
aracne

MARA FLORIS

LANGUAGE AND CATEGORISATION IN INFANCY

aracne

ISBN
979–12–218–0400–3

FIRST EDITION
ROMA 20 DECEMBER 2022

INDEX

INTRODUCTION

According to the Sapir-Whorf Hypothesis, also called Linguistic Relativity, the language we speak influences the way we think, at least, this is what states the most popular version of this theory.

The exact content of this theory is unclear; the name Sapir-Whorf Hypothesis itself is improper, as Edward Sapir and Benjamin Lee Whorf never proposed it as a co-authored theory, although Sapir was Whorf's mentor[1]. It is doubtless that Sapir had a significant influence on Whorf's work, but it is also undeniable that Whorf independently developed the hypothesis, this is the reason why it is often referred as the Whorfian hypothesis.

Whorf's most famous fragment states that:

> We dissect nature along lines laid down by our native language. The categories and types that we isolate from the world of phenomena we do not find there because they stare every observer in the face; on the contrary, the world is presented in a kaleidoscope flux of impressions which has to be organized by our minds and this means largely by the linguistic systems of our minds. We cut nature up, organize it into concepts, and ascribe significances as we do, largely because we are parties to an agreement to organize it in this way an agreement that

[1] It is also unclear who is responsible for the diffusion of the name Sapir-Whorf Hypothesis. Some claim that it is due to the linguists Eric Lenneberg and Roger Brown, some others claim that the linguist Harry Hoijer mentioned it in a paper. The high diffusion it had been due to the psychologist John Carrol (Koerner, 1992).

> holds throughout our speech community and is codified in the patterns of our language. The agreement is of course, an implicit and unstated one, but its terms are absolutely obligatory; we cannot talk at all except by subscribing to the organization and classification of data that the agreement decrees. We are thus introduced to a new principle of Relativity, which holds that all observers are not led by the same physical evidence to the same picture of the universe, unless their linguistic backgrounds are similar, or can in some way be calibrated. (Whorf, 1940)

Linguistic Relativity, initially, has been accepted as an undeniable fact and psychologists and sociologists studied it as an axiom; only in the Seventies, the increasing interest for psychological universalism cast some doubt on its validity. It has then been the rejected, especially in its stronger version which is linguistic determinism.

Linguistic determinism is the theory that states that everything we can think is determined by the language we speak. It is impossible to identify who proposed this version of Linguistic Relativity, as it cannot be inferred from Whorf's work, and no one ever claimed its authorship. Anyhow, it is rather implausible that language determines all our cognitive activity, and therefore it is not surprising that it was rejected. However, in the Nineties, a new interest for Linguistic Relativity rose, and the empirical investigations on the topic become kept growing, as witnessed by the publication of some volumes about it (Gumperz & Levinson, 1991; Niemeier & Dirven, 2000; Verspoor & Putz, 2000).

Nowadays, weaker versions of the hypothesis are still being tested, and some theoretical advancements were made. The existing studies tackle different areas where effects of Linguistic Relativity could be found; in particular, it is important to define what part of language affects what cognitive aspect.

A possible hypothesis is that language could affect reasoning, for example, counterfactual thinking could depend on the use of the subjunctive, which is not an element of every human language[2]. However, most of the existing studies focus on whether language can

[2] This hypothesis turned out not to be true (Au, 1983; Bloom, 1981; Liu, 1985).

impact perception, conceptualisation, and categorisation, which, as we will see, are not the same thing. Among the most common topics there are studies on colours (e.g., Franklin *et al.*, 2008; Regier *et al.*, 2007; Winawer *et al.*, 2007), on the effects of grammatical gender (e.g., Cubelli *et al.*, 2011), on object perception (e.g., Malt *et al.*, 1999) and motion perception (e.g., Athanasopoulos *et al.*, 2015).

The above-mentioned studies are different for topics and methods, but they all identify some common aspects of language on thought:

1. The effects of language are on-line; namely, they are active as long as the language is being used. It means that these effects disappear when participants of an experiment are given a verbal interference task.
2. Some effects are active as long as verbalisation is required; for example, during an experiment, some effects may be available only if the participant knows that she will have to give a verbal answer.
3. The effects of language on thought are not rigid. Language does not permanently and deeply affect cognition, as Linguistic determinism claims. The effects of language create habits rather than rigid schemes.
4. Bilingual speakers can switch from habitual schemes when they change the language.

Finally, a helpful distinction was draw by Lucy (1997), who identified three different levels at which language can influence cognition:

1. Having a language, any language, may affect thought in comparison with animals or pre-verbal infants.
2. Speaking a specific language could make a difference; for instance, English or Italian could affect cognition in different ways.
3. Inside the same language, there could be differences depending on the linguistic abilities of the speakers.

The cases described by Whorf, and most of the studies, usually address the second option, but the case study that I will analyse in this book belongs to the first and the third options. This is possible because

the participants of the studies I examine are mainly prelinguistic infants. The experiments, or at least some of them, compare the effects of labels on categorisation with categorisation in silence, and they also test whether having different labels shapes categories.

The debate on Linguistic Relativity intersects another debate, the one on the Cognitive Penetrability of Perception (CPP); the thesis of CPP is that perceptual experience can be influenced by our beliefs, desires, or mental states. It is a controversial and debated thesis both on the theoretical and on the empirical side on many different levels. On the theoretical side, for instance, CPP would have a crucial fallout on epistemology: if higher levels of cognition impact perception, its role as a «truth-preserving source of knowledge of the world» is not guaranteed (Vetter & Newen, 2014).

Those who claim cognitive impenetrability think that perception is a module (e.g., Carruthers, 2006; Fodor, 1983, 2000; Sperber & Wilson, 2002) and that its processes are encapsulated, which preserves their role as a source of reliable knowledge. On the contrary, those who claim that perception can be penetrated also deny the existence of modules in a strict sense and accept that knowledge is grounded on perception, even if there is not a truth preserving perception mechanism.

The debate on what does it mean that perception is penetrated led to a fine-grained description of what is perception, what is cognition and where is the boundary between the two; the most recent studies even started questioning the existence of such a boundary (see Beck, 2018; Burnston, 2017; Montemayor & Haladjian, 2017; Vetter & Newen, 2014). The core part of this debate is focused on whether early vision can be affected by higher processes; even if perception as a whole is discussed, most of the studies focus only on vision.

For the purposes of this book, it is important to keep in mind the existence of these two debates because they both help in framing the effect of labels on categorisation. Linguistic Relativity and CPP have an intersection: Linguistic Relativity holds that language can affect any level of cognition; CPP holds that perception is affected by higher levels of cognition. The common subset is the one where language affects perception. The experiments I will describe in this book belong to this intersection; they investigate whether only one specific aspect of

language – naming – can affect a specific cognitive process – categorisation.

Before discussing the effects of labels on categorisation it is fundamental to define what are labels and, in particular, what are categories. Concepts and categories are often used as synonyms, especially by psychologists, but it is worth to disambiguate their use.

The core issue I will deal with in this book is whether the effects on categorisation stem form top-down processes because labels refer, or if they can impact categorisation also in a bottom-up manner because can count as additional perceptual features. Following Plunkett *et al.* (2008), in this book I will rarely use the term names because names refer, and it may be the case the infants do not consider labels as referent yet. Label is a neutral term which does not imply any commitment on its role. Labels will be called names when there is evidence that they are used in a referential way.

The notion of concept is as pervasive in cognitive science as it is unclear. Machery (2009) described the currently available definitions of concepts and claimed that cognitive scientists should abandon the very notion of concepts and replace it with the terms which refer to what he calls the fundamental kind of concepts: prototype, exemplar, and theory. He defines concepts as:

> Within cognitive science, a concept of x is a body of information about x that is stored in long-term memory and that is used by default in the processes underlying most, if not all, higher cognitive competences when they result in judgments about x. (Machery, 2010, pp. 195-196).

Machery (2010) also describes the properties of concepts:

- Concepts can be about classes of objects (such as cat), events (such as running), substances (such as gold) and individuals (such as Immanuel Kant).
- Concepts can be used in multiple processing: they can be used for categorisation, induction, linguistic comprehension, and others.
- Concepts can vary over time and are different across individuals.
- Concepts are used by default by cognitive processes.

The experiments I will describe are mainly conducted with infants, the notion of concepts that is needed to describe their behaviour is minimal. What infants are required to do is to look at sets of images, or plastic toys, which are more or less similar, and then with an experimental procedure called novelty preference task it is assessed whether they consider some new items as familiar or not. Depending on their preferences, it is possible to infer whether they formed one or more categories.

There is no need to posit any form of representational content for these objects because there is no information stored other than their physical appearance. This is the reason why I am reluctant to claim that labels impact concepts, even if psychologists often state it. The effects of labels on categorisation may be the basis of concept learning, but this is a question which goes beyond the purpose of this book.

Although concepts and categories are often interchangeable, in this book I will try to keep them separate for the above-mentioned reason: when I claim that infants can categorise objects, I do not want to commit to the fact that they possess concepts for those objects, even if it is possible. Categorisation as a process is not identical to have categories. A minimal definition of categorisation needs to account for the fact that even infants can sort objects into classes, without possessing any information other than their physical appearance. Furthermore, the categories they form do not need to be stable over time.

Categorisation, which I will consider just as the ability to sort the objects into classes, is an essential process both for animals and human beings. To understand the categorisation as it is intended in the experiments described in this book, another useful distinction is the one between conceptual categories and perceptual categories (see Mandler, 2007). In the experiments infants are exposed only to the physical features of the objects, therefore, the only way to cluster them in categories is based on their physical similarity. They are tasks of object identification and not conceptual understanding.

It is plausible to think that infants can learn perceptual categories well before they display conceptual abilities, at least because this faculty is shared with other animals (Mandler, 2007). This does not mean that

infants do not possess conceptual abilities at all, what I am claiming is that it is not necessary to postulate any conceptual understanding to explain the results of the experiments discussed in this book.

The first chapter concerns the role of labels in a strict sense. I will first review the existing studies to show for the idea that the effects in categorisation depend on labels and not on their being sounds or language. Secondly, I will describe two effects, e grouping effect and a segregation effect. The second chapter reviews the theories which claim that labels act in a top-down manner. The third chapter, finally, will review the theories which claim that labels, instead, act as bottom-up stimuli, with particular attention to the use of neural networks as part of the explanations of these theories.

CHAPTER 1

ASSESSING THE ROLE OF LABELS

1.1. Labels and Other Auditory Inputs

When dealing with the role of labels in categorisation, the first step is assessing whether the effects of language, if any, actually depend on labels and not on a general auditory input. In this section, I will argue that the effects on categorisation initially depend on a broad variety of auditory stimuli that becomes increasingly narrow during development. By the second year of life, in fact, only count nouns affect the categorisation process. First, I will analyse the studies in which there is a comparison between the categorisation process in silence and the very same process (with the same stimuli) in the presence of a verbal label. Then I will consider the studies in which there is a comparison between the effects of sounds and those of non-labelling expressions. Finally, I will focus on the studies that highlight the specific role of count nouns as compared to language in general and adjectives. None of the studies conducted so far includes a direct and systematic comparison of all these variables.

The purpose of this section is to show that consideration of the existing literature and a comparison of the studies supports the claim that labels do play a role in categorisation. I consider 17 studies, published between 1995 and 2016, largely uniform with respect to their research design. Most of the experiments on this topic are eye-tracking studies with infants (3 to 26 months) using a novelty preference procedure; only two of them had a different research design.

The novelty preference task relies on the principle that infants show a preference for novelty. Usually, in the first phase of the experiments, infants are familiarised with a set of visual stimuli (such as drawings of animals) all belonging to the same category. Items are presented one at the time, they are either drawings shown on a screen or plastic toys. In the second phase, infants are tested with two new objects, one belonging to the familiarised category and one completely novel. If the participant shows a preference for the novel object, this is taken as a sign that the other object is considered similar to the familiarised examples, and it is meant to belong to the same category. If, on the contrary, the participants prefer the within-category object or shows no preference, it is inferred that the category presented in the familiarisation phase was not learnt. The experiments vary in the type of auditory stimuli presented in the familiarisation phase (e.g., tones, sentences, novel nouns) along with the visual items, and in the age groups that have been tested.

1.1.1. Experiments with a Silence Control Condition

The first result to assess is whether the presence of a label in the familiarisation phase makes a difference in the categorisation process when compared to a condition in which the stimuli were presented in silence. Only 3 out of the 17 studies considered in this section had a Silence control condition: most of them just compared the presence of a labelling expression to the one of a non-labelling expression or a sound. In order to state that labels have an advantage in categorisation, however, including a Silence condition is crucial. For example, in some studies, there seems to be an advantage of labels over sounds, but no Silence baseline is employed. In the absence of such a condition, the claim that labels enhance categorisation remains unsupported. It could be that sounds hinder categorisation (so-called "overshadowing effect", Best *et al.* (2011)) and that labels do not have any effect. The advantage of labels would then be only apparent and seem to be at work just because the comparison with categorisation in silence is lacking. The only three studies that overcome this problem are Plunkett *et al.*

(2008), Althaus & Mareschal (2014) and Althaus & Westermann (2016)[1].

In Plunkett *et al.* (2008) two sets of stimuli were presented in the familiarisation phase, a Broad and a Narrow Condition. All the stimuli were sketched animals that varied in the size of the neck, legs, tail, and ears. In the Broad Condition, the four features combined randomly, whereas in the Narrow Condition they were correlated (e.g., long neck with short legs and vice-versa) in order to form two clusters of stimuli. If the stimuli were presented in silence, the Broad Condition would lead to the formation of one single category and the Narrow Condition to the formation of two categories. The Narrow Condition yielded a binary categorisation again if paired with two consistent labels, while if the labels were randomly assigned, it was not possible to measure any proof of categorisation with the novelty preference task. Finally, if the Narrow Condition was paired with a single label, the stimuli were then considered as belonging to the same category. Althaus & Westermann (2016) used a similar research design: their set of stimuli consisted of drawings of invented animals, and it was possible to segregate them in two visual categories in much the same way as in the Narrow Condition used by Plunkett and colleagues. When the stimuli were presented in silence or with a single label, in the test phase, the overall average stimulus was considered familiar, and only one category was formed. When the stimuli were presented with two consistent labels, the two sub-category prototypes were considered familiar, and two categories were formed. When the stimuli were presented with two consistent sounds (a tingling bell and a xylophone tone sequence), then it was not possible to measure any preference at testing.

[1] Actually, there are other studies (Balaban & Waxman 1997; Haaf *et al.* 2003) in which some of the stimuli presented in the familiarisation phase are presented in silence, rather than with a sound or a linguistic expression, but in these experiments the stimuli presented silence are not tested with a separate novelty preference task. The only available finding is that there is a quick decrease of attention for the stimuli presented in silence when compared those presented paired with language. This lack of attention was usually considered enough to conclude that labels do have an effect as compared to silence, but they should have tested the categorisation process of the item presented in silence rather than accepting just the decreasing of attention in the familiarization phase as significant.

The studies just mentioned were both conducted on 10-month-old infants, whereas the experiments of Althaus & Mareschal (2014) concerned a group of 8-month-olds and one of 12-month-olds in four conditions: Silence, Labelling expression, non-labelling expression, Sound. With the first group, it was not possible to measure any categorisation of the visual stimuli presented. With the second group, instead, categorisation was achieved both with a labelling and a non-labelling expression, but not in the absence of any auditory stimuli or with a non-linguistic sound.

These three experiments show that categorisation occurs at least sometimes even in silence, but that labels can disrupt categories that would be formed otherwise or enable categorisation not taking place in silence. Appropriately, at least some of the above-mentioned experiments (Althaus Westerman 2016; Althaus Mareschal 2014) did have comparisons between a Silence condition, a Sound condition, and a Label condition. These cases show that the effect of labels is not only apparent, as it would be if sounds hindered categorisation.

1.1.2. Experiments Comparing Sounds

The number of experiments that included a condition in which a sound was compared to language is more substantial. Also, the variety of sounds used in these experiments is quite broad. Eight out of 17 studies in this section compared sounds to labelling and non-labelling expressions. It is crucial to prove that the facilitative effect on categorisation does depend on language (or labels) and not merely on the presence of any auditory input. In principle, any sound could help in focusing attention, thereby leading to a positive outcome in categorisation.

Balaban & Waxman (1997) had familiarised 9-month-olds with a set of visual stimuli paired either with a tone (a 400 Hz sine wave tone) or with a noun phrase ("A pig!" or "A rabbit!"). The proportion of infants looking at the novel object at test was higher for those in the Word condition than for those in the Tone condition.

Haaf *et al.* (2003) tested two groups of infants, 9-month-olds and 15month-olds. Each of the two groups was in turn split into two

conditions: basic-level and superordinate-level. In the familiarisation phase, they were exposed to some visual stimuli (20 plastic toys, animals, or vehicles) accompanied by a labelling phrase ("Look at the toma/bicket"), a five-note melody or non-labelling repetitive mouth sounds. The data suggest that labelling phrases facilitated global categorisation, but not basic-level categorisation (that was always achieved), over non-labelling sounds both at 9 and 15 months of age. There is also a sensitivity to the source of the auditory stimuli, and it undergoes some changes as infants grow up: 9-month-olds accomplish categorisation at global level, despite the source of the auditory input; 15-month-olds achieved global categorisation only when the experimenter directly uttered labelling phrases. According to the authors, the fact that basic-level categorisation was achieved despite the presence of an auditory stimulus may depend on the low perceptual variability among the stimuli: a higher perceptual similarity among stimuli makes the category easier to detect.

Similar results were reported by Fulkerson & Waxman (2007). They tested a group of 6-month-olds and a group of 12-month-olds with a set of figures depicting dinosaurs. The auditory stimuli were presented to half of the infants accompanied by a naming phrase ("Oh look, it's a toma/modi" or "Do you see the toma/modi?") and to the other half with two sequences of pure tones (400 and 800 Hz). Naming phrases were uttered by a female voice in the infant-directed speech register and recorded for presentation; the tone sequences were created to match the naming phrases in timing, duration, and volume. In the test phase, 12-month-olds in the Word condition demonstrated a reliable novelty preference, whereas those in the Tone condition performed at chance level; 6-month-olds showed the same effect.

A more recent study (Ferry *et al.*, 2013) used the same set of stimuli as Fulkerson & Waxman (2007), but with a group of 3/4-month-olds. Their results were similar to those of the previous study: labels do have a facilitative effect on categorisation that one does not achieve with tones. Finally, both Althaus & Mareschal (2014) and Althaus & Westermann (2016) had a Sound condition after which it was not possible to measure any preference in the test phase. All the

experiments considered here point in the same direction: sounds do not improve infants' performances in categorisation tasks.

1.1.3. Experiments with ecologically plausible sounds

The sounds used in the experiments discussed above were mainly tones. In this section, I will discuss some experiments in which other kinds of sounds were used. The reason why I keep them separate is that the complexity of this last group of sounds may make them ecologically more plausible. It may appear unsurprising that pure tones fail to affect categorisation, for they are usually not employed as communicative signals. Even if it is established that infants are able to detect their native language when they are born (J. Werker & J.F. Gervain, 2013), there might still be many variations in the kind of signal that affects their categorisation process. An ecologically plausible sound may thus be necessary to impact categorisation.

In a study already mentioned, Balaban & Waxman (1997) tested a group of 12/13-month-olds and one of 9-month-olds. They had a Tones condition, a Words condition, and a Content-filtered words condition. The content filtered words were obtained by filtering the original, computer digitised, phrases with an electronic filter system in order to remove high frequencies. These stimuli were recorded on tape for presentation and were matched in loudness to the other word phrases and tone sequence. man (1997) found that during the test phase the preference for the novel object was stronger for those who heard proper words; content-filtered words enhanced the preference for the novel item only if compared to tones, their effect was not as strong as words.

An interesting result was found by Hespos and Waxman (2013): they provided evidence for the idea that infants up to 4 months may accept non-verbal sounds as communicative signals. The set of stimuli they used is the same as Fulkerson & Waxman (2007), their participants were divided into three groups: 3-month-olds, 4-month-olds, and 6-month-olds. The three groups were tested in two different conditions: lemur vocalisations and backward speech. The lemur vocalisations were chosen because, although they differ from human