionary account of why humans acquired this predisposition in the phylogenesis (that manifests itself in the ontogenesis), that is, what advantages our Kripkean ancestors had over their descriptivist rivals that allowed them to reproduce more profusely. Or this may be a brute fact that occurred by chance: humans could have been descriptivists, but they (or most of them, anyhow) just happen to be Kripkeans. In any case, Kripke never mentions human biology. In the Preface to *Naming and Necessity* he says that proper names are rigid designators *de jure* because their reference "is *stipulated* to be a single object, whether we are speaking of the actual world or of a counterfactual situation" (Kripke 1972/1980, 21). The question is what guarantees that in every language this stipulation for the use of proper names is made? Talking of "stipulation" seems to imply that things could have been different: we could decide not to stipulate this and stipulate another thing instead. If this is a universal rule present in every one of the over 6000 languages now spoken around the world and in the much larger number that have ever existed, there must be some universal pressure that accounts for it, and the search for it should be a central concern for a theory of reference. Most philosophers agree that the reference of proper names is not a biological question, but depends on conventions, rules, practices or language games. However, we find diversity in other fields with these same features. Why not in the rules that establish how proper names refer? Just before the famous passage where he exposes his new picture, Kripke admits that we are free to stipulate that our names work as the descriptivist says that they in fact work. There is then no human (let alone physical,

[5] I don't mean to deny that children are born with a pre-linguistic capacity for individuating objects, but only that this capacity determines by itself the correct theory of reference for every human language.

metaphysical or logical) necessity that guarantees that all humans will stipulate that names work causal-historically:

> So what makes my use of 'Cicero' into a name of *him*? The picture which leads to the cluster-of-descriptions theory is something like this: One is isolated in a room; the entire community of other speakers, everything else, could disappear; and one determines the reference for himself by saying—'By "Gödel" I shall mean the man, whoever he is, who probed the incompleteness of arithmetic'. **Now you can do this if you want to. There's nothing really preventing it. You can just stick to that determination.** If that's what you do, then if Schmidt discovered the incompleteness of arithmetic you *do* refer to him when you say 'Gödel did such and such'. (Kripke 1972/1980, 91; my boldface added)

Here Kripke is clear that it is possible in principle to create names that work in the way the descriptivists think they work in English. His claim is that this is not how we use them as a matter of fact and that there are good reasons for using them as we do. He suggests that only a weird speaker would use his names this way, in a sort of private language. This would open the door to an individualistic descriptivism where a speaker applies her own descriptions without taking into account the descriptions of others. But what prevents a whole community from creating a Descriptiranto in which names work descriptively for every speaker? And what allows us to discard that in some actual languages, with no deliberate decision, but in the tacit way in which many conventions are established, names work descriptively? In fact, the descriptivist philosopher thinks that English is such a language; we can reverse the question: What prevents proper names from working causal-historically in some languages? I take this as being the basic challenge that arises from the controversy initiated by Machery *et al.* (2004).

My own answer is that it is possible that the existence of proper names that work causal-historically is a linguistic universal, but that we cannot take it for granted and that the question has to be decided empirically. Moreover, I believe that Kripke's view is right for most names in English, but it is good to remember that there are English-speaking philosophers that are descriptivists concerning the functioning of proper names in their own language. This shows the magnitude of the problem: if it is difficult to settle the question for languages in which scholars are fluent, the difficulty

can be higher for non-familiar languages, not to mention for human language in general.

Linguist Daniel Everett offers an analogy that illustrates the kind of functional account that could be alleged for the features that are found in all, most or many, languages: the analogy with the independent invention of the bow and arrow in many different parts of the world. It seems absurd to postulate an innate faculty for making bows. But there exists a general pressure, killing protein that moves faster than we do (Everett 2013, 17), which explains why smart beings like humans find the same solution again and again. We find an account of this type in the classic Putnam (1975). His first argument against descriptivist theories of natural kind terms is based on a universal linguistic fact (the division of linguistic labor) that depends, in turn, on a hypothetically universal human practice (the division of non-linguistic labor).

Kripke suggests that his picture follows from very general facts about humans and their relationship with language. In fact, Kripke seems to consider that it follows from something as general as the fact that language is a social more than an individual tool (Kripke 1972/1980, 163). Should this be right, we could refer the universality of the causal-historical theory to the universality of language as a social tool. After all, the problem of killing protein that moves faster than we do is not more obviously universal than some of the problems whose solution is alleged to involve a use of names according to the causal-historical model, like the problem of talking about non familiar persons or places with respect to which some or most of the members of the community can have false or non individuative beliefs, or the problem of talking about what could have happened to a person, different from what really happened to her, that is, the problem of considering counterfactual scenarios about particulars.

An important non-empirical part of Kripke's work belongs to what can be called *philosophical linguistic typology*; it differs from the typology of linguists in its interest in the workings of certain types of expression in *possible* languages. So, it is not directly conditioned by the empirical findings that could be alleged as a result of the study of specific natural languages. We don't need to visit the Amazonia to do this or even to give functional reasons related to the use of certain expressions in some human practices. We can do it from our comfortable chairs of Western philosophers, at the risk of not including possibilities that in fact occur in actual languages of the

world. It is at least remotely possible that philosophers don't have the extraordinary imaginative capacities that they so often assume to have, and that allows them to visit every corner of the logical space without leaving their offices. Kripke's typology of designators belongs to this part, and the very definition of "designator" as a wide category with several subcategories. He is clear in establishing that in this first part he is neutral with respect to what types of designator actually exist in a particular language like English, or with respect to the hypothesis that we will find the same classes in every human language, i.e., that some types of designator are *semantic universals*. As is well known, his main distinction here is between *rigid* and *accidental* designators, with a subdivision of the first type in rigid designators *de jure* and *de facto*. We could also include here the distinction between *semantic reference* (the one conventionally associated with the expressions of a language) and *speaker's reference* (linked to speaker's intentions, independently of the conventional use of the expressions he uses) (*cf.* Kripke 1977).

The second moment is *empirical* and much more controversial. Kripke advances an empirical thesis about proper names and other classes of expressions as we find them in English and perhaps in any natural language, although he admits the possibility of inventing artificial languages without rigid designators. It is here where the controversy arises over whether the intuitions of competent speakers are the only source of evidence that can decide the question of semantic diversity. Machery *et al.* claim that philosophers allege only those intuitions, while Devitt and others claim that the main source of evidence for a theory of reference does not come from intuitions (be they from lay people or from philosophers of language, that he thinks are bound to be better than those of the common folk), but from the overt (spontaneous or elicited) *use* of names by competent speakers (*cf.* Devitt 2011; 2012), a stance that I basically agree with. The priority would then be to discover if all humans use proper names in the same way.

Another empirical point is the view of how a proper name or other rigid designator is connected with its bearer. Kripke claims that descriptivism derives its plausibility (apart from solving some puzzles) from the fact that it proposes a *mechanism* that removes the apparent magical character of this link.[6] If we grant that Kripke offers an alternative mechanism for reference,

[6] There can be a lack of harmony between the mechanisms that really do the work and the intuitions of the speakers, if they are biased by cultural myths about language.

then we can see a way to connect the typological and the empirical parts (hybrid theories propose combinations of both mechanisms). Some formal types of designator may not have plausible mechanisms that could realize them in the real world. It is possible that a language for supernatural beings be not constrained by facts about the natural and social world, but a language for humans clearly is. One thing is that we can invent languages which contain by stipulation one, several or all types of designator, and another is whether these languages could work or be used by beings like us in a world like ours. From Kripke's work on, it is assumed that there are different mechanisms that could work in our physical world to back the reference. It is no longer the case that the descriptivist wins because otherwise reference would be mysterious. In fact, for Kripke there are in English both expressions that work causal-historically and expressions that work descriptively (most definite descriptions).

3. Two "exotic" languages: denk nicht, sondern schau! [Don't think, but look!]

I have suggested that there may be some cross-linguistic practices that can justify the universality claim of a theory of reference (in particular, that of Kripke's) for proper names and other kinds of expressions, such as the practice of talking of individuals or substances about which one has insufficient or erroneous knowledge, the practice of ascribing mental states about individuals or substances to a person with such a defective knowledge, or the practice of considering modal situations. However, we shouldn't take for granted that names are used in the same way everywhere, or that cultural practices don't have any impact on the semantics of this class of terms. I shall offer two examples taken from field linguistics. Unlike philosophers, linguists pay little attention to proper names or to the

This is one of the criticisms made to Machery *et al.*'s position (Martí 2009). Think of a cultural belief in the magical powers of names: you can influence someone through her name. For these people it would be natural to claim that there is a magical bond between a name and its bearer, beyond any description or something as prosaic as a causal-historical chain. But surely the more plausible real bond is one of those proposed by naturalist theories of reference.

problem of reference, but some of the things they say have obvious consequences for the theory of reference. Both examples show that proper names are used in other cultures in quite different ways from ours and that it is risky to claim that those differences will never affect the heart of our preferred theory of reference (there are probably more radical cases). I think that these kind of cases shed more light on the theory of reference (taken as a form of "experimental semantics") than the method of consulting the intuitions, which are not even mentioned in these studies (although it would be interesting to test the intuitions of speakers of these languages).

3.1. Jesus' name

The first case is taken from a study of the Amazonian language Pirahã (Everett 2008). Everett's controversial thesis (firstly proposed in Everett 2005) is that a cultural principle permeates the form of life of the Pirahãs and is responsible for many features in the grammar of their language, the *principle of immediacy of experience* (PIE), according to which the communication is limited to the immediate experience of speakers. The idea of immediacy doesn't imply something so radical as to stick to the present moment. An experience is immediate in Pirahã if it has been seen or recounted as seen by a person at the time of telling (Everett 2005, 622).[7] The argument for taking culture as the causal factor is that there are many rare or unique features of the Pirahã language that are formally very different but that can be connected and explained by this unique cultural principle.[8] The claim is that the Pirahã culture has a holistic impact on their language, showing up in various aspects of it as a coherent way of speaking in accord with the aforementioned principle. Everett doesn't say much about the use of proper names in Pirahã, but clearly the PIE affects them, given that it prevents talking about particulars with respect to which no member of the

[7] This cultural principle implies a restriction to the epistemology of testimony. Epistemology is important for the philosophy of reference. One of the disputed questions is whether we must know the referent well enough for a name to refer to it. Far from thinking that Pirahã's epistemology is defective, Everett argues that it is better than ours (which is, of course, questionable).

[8] The most controversial idea is that Pirahã lacks recursion altogether, but Everett also points to other shocking absences (and some presences) such as the lack of numbers, quantifiers or fixed color terms.

community has had a direct familiarity. For this reason, one class of names that responds well to the Kripkean view, the names of historical characters, doesn't exist in Pirahã. In fact, they don't have creation myths or a mythical or real history. Everett's life with the Pirahã people began as a project of evangelization aimed at converting them to Christianity, but he himself was converted to atheism, after not being able to convert a single Pirahã.[9] What is interesting for us is his failure to introduce the name "Jesus" in the naming practices of the Pirahã (they don't have a name for God either, and he used *Baíxi Hioóxio*, "Up-high father"). Many Westerners think that "Jesus" is like "Jonas", the name of a real but legendary man who wasn't born from a virgin, didn't work miracles, didn't resurrect, etc. When the Pirahã asked if he himself had seen Jesus and he tried to explain that Jesus had lived a long time ago, he couldn't get them to understand him or take him seriously. All the names of persons in this culture are names of live people or of someone closely known by a live speaker: "the Pirahãs believe only what they see. Sometimes they also believe in things that someone else has told them, so long as that person has personally witnessed what he or she is reporting" (Everett 2008, 266).

I don't want to claim that proper names don't work causal-historically among the Pirahã. That depends on the existence of practices like speaking of somebody using her name in cases of ignorance or error (which is less probable given their cultural emphasis on the evidence), ascribing mental states in these circumstances, or considering counterfactual scenarios (I assume that these arguments for Kripkeanism are correct); the PIE could also prevent all this, but I ignore it. What the case shows is that there can be restrictions to the causal-historical links between a name and its bearer, that is, to the intention of fluent speakers of making the reference of some types of expression depend on these links. What is more, proper names seem to work among the Pirahã in the same way as Strawson (1959, 82) describes our own use: some speakers, with respect to some names, can "pass the buck" to others who are supposed to know better, but there is always at

[9] The Summer Institute of Linguistics entrusted Everett with the study of Pirahã to translate the Bible into this language. According to him, "missionaries had been trying to convert them for over two hundred years", but "no Pirahãs are known to have 'converted' at any period of their history" (Everett 2008, 269).

least one member of the community that can give an accurate and individuative description of the referent.[10] In any case, part of what Everett says about the Pirahã, such as that regarding the complete absence of fiction or myths among them, point to severe restrictions to counterfactual discourse in this community, which can undermine the universal application of the "modal argument" for the causal-historical view on proper names. The question as to whether Pirahã names work descriptively or causal-historically (with strict epistemological restrictions) requires a more careful empirical scrutiny.

3.2. Naming in Madagascar

We travel from the Amazonia to Madagascar and from Pirahã to Malagasy, an Austronesian language with 18 million native speakers in 2007. My source now is Och (1976), a field work on the conversational practices in traditional communities in Madagascar. I shall consider only what she says about names, but her general aim is to question the universality of Grice's conversational maxims. Again, for cultural reasons, some of these practices affect the use of names for persons. The natives think that if someone is called by his name, the spirits can cause him harm through it. For this and other cultural reasons, they usually avoid the reference by the name, letting it remain implicit, or use descriptions as the equivalent of "the builder" or even general descriptions like "the person", so that it is normal for a mother to ask her son, referring to her husband: "Has the person already come?" For the same reason, they change their names when they suffer a misfortune. In recent times the authorities have forbidden

[10] An anonymous reviewer objects that the difference between English and Pirahã could be "a difference in the standards of testimonial justification rather than in the language". But it should be borne in mind that according to Everett the PIE has a holistic impact on the Pirahã language, affecting many aspects of its structure. So, this would not be a minor epistemic restriction on some autonomous linguistic mechanisms. This culturally entrenched epistemology is supposed to be widely codified in the language and constantly reinforced through its use. All I am adding here is that if Everett is right, the PIE also affects the rules for proper names. More generally, Everett's view is akin to the Wittgensteinian idea of the intertwining of language and life forms, against the idea of the autonomy of language with respect to culture (Everett 2005, 622). In Wittgenstein's terms, "Jesus" doesn't find a place in the language games of the Pirahã.

changing one's name more than three times in the course of a life, because in the past it was usual to change it six or seven times.

As before, we can ask what the implications are of all this for a theory of reference with universalist ambitions; at least, the functioning of proper names among these people is not exactly like ours, as the causal-historical chains that can be invoked are shorter and more dubious than the ones associated with our proper names. The feeling that there is no profound difference here with our linguistic practices stems from the assumption of an ontology of persons according to which a change of name does not imply a change of referent in these cases. But in many instances of name change in languages like this, the assumption is that the referent is not the same or exactly the same after the change; that is, there can be ontological differences associated to this linguistic diversity in the naming practices. To appreciate this, we can resort to some exceptional similar examples among us, like the "Cassius Clay" / "Muhammad Ali" case. A person who changes her name after her religious conversion can say very seriously that she is not the same person as before. The apparently trivial application of the law of identity in "Cassius Clay is Muhammad Ali" would be problematic if there were social consensus that the individuals are different. At the very least, in many cultures a change of name is linked to a change of social status, in the social identity of the individual. Here, language performatively creates new social reality (new social persons).[11] This idea that linguistic diversity implies diversity in the kinds of social reality that language

[11] The inconstancy in the use of proper names is present in other languages. In fact, we find it in Pirahã. Everett claims that once, after a prolonged absence, he addressed a pirahã using what he thought was still his name, and that the pirahã did not react. The following quote also illustrates other (for us) oddities in the institution of names among this people:

> One of the men, Kaaboogí, […] addressed me in very rudimentary Portuguese: "*Pirahã chamar você Xoogiái*" (The Pirahãs will call you OO-gi-Ai). I had received my Pirahã name.
>
> I knew that the Pirahãs would name me, because […] they name all foreigners, since they don't like to say foreign names. I later learned that the names are based on a similarity that the Pirahãs perceive between the foreigner and some Pirahã. Among the men there that day was a young man named Xoogiái, and I had to admit that I could see some resemblance. Xoogiái would be my name for the next ten years, until the very same Kaaboogí, now called Xahóápati, told me that my name

can create is one of the "new directions" in the study of linguistic relativity (Enfield 2015, 216). Again, even if the general mechanism for reference were causal-historical (allowing, for example, a certain type of modal discourse about individuals), there can be nontrivial differences across languages in the details of the implementation of this common mechanism, connected to different functions of words in cultural practices.

4. Linguistic relativity

Linguistic relativity is often defined (by its detractors) in ways that make it seem a radical and implausible idea. I take it simply as the claim that linguistic diversity (the different realizations of human language) has a nontrivial impact on cognitive diversity (the various styles of thinking in humans). In Blanco Salgueiro (2017) I provide a map of the many forms that this hypothesis may take, assuming that it is the conclusion of an argument whose two premises can be formulated in different ways. The radicalism or moderation of the hypothesis is the result of adopting one or other version of the premises. This is my reconstruction of the general argument, with many of the complexities in brackets:

Premise 1 *Cognitive Impact of Language* (CIL): Language [such and such aspects, features, levels or mechanisms of any particular language] has [strong or weak, qualitative or quantitative] effects on thought [in such and such cognitive domains; in the most classic version, on the interpretation or construction of reality] and on behavior.

Premise 2 *Linguistic Diversity* (LD): But the different languages [or linguistic variants] differ among them [little or much] in [some or all] the aspects that cause those cognitive or behavioral effects.

---

was now too old and that my new name was Xaíbigaí. (About six years after that my name was changed again to what it is today, Paóxaisi – the name of a very old man). As I learned, the Pirahãs change names from time to time, usually when individual Pirahãs trade names with spirits they encounter in the jungle. (Everett 2008, 9).

Conclusion *Linguistic Relativity* (LR): So, there exist cognitive and behavioral differences between speakers of different languages [or linguistic variants].

I shall not argue here for the plausibility of some of the argument's versions; my aim is just to connect it with the problem of reference. But three related points must be highlighted. Firstly, most current relativists reject deterministic versions of Premise 1 ("linguistic determinism"), and argue for a weaker but nontrivial *influence* of language on thought ("linguistic influencism"). Secondly, although often the question discussed is the relative priority of *language* and *thought*, probably the relationship is dynamic: not language affecting thought (L→T), or thought affecting language (T→L), but both interacting in complex ways (L↔T). Moreover, further factors, like culture, could play a key role. This last possibility has gained strength in linguistics, partly thanks to Everett's work on the Pirahã language which, in his view, defies for cultural reasons the most basic ideas about Universal Grammar (like the universality of recursion). But in his first controversial work, Everett explicitly rejects LR:

> […] against the simple Whorfian idea that linguistic relativity or determinism alone can account for the facts under consideration. In fact, I also argue that the unidirectionality inherent in linguistic relativity offers an insufficient tool for language-cognition connections more generally in that it falls to recognize the fundamental role of culture in shaping language. (Everett 2005, 623)

To this, a neo-Whorfian responds what will guide my following remarks:

> […] a language of course is a crucial part of a culture and it is adapted to the rest of it […] The question that neo-Whorfians are interested in is how culture gets into the head, so to speak, and here language appears to play a crucial role: it is learnt far earlier than most aspects of culture, is the most highly practiced set of cultural skills, and is a representation system that is at once public and private, cultural and mental. It is hard to explain nonecologically induced uniformities in cognitive style without invoking language as a causal factor. (Levinson 2005, 638)

Previously, in the Introduction to a classic in the neo-Whorfian literature (Gumperz & Levinson 1996, 1), the editors set a link between culture, language and thought in their very definition of LR as the idea that culture, *through* language, affects the way we think. I take this as the orthodox stance in modern defenses of LR. This is close to what Whorf himself claims in Whorf (1939), where he gives diachronic priority to the culture (he sees language as a cultural construction), but also insists that language is the strongest factor synchronically, accounting for how individual thought is affected by social factors. In non-biological approaches to language the distinction between language and culture is not neat.

Machery *et al.* (2004) assume without question that the key factor that explains the differences in intuitions is culture, not language. They are so sure that they don't see a problem in conducting their experiments in English, when comparing American English speakers and Hong Kong speakers, whose mother tongue is Cantonese, for whom English is a second language, and who could make transfers from their first to their second language, a well-known phenomenon in the study of second languages ("false friends"); so it is unclear that the differences are due to culture and not to language. Lam (2009) criticizes this part of their methodology. In reaction, Machery *et al.* (2010) repeat the experiments using the native tongue for each group; for the Chinese group the stories appear now in Chinese writing, common to Cantonese and Mandarin. They claim that the results are analogous to those of the original experiment. In my view, the most plausible hypothesis (following Levinson's argument) is that language diversity is the direct cause of the cognitive differences in this case, if they are confirmed, although the cultural forms of life may be the ultimate causes. If there were differences in the referential conventions associated to the designators of different languages, and taking into account the early acquisition and habitual character of the practices that involve the use of names, it is plausible that this has some impact on the differences in cognitive style, i.e., that we habitually think using the same conventionally established referential mechanisms that we use when we speak.

5. From linguistic reference to mental aboutness

But we shouldn't rush to accept that we will find diversity in this field. If it were true that there is a universal pressure that guarantees the universality of the mechanisms for reference, then the possible impact of language on thought could also be universal. That is, as in other domains, Premise 1 (CIL) of the argument for LR can be held independently of the truth of Premise 2 (LD). Many authors have claimed that the most important impact of language on thought is transversal to languages. It would be the fact that we are verbal beings and not the fact that we speak a particular language, which explains the human cognitive singularity (our capacity for planning, regulating our actions, thinking about thinking, nonmodular thinking, active thinking, etc.).[12]

Let's assume that the Kripkean view is correct and that this follows from our need to invoke it to explain how speakers talk counterfactually (they keep applying a name to a particular even when they imagine that the descriptions associated to it were false); or to explain how they refer to a particular using a name in cases of ignorance or error (and of semantic reference, not of speaker's reference); etc. Then, we can formulate a special case of Premise 1, the claim that human language is what allows or at least fosters or facilitates counterfactual thought, or thought about particulars in cases of ignorance and error in humans. The hypothesis predicts that nonverbal beings don't have, or are less good at, those kinds of thought.[13] The mechanism involved could be the same as the one invoked in much current research on linguistic relativity,[14] namely, *habit*. Habits afford a nontrivial

[12] This is what Lucy (1996) calls "semiotic relativity". Jackendoff (1996), Clark (1998) or Carruthers (2002) argue for a non-trivial impact of language on thought, but avoid completely the question of diversity (be it linguistic or cognitive), or even argue against it.

[13] When confronted to a correlation of language and thought phenomena, the advocate of CIL has to show that this correlation is at least in part the result of an impact in the direction language → thought. He doesn't need to (although he can) argue for something as strong as the thesis that some forms of thought originate with language. Perhaps language only augments or facilitates some pre-linguistic capacities.

[14] Counterfactual reasoning is one of the classic areas of research, although the usual focus is on grammatical features such as if-clauses or verb tense/aspect/mood (*cf.* Bloom 1981), not on names.

but non-deterministic impact of language on thought. In the present case, the constant use of linguistic resources for counterfactual discourse (by hypothesis, present in every human language) arguably fosters counterfactual thought habits that show up even when we are not thinking for speaking (and, of course, also when we are thinking for speaking or for understanding other people's speech). A key point is that this questions the idea defended by Fodor or Searle that in every domain *original aboutness* belongs to mental representations while linguistic expressions have only *derived aboutness*. The idea is that only by internalizing social linguistic practices in which names intervene, humans acquire the cognitive tools that are involved in at least some human forms of thought.[15]

6. A diversity of linguistic diversities

Premise 2 is a necessary step in the argument from the claim that language affects thought to the conclusion that this impact is not homogeneous. The following series of quotes illustrates the possible stances on the topic of linguistic diversity in linguistics. I add a first stance that may be tempting for philosophers (we find it in the *Tractatus*). Chomsky's universalism has at least the restriction of human nature:

Linguistic hyper-universality
There are some features that we can expect to find in any language, natural or artificial, human, divine or alien.

Linguistic (Chomskyan) universality
"We can be pretty confident that the different stages that are attained by the language faculty are only different in a superficial fashion and that each one is largely determined by the common language faculty. The reason for believing that is pretty straightforward. It is simply that relevant experience is far too limited." (Chomsky 2000, 6)

---

[15] The role of speech in the socialization of thought constitutes Vygotsky's fundamental idea (Vygotsky 1962). The internalized language that is used as a cognitive tool is for him the public language and retains many of its public characteristics. That would affect the use of names when we think in inner speech.

Cross-linguistic diversity
"In actuality, no person speaks 'language in general' but always a particular language with its own characteristic structure of meaning." (Lucy 1996, 41)

Radical cross-linguistic diversity
"The more we discover about languages, the more diversity we find." (Evans & Levinson 2009, 436)

Intra-linguistic diversity (diversity of linguistic variants)
"Strictly speaking, nobody speaks a language; we all speak a linguistic variety." (Moreno Cabrera 2000, 47)

Idiolectal diversity, and Intra-individual linguistic diversity
"There aren't two people that speak exactly the same way. Even the same person doesn't speak the same way during her life, or in different moments of the same day." (Bernárdez 1999, 26)

What is the case with proper names and other types of expressions for which it has been argued that Kripke's view is correct? We've seen that Kripke rejects A), but most philosophers accept B), although for functional more than biological reasons. The more obvious form of LD would be cross-linguistic: as Pirahã and Malagasy perhaps show, different languages can incorporate different mechanisms of reference for some terms. But possibly there is also intra-linguistic diversity (of dialects, sociolects or idiolects), and even intra-individual diversity. It could be the case that a speaker, even a typical one, associates two conventions with a name, one descriptivist, the other causal-historical, and that he uses them in a flexible way according to the context. Some experimental philosophers have defended recently that natural kind terms and names are ambiguous between a descriptive and a causal-historical reading (Nichols *et al.* 2016).

7. Diversity in referential mechanisms and relativity of mental aboutness

A recurrent topic in the work of experimental philosophers is that the variability in intuitions is too anarchic. I don't accept that speakers'

intuitions are the only available evidence for a theory of reference. But the evidence of whatever kind could convince us that there is a great linguistic diversity in the referential mechanisms. Let's assume that this diversity originates in cultural life forms, but reaches the uses of individual speakers largely through their acquisition of a public language. We have then different ways in which the general cognitive impact of language could vary according to the distinct forms of LD. The most obvious form would be the one that arises from cross-linguistic diversity. Do the Pirahã have only a problem with the name "Jesus" or are they incapable of thinking about Jesus? I find it plausible that they have difficulties with Jesus' thoughts mainly because of the way they talk. Another possibility (that could account for the percentages detected by Machery *et al.*) is that in some languages there are two internal varieties, one causal-historical and the other descriptivist, or hybrid, or whatever. Then, one group could talk and think in one way and another group could talk and think in a different way. A third possibility is that there are members of the community that aren't well acculturated or have atypical intentions, that is, cases of idiolectal diversity or just incorrect use; this could be the cause of cognitive idiosyncrasies. Finally, if some terms are ambiguous, this could affect the corresponding concepts. As I said, some experimental philosophers have recently argued that natural kind terms and names are ambiguous (it is not clear if in English or in any language) with a descriptivist reading and a causal-historical one:

> […] our proposal is that natural kind terms (and plausibly names as well) are ambiguous, such that in some cases the reference is determined descriptively and in other cases the reference is determined non-descriptively (Nichols *et al*. 2016, 160)

If this were so, we could have here a case of cognitive impact similar to the one that affects the expression "time" and the concept TIME. We can choose to describe or to think about a situation as "a too *long* talk" or as "*wasting* too much time", depending on our use of the metaphor TIME IS SPACE or the other metaphor TIME IS MONEY, both present in the repertory of conventional metaphors that we acquire through the learning of English, according to our communicative or cognitive purposes. The second metaphor at least is absent in many cultures and languages (where

money doesn't exist), so we also have cross-linguistic diversity here. Analogously, Nichols *et al.* argue that depending on the context we can use a natural kind term causal-historically or descriptively. A defender of Premise 1 would only have to add the hypothesis that this makes us capable of a correlative cognitive flexibility in the mental use of our natural kind concepts (and of switching our metaphysics from realism about natural kinds to a more Lockean metaphysics).

With this move, experimental philosophers finish the exploration of the varieties of diversity that linguists have distinguished. In footnote 25 they are explicit with respect to a change from an emphasis on C) and E) to an emphasis on G):

> In an earlier paper […] two of us reject the assumption that there is a single set of reference intuitions in the population. In that paper, we allowed that different people might have intuitions that support different theories of reference […]. But we did not explore the possibility that within each of us, there are (at least) two ways of thinking about the reference of kind terms. (Nichols *et al*. 2016, 161)

Blanco Salgueiro (2017) points out that this kind of LD (intra-individual LD) can be used to avoid the radical view known as "linguistic determinism". It is possible for a language to contain versatile enough tools that respond to the current context in flexible ways. Be it plausible or not, Nichols *et al.*'s proposal suggests a new relativistic hypothesis: languages with two referential conventions associated with proper names and other types of expressions not only allow for using them in two different ways when speaking, but also using their mental correlates in different ways when thinking. These different conventions surely come ultimately from different cultural practices.

8. Conclusion

I have tried to show that there is an important connection between the philosophy of reference and the language-thought relation debate, in particular with the controversy about the plausibility of the linguistic relativity hypothesis.

The implications of the discussion are far-reaching. Our ability to share thoughts with people who speak other languages (and the possibility of translating our language into theirs) is not in question if the main impact of language on thought were universal. If the mechanisms for reference are shared cross-linguistically, then it can be argued for a semiotic impact of language on thought: perhaps the ability to think counterfactually about objects and kinds depends in part on language, but it happens that *any* language has resources that promote this ability. Arguably, nonverbal beings have little or no capacity for displaced thinking (not attached to the actual situation), and language can contribute to explain this human cognitive singularity.

The hypothetical impact of intra-individual diversity evidences the flexibility of human thought, its capacity to change its frames to respond to the actual context. If your ability to think of an individual descriptively or causal-historically depends on your having two linguistic conventions (and this may not be a language universal), this also makes you capable of understanding both uses of the terms, although there is a risk of misunderstanding with others (or even with yourself in different moments) if you are applying a different convention from your interlocutor; in a Gödel-type scenario the referent will be different, depending on which convention is applied.

In the case of cross-linguistic or cross-variant diversity, there can be systematic differences in the habitual ways of thinking about individuals and kinds due to linguistic diversity. This does not necessarily mean that the misunderstanding is insurmountable. If language influence is a question of promoting particular habits of thinking (as argued by current linguistic relativists), then you can grasp other ways of talking and thinking paying more attention, dedicating more mental resources, or using your imagination. And, of course, you can learn other languages, or new linguistic rules. Cognitive habits are reversible, but can also be persistent, so that you have to make an effort to understand and pursue ways of thinking you are not used to. Nevertheless, as the Jesus-case shows, some features in the language games can be so entrenched that it is near impossible for a speaker of a language to think of an individual or kind in a way not permitted by her language. For instance, because there is no place in Pirahã for Jesus' name, it is very difficult to find a place for Jesus in a pirahã's mind. Here, the question is not (as in the Gödel case) which should be the referent of

the name in a counterfactual scenario, but the very possibility to refer to an individual that existed far in the past.

All these possibilities are consistent with current research on the language-thought relation and the linguistic relativity hypothesis. This research has paid little attention to names and their potential cognitive impact. More empirical work on this topic (in particular, field work on the conventions for the use of names in many languages) is needed to properly answer the questions addressed.

What about Machery *et al.*'s position? I see it as a hypothesis on the influence of particular languages (and only indirectly of particular cultural practices) on a certain cognitive domain: the intuitions of ordinary speakers in imaginary scenarios such as the Gödel-case. This hypothesis is controversial for reasons alleged by their critics; there could be other factors that explain the difference in intuitions: the extraordinariness of the imaginary cases, the influence of folk theories or myths about language, etc. The focus on intuitions seems wrong, if we try to set the influence of the linguistic mechanisms for reference on the cognitive mechanisms for aboutness. We should focus instead on the differences in the ordinary use of referential terms to settle the question of linguistic diversity, and on the possible influence of these differences in cognitive tasks that involve mental reference (like counterfactual reasoning). But, of course, it is possible (in fact, I take it as a good hypothesis) that some of the differences in intuitions are due, in part, to differences in the conventions for the use of proper names.

Acknowledgments

I am grateful to two anonymous reviewers for their helpful remarks on a previous version of this paper. The work has been supported by the Spanish Ministry of Economy and Competitiveness in the framework of the research project FFI2014-52244-P.

References

BERNÁRDEZ, E. (1999): *Qué son las lenguas.* Madrid: Alianza.

BLANCO SALGUEIRO, A. (2017): *La relatividad lingüística (variaciones filosóficas).* Madrid: Akal.

BLOOM, A. (1984): *The Linguistic Shaping of Thought: A Study in the Impact of Language on Thinking in China and the West*. Hillsdale, New Jersey: L. Erlbaum.

CARRUTHERS, P. (2002): The Cognitive Functions of Language. *Behavioral and Brain Sciences* 25, 657-674.

CHOMSKY, N. (2000): *The Architecture of Language*. Oxford: Oxford University Press.

CLARK, A. (1998): Magic Words: How Language Augments Human Computation. In: Carruthers, P. & Boucher, J. (eds.): *Language and Thought: Interdisciplinary Themes*. Cambridge: Cambridge University Press.

CRYSTAL, D. (2000): *Language Death*. Cambridge: Cambridge University Press.

DEVITT, M. (2011): Experimental Semantics. *Philosophy and Phenomenological Research* 82(2), 418-435.

DEVITT, M. (2012): Whither Experimental Semantics? *Theoria, Revista de Teoría, Historia y Fundamentos de la Ciencia* 27(1), 5-36.

DEUTSCH, M. (2009): Experimental Philosophy and the Theory of Reference. *Mind & Language* 24(4), 445-466.

ENFIELD, N. J. (2015): Linguistic Relativity from Reference to Agency. *Annual Review of Anthropology* 44, 207-224.

EVANS, N. & LEVINSON, S. C. (2009): The Myth of Linguistic Universals: Language Diversity and Its Importance for Cognitive Science. *Behavioral and Brain Sciences* 32, 429-492.

EVERETT, D. (2005): Cultural Constraints on Grammar and Cognition in Pirahã. Another Look at the Design Features of Human Language. *Current Anthropology* 46(4), 621-646.

EVERETT, D. (2008): *Don't Sleep, there are snakes. Life and Language in the Amazonian Jungle*. London: Profile Books.

EVERETT, D. (2013): *Language: The Cultural Tool*. London: Profile Books.

GUMPERZ, J. J. & LEVINSON, S. C. (1996): *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press.

HANNON, M. J. (2017): Intuitions, Reflective Judgements, and Experimental Philosophy. *Synthese* http://doi.org/10.1007/s11229-017-1412-1.

HECK, J. G. (2017): Speaker's Reference, Semantic Reference, and Intuition. *Review of Philosophy and Psychology*, http://doi.org/10.1007(s13164-017-0362-3.

ICHIKAWA, J. MAITRA, I. & WEATHERSON, B. (2011): In Defense of a Kripkean Dogma. *Philosophy and Phenomenological Research* 85(1), 56-68.

JYLKKÄ, J., RAILO, H. & HAUKIOJA, J. (2009): Psychological Essentialism and Semantic Externalism: Evidence for Externalism in Lay Speakers' Language Use. *Philosophical Psychology* 22(1), 37-60.

JACKENDOFF, R. (1996): How Language Helps Us Think. *Pragmatics and Cognition* 4(1), 1-34.

KNOBE, J., BUCKWALTER, W., NICHOLS, S., ROBBINS, P., SARKISSIAN, H. & SOOMERS, T. (2012): Experimental Philosophy. *Annual Review of Psychology* 63, 81-99.

KRIPKE, S. (1972/1980): *Naming and Necessity*. Oxford: Blackwell.

KRIPKE, S. (1977): Speaker's Reference and Semantic Reference. *Midwest Studies in Philosophy* 2, 255-276.

LAM, B. (2010): Are Cantonese-Speakers Really Descriptivist? Revisiting Cross-Cultural Semantics. *Cognition* 115, 320-329.

LEVINSON, S. C. (2005): Comment to Everett (2005). *Current Anthropology* 46(4), 637-638.

LIAO, S. M. (2008): A Defense of Intuitions. *Philosophical Studies* 140, 247-262.

LUCY, J. A. (1996): The Scope of Linguistic Relativity: An Analysis and Review of Empirical Research. In: Gumperz & Levinson (eds.), 37-69.

LUDWIG, K. (2010): Intuitions and Relativity. *Philosophical Psychology* 23(4), 427-445.

MACHERY, E. (2012): Expertise and Intuitions about Reference. *Theoria, Revista de Teoría, Historia y Fundamentos de la Ciencia* 27(1), 37-54.

MACHERY, E., DEUTSCH, M., MALLON, R., NICHOLS, S., SYTSMA, J. & STICH, S. P. (2010): Semantic Intuitions: Reply to Lam. *Cognition* 117, 361-366.

MACHERY, E., MALLON, R., NICHOLS, S. & STICH, S. C. (2004): Semantics, Cross-Cultural Style. *Cognition* 92, B1-B12.

MACHERY, E., MALLON, R., NICHOLS, S. & STICH, S. C. (2013): If Folk Intuitions Vary, Then What? *Philosophy and Phenomenological Research*, 56(3) 618-635.

MACHERY, E., OLIVOLA, C. & DE BLANC, M. (2009): Linguistic and Metalinguistic Intuitions in the Philosophy of Language. *Analysis* 69, 689-694.

MACHERY, E., STICH, S., ROSE, D., CHATTERJEE, A., KARASAWA, K., STRUCHINER, N., SIRKER, S., USUI, N. & HASHIMOTO, T. (2015): Gettier Across Cultures. *Noûs* 51(3), 645-664.

MACHERY, E., SYTSMA, J. & DEUTSCH, M. (2015): Speaker's Reference and Cross-Cultural Semantics. In: Bianchi, A. (ed.): *On Reference*. Oxford: Oxford University Press, 62-76.

MALLON, R., MACHERY, E., NICHOLS, S. & STICH, S. (2009): Against Arguments from Reference. *Philosophy and Phenomenological Research* 59(2), 332-356.

MARTÍ, G. (2009): Against Semantic Multi-Culturalism. *Analisis* 69(1), 42-48.

MORENO CABRERA, J. C. (2000): *La dignidad e igualdad de las lenguas. Crítica de la discriminación lingüística*. Madrid: Alianza.

NADO, J. & JOHNSON, M. (2016): Intuitions and the Theory of Reference. In: Nado, J. (ed): *Advances in Experimental Philosophy and Philosophical Methodology*. New York: Bloomsbury, 125-154.

NICHOLS, S., PINILLOS, N. A. & MALLON, R. (2016): Ambiguous Reference. *Mind* 125(417), 145-175.

NISBETT, R. E. (2003): *The Geography of Thought: How Asians and Westerners Think Differently … and Why*. New York: Free Press.

NISBETT, R. E., PENG, K., CHOI, I. & NORENZAYAN, A. (2001): Culture and Systems of Thought: Holistic vs. Analytic Cognition. *Psychological Review* 108, 291-310.

OCHS, E. (1976): The Universality of Conversational Postulates. *Language in Society* 5(1), 67-80.

PUTNAM, H. (1975): The Meaning of 'Meaning'. In: Gunderson, K. (ed.): *Language, Mind and Knowledge, Minnesota Studies in the Philosophy of Science VII*. Minneapolis: University of Minnesota Press, 131-193.

STICH, S. & TOBIA, K. (2016): Experimental Philosophy and the Philosophical Tradition. In: Buckwalter, W. & Sytsma, J. (eds): *Blackwell Companion to Experimental Philosophy*. Oxford: Blackwell.

STRAWSON, P. (1959): *Individuals*. London: Methuen.

SYTSMA, J., LIVENGOOD, J., SATO, R. & OGUCHI, M. (2015): Reference in the Land of the Rising Sun: A Cross-Cultural Study on the Reference of Proper Names. *Review of Philosophy and Psychology* 6, 213-230, doi: 10.1007/s13164-014-0206-3.

VAESEN, K., PETERSON, M. & VAN BEZOOIJEN, B. (2013): The Reliability of Armchair Intuitions. *Metaphilosophy* 44(5), 559-578.

VYGOTSKY, L. (1962): *Thought and Language*. Cambridge, Mass.: MIT Press.

WEINBERG, J., NICHOLS, S. & STICH, S. (2001): Normativity and Epistemic Intuitions. *Philosophical Topics* 29(1&2), 429-460.

WITTGENSTEIN, L. (1953/1958): *Philosophische Untersuchungen / Philosophical Investigations*. Oxford: Blackwell.

WHORF, B. L. (1939): The Relation of Habitual Thought and Behavior to Language. In: Whorf, B. L. (1972): *Language, Thought and Reality* [1956]. Cambridge, Mass.: MIT Press, 134-159.