The influence of language in conceptualization: three views

Agustin Vicente, Ikerbasque: Basque Foundation for Science &
University of the Basque Country, UPV-EHU
Fernando Martinez-Manrique, University of Granada

In the fields of the lord, stood Abel and Cain
Cain slew Abel 'neath the black rain
At night he couldn't stand the guilt or the blame
So he gave it a name
So he gave it a name
So he gave it a name
Gave it a name
Bruce Springsteen

Abstract

Different languages carve the world in different categories. They also encode events in different ways, conventionalize different metaphorical mappings, and differ in their rule-based metonymies and patterns of meaning extensions. A long-standing, and controversial, question is whether this variability in the languages generates a corresponding variability in the conceptual structure of the speakers of those languages. Here we will present and discuss three interesting general proposals by focusing on representative authors of such proposals. The proposals are the following: first, that the effect of language in conceptualization is general and deep; second, that the effect is local, transient, shallow and easily revisable; and third, that there is no proper effect of language on conceptualization, although there is surely some cognitive impact of language: many conceptual tasks engage language one way or another.

Introduction

Many authors have held that language is necessary for thought: without language, there are no concepts. This language-first thesis is typically associated with German romantics, but it is by no means exclusive of them. Many philosophers of the analytic tradition have endorsed it, and it has even been claimed that the Sapir-Whorf hypothesis was inspired less by the romantics than by early analytic philosophers (Joseph 1996). However, the hypothesis that language somehow shapes thought, or our conceptual structure, does not require a thesis as strong as the language-first thesis. Many contemporary psychologists believe that non-linguistic babies have concepts, but some argue that the acquisition of language effects a kind of re-calibration of conceptual categories, and some hold that language is a source of concepts.

There are many issues that are usually run together in this debate. We want to steer clear from some of them. First of all, we want to avoid the issue of whether language is or is

not either a vehicle of thought[1]. If language is a vehicle of thought, then presumably language does have an impact on conceptual structure. However, language can have an impact on conceptual structure without being a vehicle of thought. Secondly, some authors have distinguished between two general ways in which language can have an effect on cognition: language-as-a-lens and language-as-a-tool-kit (see Gentner and Goldin-Meadow 2003; Gomila 2012). We will not be concerned with the issue of language-as-a-tool-kit, which has to do more with whether there are certain things that minds cannot do without using language. For instance, it has been claimed that language is involved in intermodular thinking (Spelke 2003), or that it is indispensable for conscious thought (Jackendoff 2012). These are not the kinds of issues we will be concerned with here.

Our focus is rather on the issue of language-as-a-lens. In a first pass, this is the issue of whether or not languages affect the way we conceptualize the world. But this is too broad. We want to focus on the issue of whether languages have some effect in conceptualization in those domains where it seems that there can be categorization without language. At least, it is required that there is a domain "out there" that can be categorized in different ways by different languages. This means, in particular, that we will leave aside the much-debated case of numbers and of arithmetic cognition. This case is one where, apparently, there cannot be number concepts –that is, a whole domain of concepts– without language because number concepts seem to go hand-in-hand with the linguistic symbols we use for them, instead of drawing on perceptual or core-conceptual information (see Carey 2009; Gordon 2004, 2010). If your language has incorporated natural numbers, you can have number concepts and count; otherwise, there seems to be no way to have any properly arithmetic ability at all. Cases like this do not sit well with the spirit of the "lens" metaphor, which is best exemplified by the color case. According to many, different languages carve the color spectrum in different ways, probably transforming a pre-existent way of categorization. In fact, the numbers case seems to exemplify better the "tool-kit" metaphor, since we seem to engage the linguistic system in counting.

2. Deep, transient, or no effects? Enthusiasts, moderates and skeptics

[1] For this issue, we refer the reader to Vicente and Martinez-Manrique (2005, 2008).

Some authors claim that language does not affect conceptualization at all, many others accept that it may affect conceptualization to some extent, and others want to hold that the effect of language on conceptualization is vast and deep. One initial problem is that it is not entirely clear what it takes for an effect to be deep. On the one hand, one could expect that if language had a profound impact, it would be roughly along the lines of the resilient changes that language's phonetics brings about in the perception of phonemes – as pointed out by Gleitman and Papafragou (2005) or Hespos and Spelke (2004). Yet the evidence shows that grammar and semantics seem to bring about changes in the conceptual structure that are quite unlike the changes that phonetics imposes. On the other hand, perhaps such a demand is too strong for the conceptual sphere, where – unlike the phonetics case– there are other factors that intervene in its organization. Here we will present and discuss three interesting general proposals: the enthusiasts, who hold that the effect of language in conceptualization is general and deep; the moderates, for whom the effect is local, transient, shallow and easily revisable; and the skeptics, who see no proper effect of language on conceptualization, although they admit some cognitive impact of language in the sense that many conceptual tasks engage language one way or another. These three seem to be the relevant contenders in the contemporary discussion, and for each of them we will select one the most recent representative lines of research: work from Casasanto et al. for the enthusiasts, from Landau et al. for the moderates, and from Li et al. for the skeptics. Of course, there are other areas of research, and we will mention briefly some of them, but they have been previously covered in other papers (e.g., Bloom and Keil 2001; Gleitman and Papafragou 2005). In addition, we will devote relatively more space to the enthusiasts' camp, given that it is here where most of the polemic points arise.

2.1. The enthusiasts

The initial revival of Linguistic Relativity (LR, from now on) purported to show that language affects adult categorization. For instance, Lucy (1992) tried to show that speakers of classifier languages (*vs*. speakers of languages with the count noun/ mass noun distinction)[2] categorized more in terms of substance than in terms of shape, while Levinson and his colleagues picked out the domain of spatial coordinates and compared

[2] Classifier languages count objects relying on classifier particles that vary according to the kind of object it is –animate/inanimate, substance it is made of, shape, etc– as in 'one [animal] pig'. Languages with count nouns can apply numerals directly to them ('one pigs, two pigs').

the categorization of speakers of languages with dominant egocentric relative coordinates *vs*. that of speakers of languages with only absolute coordinates (Levinson 1997, 2003).[3] It was held that these purported relativistic effects become manifest some time after children acquire language, a prediction implied by the thesis that language makes the difference (Lucy and Gaskins 2001). A new wave of studies by developmental psychologists suggests that children categorize the world in identical and universal ways until they master their language (Hespos and Spelke 2004).

There are plenty of domains of categorization that are under the scrutiny of experimental work now: from the classical domain of colors to the projection of grammatical gender on attributes of real-world things (Sera et al. 2002; Boroditsky et al. 2003; Kousta et al. 2008). So it seems that the evidence in favor of some degree of LR is growing. Yet, relativists still have to clear up two lingering worries about their research: 1) Do the data actually reveal that language affects non-linguistic conceptualization? 2) Can they really show that the correlation between linguistic differences and the alleged conceptual differences is due to linguistic differences causing conceptual differences? Let us develop these two issues.

Linguistic Relativity is a thesis about the impact of language on cognition. Most participants in the current debate take it that the issue is whether or not language shapes our conceptual structure. Experiments should be designed in such a way that the tasks subjects have to solve cannot be solved by linguistic means. They have to engage cognition, not language. We would not test LR if subjects were classifying objects or events in some domain by thinking "how does my language call these things/these events"? This is the worry related to whether data reveal that language affects non-linguistic cognition. We will see below that some authors doubt that classical experiments deal with this worry successfully.

The second worry is the following: many experiments show that there is a correlation between language and cognition. However, they do not show that this correlation is due to language's influencing thought. For instance, there are correlations between speaking a language with absolute coordinates and re-arranging, after turning 180º, an array of objects with respect not to the subject's own body but with respect to those coordinates

[3] Egocentric relative spatial coordinates are those that rely on the orientation of an object with respect to the bodily position of the speaker, as in 'take the road to your left'. Absolute coordinates rely on a stable frame of reference, such as a landmark ('take the road uphill') or cardinal points ('take the road north').

(Pederson 1995; Levinson 1997, 2003). Yet the effect may be due to a third factor, such as environmental circumstances (Bloom and Keil 2001).

We will concentrate on a recent line of research, developed by Casasanto (2008; Casasanto and Boroditsky 2008). Casasanto designed a series of experiments to test whether linguistic metaphors (e.g., talking about time in terms of spatial distance) leave a trace in non-linguistic tasks. He first shows that there is a pattern of interference from the spatial over the temporal cognitive domain. A typical task is: English speakers watch a line growing in the horizontal axis in a screen. When the image is removed, they make an estimation of the time the line took to grow by clicking a button twice. The task seems to be entirely non-linguistic: it does not recruit linguistic representations of time in terms of spatial distance. Yet, he observed that, keeping the time fixed, the longer the line was, the longer it was estimated to be the time of its growth. On the other hand, speakers of languages that do not use the time-as-distance metaphor as much as English did not show this pattern of interference in these tasks.

However, the most intriguing part of Casasanto's approach is the next step: in order to show that the observed correlation between language and the conceptualization of time is due to a causal influence of language, he trained his subjects for 30 minutes, making them use unusual spatial linguistic metaphors. English speakers were forced to use the time-as-amount metaphor, while Greek speakers, who make relatively little use of the time-as-distance metaphor, were forced to use this latter linguistic metaphorical mapping, instead of their preferred time-as-amount metaphor. Then he ran the experiments again, and the pattern of interference was reversed!

Casasanto (2008) claims that his results show not only that language shapes conceptualization but also that language has a *profound* influence in thought. It is an issue in this debate what mechanism could explain the effects of language on thought (see below). Casasanto puts the burden on associations, and explains how relativistic effects are possible, and how they can be overridden, in terms of Hebbian long-term strengthening of frequent associations and short-term adjustments due to immediate physical or linguistic experience.

Casasanto's claims raise some interesting issues. First, it is somewhat paradoxical that to show that language has a profound influence on thought, he has to show that such a profound influence can be reversed after only 30 minutes of training. Moreover, Casasanto claims that language is only a part, even if important, of the way experience shapes our way of thinking. For instance, cultural features such as directionality of

writing can also shape the way we think about time –as running from left to right *vs*. running from right to left (Casasanto and Bottini 2010). In fact, he takes it that "if language plays a privileged role in shaping thought, it is perhaps *only* by virtue of being a ubiquitous and highly systematic form *of experience*" (Casasanto 2008: 75, our italics). This means that what molds our thought is not language but experience,,which in turn means that thought can be molded by other forms of experience. So if Casasanto's claims about the linguistic metaphors of time could be generalized to other domains, language could still be considered at most *a* mold of thought.

On the other hand, the molding effects of linguistic experience seem to be easily reversed not only by manipulating linguistic experience, as Casasanto does. Data collected by Li and collaborators (Li and Gleitman 2002; Li et al. 2011) on frames of reference –where results can be reversed by changing the experimental conditions– point in this direction.

Now, there is another challenge for defenders of the deep effects thesis: the challenge of the mechanism (see Malt et al. 2003). They must explain how linguistic representations of categories in the world come to have deep effects in the way subjects conceptualize the world. Casasanto's suggestion of Hebbian associative learning is by no means the only mechanism that has been proposed. For instance, Boroditsky et al. (2003) suggest that the fact that we have a label makes us strengthen similarities between all those instances that fall under the label and thus reshapes the similarity-space. Evans (2009), in turn, talks about *linguistic indexing* as a process in which language interacts with the conceptual system in a language-specific way, producing thus linguistic relativity effects. Finally, Majid et al. (2004) gesture towards the idea of a representational re-description (Karmiloff-Smith 1992), and even a non-relativist like Fodor (1975) speculated with the idea of linguistic-based information chunking.[4]

An interesting proposal for a mechanism comes from the developmental psychology camp and puts the weight in the process of labeling. It has long been noted that linguistic labeling helps in categorization tasks. Waxman and Markow (1995) showed that 12-month-olds are sensitive to two objects being labeled either equally or differently. When objects have equal labels, children look for a categorical similarity at some level in the subordinate/superordinate hierarchy, whereas for objects with

[4] Representational re-description is a process by which information originally encoded in a representational format is "translated" to a different mode of representation. Chunking is the process of treating a collection of representations as a single unit.

different labels they create representations of two distinct individuals (see also Gelman and Markman 1986, 1987; Graham, Kilbreath and Welder 2004; Graham and Kilbreath 2007; Welder and Graham 2001; Waxman and Braun 2005). Xu (2002) has shown that these effects of labeling can be observed in 9-month olds: giving different labels to two different objects that go in and out of a screen, one at a time, makes babies "see" two objects instead of just one.

Labeling seems to have an effect in creating kind concepts, that is, concepts that are, as Carey (2009) puts it, "inductively deep" (see Weiskopf 2008). It is customary to speak of concepts as bodies of knowledge stored in long-term memory that we use to induce properties of new objects. Thus, it seems that labeling has a highly relevant role in generating a class of concepts that is distinctively important. In fact, it is so important that most talk of concepts concerns that class. For instance, in a classic experiment, Gopnik and Sobel (2000) introduced a "blicket detector" to two-year olds, along with four objects, two of which were called 'blickets'. One of the "blickets" was put on top of the detector, and a light turned on while a tune was sung. Then a "non-blicket" was put on top of the machine, and nothing happened. The children were told to choose another object that would turn the detector on: they took the other "blicket". That is, they induced that only "blickets" could turn the machine on.

One impressive result in this area is reported in Dewar and Xu (2009). Dewar and Xu found that the effects of labeling in creating categories arise as early as 10 months. A striking fact is that the effects of labeling can override not just perceptual similarity but even identity. Dewar and Xu presented different pairs of objects to different infants. The objects could be either perceptually dissimilar or identical. Then they introduced either a novel name applying to both objects, or two novel names, one for each. Thus, they had four conditions: dissimilar objects, one common name; dissimilar objects, two names; identical objects, one common name; and identical objects, two names. After that, the infants could hear one of the objects making a sound. The experiment tested the infants' expectancies about whether the second object should make the same sound or not in all four conditions.

Infants in a non-labeling condition expect two identical objects to make the same sound and two different-looking objects to make different sounds. However, in the labeling condition expectancies reversed dramatically: infants who heard two distinct labels expected the object pair to make different sounds, whereas infants who heard one

repeated label expected the object pair to make the same sound, regardless of whether they looked different or identical.

Xu (2007) thinks that infants expect count nouns to map onto kinds of objects –an expectation that leads them to use labeling as a source of evidence in identifying kinds. We take it that this means that labeling can contribute to create concepts (and this is in fact the usual interpretation of labeling: see Waxman and Markow 1995). These linguistically created categories may reach stability if they are able to support inductive generalizations. If we assume that languages do not usually carve the world in gerrymandered ways, it seems that word learning could have a profound effect in our conceptual system (but see Malt et al. 2011). Long et al. (2012, 241) summarize their results in the idea that "labels act as a kind of *conditional* ''cumulative cue''. If labels are the *only* evidence favoring a conceptual relation, and there is other conflicting evidence or simply insufficient support for the labels, children might ignore them". So, labels can be ignored, but they are not ignored if they carve the world in ways that are interesting to our inferential practices.

Labeling, thus, could explain at least some eventual relativistic effects (see Casasola 2005). If labeling has the role Xu and others suggest, and if linguistic categories are not gerrymandered, a good part of our kind concepts may be provided by our language, i.e., we may use linguistic categories for thinking. Even though there are other sources of evidence that we can use to acquire kind concepts, the importance of language eventually means that we may have many concepts that correspond to linguistic categories. In turn, there may be differences in the conceptual repertoires of speakers of different languages, given that not all languages carve the world in the same kinds[5]. In addition, there is apparently no reason why the effects of labeling should be restricted to kind concepts. In fact, Casasola (2005) argues that labeling affects spatial concepts[6], and Russell and Widen (2002) extend its influence to emotion recognition.

However, it is possible to rule out the importance of labeling on account of findings of Barbara Malt and colleagues. Malt et al. (2003) present two kinds of studies: on motion events, and on artifactual categories. Here we will focus on the second study (we describe motion events below). Malt et al. say that there is huge cross-linguistic

[5] However, note that if labeling proceeds in the way we have told, we are already justified in claiming that language shapes thought. The issue of LR comes afterwards: labeling + linguistic variability = LR.

[6] Casasola proposes that the relativistic effects in spatial relations (Korean *vs.* English) could be explained as results of the effects of labeling. See Hespos and Spelke (2004) and McDonough, Choi and Mandler (2003) on the issue of spatial relations.

variation in naming artifactual kinds. They collected up to 60 items that could be called 'jars', 'bottles' or plainly 'containers' in English, and compared English naming patterns against Chinese and Argentinean Spanish naming patterns, finding notable differences. However, when asked to group the objects based on their similarity, there was no significant difference between English, Chinese and Argentinean subjects. That is, their similarity space did not seem to have been altered by the naming patterns. This result eventually contradicts Borodistky et al.'s idea that by having a label we strengthen similarities between the objects that receive the same label, thus reshaping the previous similarity-space. However, it also seems to speak against the alleged influence of labeling in creating conceptual categories.

It has to be noted, nonetheless, that Malt et al. tested *similarity judgments*. The line we have been following does not entail that linguistically generated concepts have to have an impact on the perceptual similarity space. Not all concepts can be thought as partitions in the similarity space. In fact, what Dewar and Xu's experiment suggests is that we may generate concepts that ignore perceptual similarities. In their experiment, labeling makes children not categorize by similarity but by the assumption that dissimilar objects share hidden properties –since they seem to assume that if two objects receive the same label, they have some non-obvious properties in common. So one may wonder whether Malt et al.'s results do in effect contradict at all the idea that labeling can have an effect in our conceptual repertoire.

It is however important to note that the evidence told until now falls short of proving that labeling has robust effects on an adult's conceptual repertoire. In fact, it has not even been shown that it has effects on children's *conceptual* repertoire. It is suggestive to think that labeling has the effect of "opening files" that, if things go well, become information containers. Most people working on labeling point towards the same thing (one exception is Sloutsky and Lo 1999). However, labeling may have shallower effects. As we are about to see, some authors hold that language has only a temporary influence on thought. It may be that labeling too has only a temporary effect and that the categories we apparently begin to form are just short-lived proto-categories construed on-line for the task at hand. Although it seems that no researcher takes this possibility seriously at this point, we know of no follow-up experiments that test how long the effects of labeling last. (For instance, it would be good to test kids some days after the experiment in a verbal interference condition –so that they do not recall the label).

2.2. The moderates

The shallow and transient effect approach is best exemplified by Barbara Landau and colleagues. Landau et al. (2010) focus their research on space. They hold that language has a temporary effect on spatial cognition by means of two mechanisms: selectivity and enrichment. Language encodes certain distinctions and not others: this can have an effect on what people attend to in on-line cognitive tasks. This is the selectivity mechanism. Language also encodes some information in a form that is more robust than the way this kind of information is encoded by the visual-spatial system. Language thus can turn out to be more efficient than the proper visual-spatial system in storing and retrieving this kind of information: this is the enrichment mechanism. However, linguistic encodings are recruited and used in selecting and enriching spatial representations only for brief intervals –basically, in order to perform a determinate task. Their impact is in no way permanent, i.e., spatial cognition is not changed by language in any deep way.

Concerning selectivity, Landau et al. discuss spatial tasks where prima facie being a speaker of one language rather than of another could have some cognitive import. It is known that languages represent motion events differently in terms of Figure, Ground, Motion, Manner and Path (Talmy 1985). Some languages, like English, lexicalize Manner but not Path, which is left to the prepositional system; others, like Spanish, lexicalize Path, and not Manner, which is optionally expressed by gerundive forms. Since the contemporary revival of the issue of LR (Gumperz and Levinson 1996), these differences have been widely discussed (Slobin 1996; Malt et al. 2003; Papafragou et al. 2008, among others). In general, results suggest that there are no differences in conceptualization between speakers of a Manner-language and of a Path-language. Speakers of Manner languages interpret, memorize and retrieve motion events in basically the same way as speakers of Path languages do, unless they are told to verbally describe the events after watching them. One may conclude that this shows that language is used only in linguistic tasks but Landau et al. argue that in these cases language is used to allocate attention in elements of the motion event that would not receive as much attention if they did not have to be verbalized afterwards. However, this is done just for the sake of the task.

Something along the same lines occurs in the hotly debated case of reference frames. The controversy here is whether speakers of languages that describe spatial arrangements using an absolute frame of reference differ in their spatial cognition from

speakers of languages, like English, which predominantly use egocentric and allocentric frames of reference (see Pederson 1995; Levinson 1997; Li and Gleitman 2002, among others). Landau et al. think that human beings have different reference frames at their disposal, and that they have to select which one to use. Language can shift the attentional focus to one of them and make it salient for selection.

Following this line of thinking, Landau et al. report an experiment where they tested whether language could reverse an attentional bias that has been observed. Motion events, as well as events that consist in a transference of possession, typically take place between a source (a "from") and a goal (a "to"). However, in both cases –motion and transference– we tend to focus our attention more on the goal than on the source. Landau and colleagues gave verbal hints to 3-year-old children to see whether this pattern could be easily reversed. For instance, they told a group of children to describe a transference-of-possession event using the verb 'give', while another group was required to use the verb 'get'. They found no significant difference in the way the task was complied, which, according to them, means that language can be used to reverse our usual perspective or construal of events at almost no cognitive cost.

However Landau et al.'s most original research concerns enrichment. The visual system has well-attested problems at binding together some features of a visual display. For instance, we may watch a square divided in two halves by a vertical line, the right part green, the left part red. When, after a brief exposure, we have to pick out the square from a set of four pictures with different distributions of green and red, we have real problems to retrieve the right distribution of colors. We know the colors are green and red, but do not know where they go: we do not bind colors to the left and right parts of the square.

Landau et al. tested whether binding could be improved by language. They tested 4-year olds in several conditions: experimenters labeled the whole square ("this is a dax"), or they said "the red is touching the green", or they drew kids' attention to the red part ("look at the red!"), or, finally, they directly told them: "the red is on the left". Children improved over the no-stimulus condition only in the last of these conditions. What is intriguing, however, is that these kids did not master the right/left distinction. According to Landau and colleagues, this means that they are not storing just the linguistic representation –and discarding the visual representation. Rather, the situation seems to be one in which children's attention is directed towards the relation between the two colors and that they are able to solve the binding problem because their attention was so

directed. So, language serves to enrich our non-linguistic representations by drawing our attention to elements that it can easily encode. However, this effect is temporary: ten minutes later, children were unable to distinguish the square from its mirror image. Interestingly, perhaps this approach can be extended to explain Spelke's famous reorientation data. Experiments had shown that children, as well as most animals and adults in a verbal interference condition, performed poorly in a reorientation task where they had to integrate geometrical information about a room with color information about its walls. Spelke and colleagues (see Spelke 2003) and Carruthers (2002), as well as Mithen (1995), have proposed that language can be a sort of intermodular *lingua franca*, capable of integrating information coming from a number of different, relatively autonomous, domain-specific systems. Spelke proposed that we are capable of entertaining cross-domain thoughts thanks to the combinatorial system of our language. However, in one of their latest experiments, Spelke and colleagues (see Shusterman et al. 2011) found out that kids did much better if they were simply told "I'm hiding the sticker at the red/white wall" before the task.

It has to be noted, however, that the kind of use of language that Landau and colleagues point at may not always be beneficial. As we have seen, language may enrich visual/spatial representations, and allow subjects identify displays that are hardly identifiable without a previous verbal encoding. However, researchers have obtained the opposite pattern of results related to the phenomenon known as *verbal overshadowing* (Chin and Schooler 2008; Memon and Meissner 2002). Verbal overshadowing occurs when verbalizing mental contents deteriorates the performance of a task in which those contents appear to be involved. For instance, in a classical experimental setting, all subjects watch a video about a certain salient individual that they will have to identify afterwards. After watching the film and before testing their identification capacity, some subjects had to describe verbally the target individual while others had to read an unrelated text for the same amount of time. The results showed that the subsequent performance in recognizing the individual (e.g., picking him/her out of a line-up) was poorer for those subjects that had been asked to describe the individual. The phenomenon is robust in the domain of face recognition –where it was originally demonstrated– but Chin and Schooler (2008) report that it has been observed in domains such as decision making, problem solving, analogical reasoning, or visual imagery. More work needs to be done in order to know under what conditions and in which domains verbalization is helpful or unhelpful. It is particularly relevant to

uncover the reason why in some cases verbalization is an aid while in other cases it seems to be an obstacle.

2.3. The skeptics.

Many authors do not believe that language can shape our conceptual structure in any interesting way. Among them we find Malt and colleagues (Malt et al. 2003), Pinker (2007), Carruthers (2011), Gleitman and Papafragou (2005), Li and colleagues (Li and Gleitman 2002; Li, Dunham, Carey 2009; Li et al. 2011). We have already presented some of their criticisms in our exposition of the positions favorable to the linguistic shaping hypothesis in section 2.1. Here we will focus on a study by Li, Dunham and Carey (2009) because we think it is a good exemplification of the skeptic approach taken. The strategy is to impute all the putative conceptual differences between speakers of different languages to the use of language. That is, they try to show that none of the experiments done until now shows that language has an impact on non-linguistic tasks. The possibility that subjects are using language in one way or another when performing the experimental tasks, according to these critics, has not been ruled out, and in fact it is the most likely explanation of the observed differences. Since the beginning of the revival of LR, it has been assumed that nothing would ever count as evidence for LR if the task was resolved linguistically. So, critics want to maintain that there is as yet no evidence for LR.

Lucy (1992) was the first to argue that researchers in this area should be very cautious and design experiments such that they could be certain that the task to be solved was purely non-linguistic. Previous attempts at testing LR were clearly below the standard Lucy imposed, and most of them tested just language-to-language influences, a circular and non-informative thing. Lucy himself carried out a set of experiments that, according to him, met his own demands. As we explained above, some languages, like English, have the count/mass distinction, whereas others, like Tzeltzal, a Mayan language, have only mass terms. Lucy speculated that Tzeltzal-speakers would be more sensitive to substances than to forms, so he told his subjects to classify an array of objects in pairs that would be “the same”. Tzeltzal speakers tended to classify in terms of substance, while English speakers classified more in terms of shape.

Li et al. (2009) replicated Lucy’s experiments with languages with and without count terms, but they also run several other experiments in order to put Lucy’s results in context. Two are particularly significant. In the first, they presented subjects with a

group of items of different substances and shapes. Then they asked subjects about each of the items whether the item was an object or a substance. Even though they had found out that speakers of different languages tended to classify those very items following different patterns –similarity of shape vs. shared substance– responses to their question were uniform: speakers of different languages are by no means blind to the distinction between objects and substances.

The second experiment was an elaboration on Lucy's. Subjects had to classify items in terms of "being the same". However, instead of telling subjects: "look at this; which is the same?" (Lucy's original experiment), they asked them: "look at this blicket; which is the blicket?". That is, they introduced a neutral name in a neutral construction so that the name could be denoting either a substance or a countable object. The results were similar to those obtained in the replication of Lucy's experiment. Now, Li et al.'s hypothesis is that in this new experiment subjects make use of linguistic knowledge to answer the questions. In English, nouns are typically countable, so when a new name is introduced, English speakers tend to assume that it is referring to a countable object. In the Mayan language the pattern is reversed. They call this effect "lexical statistics". Now, they suggest that in the non-labeling condition, i.e., in Lucy's experimental setting, subjects could be also making use of lexical statistical knowledge: words like 'this' and 'same' may elicit associations similar to the associations elicited by nouns. So, it may be that, after all, what Lucy tested for was an influence of language over language.

Note, however, that language may be more active in our mental life than we think. For instance, take the case of color categorization. There has been a long debate about whether color categorization can be influenced by language. Recent research on lateralization studies, aided by fMRI experiments (see Regier and Kay 2009), suggests that there is indeed an impact of color language on color categorical discrimination, but that this impact is lateralized, restricted to the right visual field and so to the left hemisphere. In fact, it seems that the reason why language can have an impact on color categorization is that the language system itself is active in the color categorization tasks that involve the right visual field: Experiments that involve a linguistic interference, plus the aforementioned fMRI studies, give support to this hypothesis. Regier and colleagues, besides, look convinced that this pattern of results should be exportable to other categorization domains (Regier and Kay 2009, Regier et al. 2010).

Lately, it has been argued that even 2-year-olds make use of internal, probably unconscious, labeling all the time (Khan et al., submitted), and so that language is playing some role or another in cognition most of the time. This suggests that it may be difficult to separate purely cognitive tasks from tasks where language is involved, but also that it may be asking too much for the defender of LR to spot a purely cognitive task not influenced by language. If it happens that we use language one way or another all the time and for everything, doesn't that show that, after all, language shapes cognition?

Conclusion

Despite the evidence amassed in the last two decades, it seems simply impossible to render a verdict on the issue of whether language shapes our conceptual structure. We have told the arguments and evidence presented by representative authors and works of three approaches: enthusiastic, moderate, and skeptic. We saw that the three camps have found some support for their respective positions, but that the issue is far from having reached any kind of consensus. We also raised some questions that we regard as interesting or intriguing for each of the three contenders or related to their views. The literature on language and thought is now incredibly vast, growing exponentially, and has connections with very many other central issues in cognitive science. This, of course, is no surprise, because the role of language in conceptualization is one of the major questions in the research on the human mind. We think there is reason to be optimistic as regards the decades to come. In that, we are on the optimistic camp.

Acknowledgements

This paper is thoroughly collaborative. Order of authorship is arbitrary. The authors' research was funded by the Spanish Ministry of Economy and Competitiveness (FFI2010-15717 to A.V., FFI2011-30074-C02-01&2, to F.M. and A.V. respectively).

References

Bloom, P. and F. C. Keil. 2001. Thinking through language. *Mind and Language* 16: 351-67.

Boroditsky, L., L. Schmidt and W. Phillips. 2003. Sex, syntax and semantics. In Gentner and Goldin-Meadow (2003), 61-80.

Carey, S. 2009. *The Origin of Concepts*. New York: Oxford University Press.

Carruthers, P. 2002. The cognitive functions of language. *Behavioral and Brain Sciences* 25 (6): 657-726.

Carruthers, P. 2011. Language in cognition. In *The Oxford Handbook of Philosophy of Cognitive Science*, edited by E. Margolis, R. Samuels, and S. Stich, 382-401. New York: Oxford University Press.

Casasanto, D. 2008. Who's afraid of the Big Bad Whorf? Cross-linguistic differences in temporal language and thought. *Language Learning* 58: 63-79.

Casasanto, D., and L. Boroditsky. 2008. Time in the mind: Using space to think about time. *Cognition* 106: 579-93.

Casasanto, D. and R. Bottini. 2010. Can mirror-reading reverse the flow of time? In *Spatial Cognition VII* , edited by C. Hölscher, T. F. Shipley, M. Olivetti Belardinelli, J. A. Bateman, and N. S. Newcombe, 335-45. Berlin Heidelberg: Springer.

Casasola, M. 2005. Can language do the driving? The effect of linguistic input on infants' categorization of support spatial relations. *Developmental Psychology* 41: 183-92.

Chin, J. M. and J. W. Schooler. 2008. Why do words hurt? Content, process, and criterion shift accounts of verbal overshadowing. *European Journal of Cognitive Psychology* 20: 396-413.

Dewar, K. M., and Xu, F. 2009. Do early nouns refer to kinds or distinct shapes? Evidence from 10-month-old infants. *Psychological Science* 20: 252-57.

Evans, V. 2009. *How Words Mean: Lexical Concepts, Cognitive Models and Meaning Construction*. Oxford: Oxford University Press.

Fodor, J. 1975. *The Language of Thought*. Cambridge, MA: Harvard University Press.

Gelman, S . and E.M. Markman. 1986. Categories and induction in young children. *Cognition* 23: 183-209.

Gelman, S. and E. M. Markman. 1987. Young children's induction from natural kinds: The role of categories and appearances. *Child Development* 58: 1532-41.

Gentner, D. and S. Goldin-Meadow, eds. 2003. *Language in Mind: Advances in the Study of Language and Thought*. Cambridge, MA: MIT Press.

Gleitman, L., and A. Papafragou. 2005. Language and thought. In *Cambridge Handbook of Thinking and Reasoning*, edited by K. Holyoak and R. Morrison, 633-61. Cambridge University Press.

Gomila, A. 2012. *Verbal Minds: Language and the Architecture of Cognition***.** Elsevier: London.

Gopnik, A. and D. Sobel 2000. Detecting blickets: How young children use information about novel causal powers in categorization and induction. *Child Development* 71: 1205-22.

Gordon, P. 2004. Numerical cognition without words: Evidence from Amazonia. *Science* 306: 496-9.

Gordon, P. 2010. Worlds without words: Commensurability and causality in language, culture and cognition. In Malt and Wolff (2010), 199-218.

Graham, S. A., C. S., Kilbreath and A. N. Welder. 2004. 13-month-olds rely on shared labels and shape similarity for inductive inferences. *Child Development* 75: 409-27.

Graham, S.A. and C.S. Kilbreath 2007. It's a sign of the kind: Gestures and words guide infants' inductive inferences. *Developmental Psychology* 43(5): 1111-23.

Gumperz, J. and Levinson, S. eds. 1996. *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press.

Hespos, S. J. and E. S. Spelke 2004. Conceptual precursors to spatial language. *Nature* 430: 453-6.

Jackendoff, R. 2012. *A User's Guide to Thought and Meaning*. Oxford University Press.

Joseph, J. E. 1996. The immediate sources of the 'Sapir-Whorf hypothesis'. *Historiographia Linguistica* 23, 365-404.

Karmiloff-Smith, A. 1992. *Beyond Modularity: a Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT Press.

Khan, M., A. Geojo, S. Wang, and J. Snedeker (submitted) Spontaneous Verbal Labeling in 24-month-old Infants: http://scholar.harvard.edu/khan/publications/spontaneous-verbal-labeling-24-month-old-infants. (17/8/2013).

Kousta, S.-T., D. P. Vinson and G. Vigliocco. 2008. Investigating linguistic relativity through bilingualism: The case of grammatical gender. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34 (4): 843–58.

Landau, B., B. Dessalegn and A. M. Goldberg. 2010. Language and space: Momentary interactions. In *Language, cognition, and space: The state of the art and new directions*, edited by P. Chilton and V. Evans, 51-78. London: Equinox.

Levinson, S. 1997. From outer to inner space: linguistic categories and non-linguistic thinking. In *Language and Conceptualization*, edited by J. Nuyts and E. Pederson, 13-45. Cambridge: Cambridge University Press.

Levinson, S. 2003. Language and mind: Let's get the issues straight. In Gentner and Goldin-Meadow (2003), 25-46.

Li, P., L. Abarbanell, L. Gleitman and A. Papafragou. 2011. Spatial reasoning in Tenejapan Mayans. *Cognition* 120: 33-53.

Li, P., Y. Dunham and S. Carey. 2009. Of substance: The nature of language effects on entity construal. *Cognitive Psychology* 58: 487-524.

Li., P., and L. Gleitman. 2002. Turning the tables: language and spatial reasoning. *Cognition* 83(3): 265-94.

Long C., X. Lu , L. Zhang, H. Li , G.O. Deák. 2012. Category label effects on Chinese children's inductive inferences: modulation by perceptual detail and category specificity. *Journal of Experimental Child Psychology* 11: 230-45.

Lucy, J. 1992. *Language Diversity and Thought*. New York: Cambridge University Press.

Lucy, J. and S. Gaskins. 2001. Grammatical Categories and the Development of Classification Preferences: A Comparative Approach. In *Language Acquisition and Conceptual Development*, edited by S. Levinson and M. Bowerman, 257-83. Cambridge: Cambridge University Press.

Majid, A., M. Bowerman, S. Kita, D. B.M. Haun and S. Levinson. 2004. Can language restructure cognition? The case for space. *Trends in Cognitive Sciences* 8 (3): 108-14.

Malt, B., S. Sloman and S. Gennari 2003. Speaking versus thinking about objects and actions. In Gentner and Goldin-Meadow (2003), 81-112.

Malt, B. and Wolff, P. eds. 2010. *Words and the Mind: How words capture human experience*. New York: Oxford University Press.

Malt, B., Ameel, E., Gennari, S., Imai, M., Saji, N., and Majid, A. 2011. Do words reveal concepts? In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society,* edited by L. Carlson, C. Hölscher, and T. Shipley, 519-24. Austin, TX: Cognitive Science Society.

McDonough, L., S. Choi and J. Mandler. 2003. Understanding spatial relations: Flexible infants, lexical adults. *Cognitive Psychology* 46: 229–59.

Memon, A. and C. A. Meissner. 2002. *Special issue: Investigations of the effects of verbalization on memory*, *Applied Cognitive Psychology*, 16(8).

Mithen, S. 1995. *The Prehistory of Mind*. London: Thames and Hudson.

Papafragou, A., J. Hulbert, and J. Trueswell. 2008. Does language guide event perception? Evidence from eye movements. *Cognition* 108: 155-84.

Pederson, E. 1995. Language as context, language as means: Spatial cognition and habitual language use. *Cognitive Linguistics* 6: 33-62.

Pinker, S. 2007. *The Stuff of Thought*. New York: Viking.

Regier, T. and P. Kay. 2009. Language, thought, and color: Whorf was half right. *Trends in Cognitive Sciences* 13: 439-46.

Regier, T., P. Kay, A. L. Gilbert, R. B. Ivry. 2010. Language and thought: Which side are you on, anyway? In Malt and Wolff (2010), 165-82.

Russell, J. A. and S. C. Widen. 2002. A label superiority effect in children's categorization of facial expressions. *Social Development* 11: 30-52.

Sera, M. D., C. Elieff, M. C. Burch, J. Forbes and W. Rodríguez. 2002. When language affects cognition and when it does not: An analysis of grammatical gender and classification. *Journal of Experimental Psychology: General* 131: 377-97.

Shusterman, A., S.A. Lee and E. Spelke. 2011. Cognitive effects of language on human navigation. *Cognition*, 120: 186-201.

Slobin, D. I. 1996. From "thought and language" to "thinking for speaking." In Gumperz and Levinson (1996), 70-96.

Sloutsky, V. M. and Lo Y-F. 1999. How much does a shared name make things similar? Part 1. Linguistic labels and the development of similarity judgments. *Developmental Psychology* 35: 1478-92.

Spelke, E. 2003. What makes us smart? Core knowledge and natural language. In Gentner and Goldin-Meadow (eds.), 277-312.

Talmy, L. 1985. Lexicalization patterns: semantic structure in lexical forms. In *Language typology and syntactic description*, second edition. Vol. III: *Grammatical categories and the lexicon*, edited by T. Shopen, 57-149. Cambridge: Cambridge University Press.

Vicente, A. and F. Martínez-Manrique 2005. Semantic underdetermination and the cognitive uses of language. *Mind and Language* 20: 537-58.

Vicente, A. and F. Martínez-Manrique 2008. Thought, language, and the argument from explicitness. *Metaphilosophy* 39: 381-401.

Waxman, S. R. and I. Braun 2005. Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition* 95: 59-68.

Waxman, S. R. and D. B. Markow 1995. Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology*,29: 257–302.

Weiskopf, D. 2008. The origins of concepts. *Philosophical Studies* 140: 359-84.

Welder, A. N. and S. A. Graham 2001. The influences of shape similarity and shared labels on infants' inductive inferences about nonobvious object properties. *Child Development* 72: 1653-73.

Xu, F. 2002. The role of language in acquiring kind concepts in infancy. *Cognition* 85: 223-50.

Xu, F. 2007. Sortal concepts, object individuation, and language. *Trends in Cognitive Sciences* 11: 400-6.