

## A FEW REMARKS ON THE RELATIONSHIP BETWEEN LANGUAGE AND THE CONCEPTUAL STRUCTURE OF THE MIND

### *ALGUMAS NOTAS SOBRE A RELAÇÃO ENTRE LINGUAGEM E A ESTRUTURA CONCEITUAL DA MENTE*

Claudia Uller<sup>1</sup>

#### ABSTRACT

In Psycholinguistics, particularly as the studies concern the language-concept relationship, it is necessary to take into account the different ways this relationship manifests itself, its nature and the direction it takes place. An objective to be pursued is to verify if (i) concepts exist spontaneously and naturally in human cognitive architecture and if language only has the role of instigating and installing them in the language; or if (ii) concepts depend on the language to exist. In the present article, we discuss the problem, consider different points of view from some proponents of alternative views and present a way to provide empirical evidence to verify their hypotheses.

**KEYWORDS:** language and thought, conceptual structure of the mind, origins of cognition

#### RESUMO

No estudo da Psicolinguística, em particular no que concerne a relação linguagem-conceito, há que se levar em conta as diferentes formas em que esta relação se manifesta, sua natureza e a direção em que ela se efetua. Um objetivo a ser atingido é verificar se (i) os conceitos existem espontânea e naturalmente na arquitetura cognitiva humana e se a linguagem tem apenas o papel de instigá-los e instalá-los na língua; ou se (ii) os conceitos dependem da linguagem para existirem. No presente artigo, discutimos o problema, analisamos as considerações de alguns proponentes com visões alternativas e apresentamos uma forma de prover evidência empírica para a verificação das hipóteses que eles levantam.

**PALAVRAS-CHAVE:** linguagem e pensamento, estrutura conceitual da mente, origens da cognição.

---

<sup>1</sup> Kingston University. Department of Psychology. Professor in Developmental Science. E-mail: c.uller@kingston.ac.uk.

*...if it isn't literally true that my wanting is causally responsible for my reaching, and my itching is causally responsible for my scratching, and my believing is causally responsible for my saying... if none of that is literally true, then practically everything I believe about anything is false and it's the end of the world.*

To Jerry Fodor (1935-2017), in memoriam

## Introduction

Philosophers and psychologists have for centuries been interested in the question of how the linguistic and the conceptual structures of mind interact, especially when babies learn knowledge, and acquire language. It is very much like the egg and chicken problem – which comes first? Do babies know of, say, the physical world before they acquire the words which label the concepts (e.g., ball, fall, roll, or do babies rely on language to represent concepts)?

At one end, those who claim that mental representations have a linguistic structure, propose that thoughts are expressed as language in the mind (Language of Thought Hypothesis (LOT), Fodor, 1975). Support for this kind of proposal rests on features of human cognition such as productivity, systematicity, and inferential coherence. These characteristics of human cognition can best (or **only**, if you are a strict Fodorian) be explained if this kind of proposal is true. This is also certainly true of the Quinean epistemological project – the child learns knowledge via what Quine (1960) proposed as “observation sentences”: the child hears “Frieda is a cat”, sees Frieda, sees a cat, and learns by observation of information which goes in to mind through the senses (visual, auditory). The Sapir-Whorf hypothesis, in its hard and soft versions, directly relates to the idea of a language of thought, and predicts that the structure of a language determines or greatly influences the modes of thought characteristic of the culture in which it is spoken.

Alternatively, psychologists and linguists have proposed that thoughts generated within a human culture can exist in the absence of a linguistic structure. We will dub this the Thought Without Language Hypothesis (TWL). One is quick to think that, for example, the same way our

ancestors expressed numbers of objects with the use of tallies (e.g., **I** for one, **II** for two, **III** for three, and so on), archaeological evidence suggests that hunter-gatherers as far back as 1.8 million years ago produced cave drawings and figurines/sculptures meant to represent animals, say, a bison, or a voluptuous woman/goddess. Whether Homo erectus had language or not to represent such concepts we can only guess. Another avenue to pursue in direct addressing these issues would be to investigate creatures whose cognitive architecture lack linguistic structure altogether.

## Sortal Concepts

Philosophers have coined the term "sortal" to refer to the concepts which underpin count nouns. A sortal is a concept which provides criteria for individuation and numerical identity. Criteria for individuation are the basis for counting; we cannot count without specifying a count noun such as "cat" or "table". Criteria for numerical identity – that is, that one individual is one and the same - are the basis for tracking individuals through time and space; we cannot ask questions about numerical identity without a count noun, e.g., "the same cat" or "the same table".

Within the past decades, research in developmental cognitive science has addressed the question whether pre-linguistic infants represent sortal concepts (e.g., Spelke, Kestenbaum, Simons, & Wein, 1995; Xu & Carey, 1996). Spelke et al. (1995) investigated whether young infants would individuate objects on the basis of spatiotemporal information. In a habituation paradigm, they presented 4-month-old infants one object emerging from the left edge of a left screen, and disappearing behind this screen. Following this emergence, a second identical object emerged from the right edge of the right screen, and went back behind the screen. No object appeared in between the screens. This sequence was repeated until a criterion of habituation was met. The screens were then removed to reveal either two objects (expected outcome) or one object (unexpected outcome). Four-month-old infants looked longer at the unexpected outcome than at the expected outcome. The researchers proposed that young human infants understood the events as involving two objects because they know that one object moves on spatiotemporally connected paths. They analyzed the spatiotemporal discontinuity and inferred that two numerically distinct objects should be behind the screens. Thus infants as young as 4 months seem to represent the sortal ‘object’.

However, the baby’s first words usually refer to more specific sortals such as ‘cup’ or ‘ball’

or ‘bottle’. Representing cup and ball as sortals require the infant to be able to use property/kind information to individuate objects and trace their identity. For example, upon seeing a cup then a ball, one at a time, adults infer two numerically distinct objects. Would infants do the same?

Xu and Carey (1996) were the first to argue that there is a general sortal, *object*, which may be the first sortal that infants represent. That is, young infants may only use spatiotemporal information to individuate objects and trace their identities over time. Spatiotemporal information includes the knowledge that one object cannot be at two places at the same time, two objects cannot be at the same place at the same time, and objects travel on connected paths. Thus, the baby sees object X at time 1, and object Y at time 2, and represent such objects as ‘object’ and not as ‘object X’ and ‘object Y’ as an adult would. What this entails is that the young infant would fail to represent numerical identity ( $X + Y = 2$ ) and would represent the occurrences at times 1 and 2 as two occurrences of the same ‘object’. To address this possibility, Xu and Carey conducted a series of experiments. The babies saw one object and then another object emerge from behind a screen; upon removal of the screen, the babies saw the two objects they had previously seen or only one object. Ten- and 12-month-old infants participated in experiments when either property/kind or spatiotemporal information was available for object individuation. Ten-month-old infants were presented one of three conditions: a baseline condition, a property/kind condition and a spatiotemporal condition. The baseline condition was devised to show whether there was any intrinsic preference for one-object outcomes or two-object outcomes. The property/kind condition tested the infants’ capacity to represent individual objects with the use of property/kind information. The spatiotemporal condition tested the infants’ capacity to use spatiotemporal information. All three conditions started with introductory trials. The babies were shown four trials in which they saw that there were objects (a cup and a toy camel) behind the screen, sometimes one object, sometimes two objects. In the property/kind condition, the infants were then given as familiarization trials a set of four emergences of each of two new toys (a ball and a toy duck). Specifically, one object emerged from behind a screen on one side and then returned behind the screen. A second object then emerged on the other side and then returned behind the screen (i.e. only one object was seen at a time). The objects were never seen coexisting at the same time. The screen was then removed to reveal one object (unexpected outcome) or two objects (expected outcome). After this first test trial, two other emergences of the objects took place, after which the second test trial followed. The whole procedure was then repeated with a second pair of toys (a

toy truck and a toy elephant). In the familiarization trials of the spatiotemporal condition, infants were presented with the same set of four emergences of each toy as for the property/kind condition, except that, on the last two emergences, the two objects were brought out from behind the screen simultaneously, providing spatiotemporal evidence that two numerically distinct objects existed at the same time. The screen was then removed to reveal either one object (unexpected event) or two objects (expected event). The whole procedure was repeated with a second pair of toys (a toy truck and a toy elephant) as in the property/kind condition. In the baseline condition, the infants were simply shown the outcomes of the experimental conditions without any familiarizations, namely, the screen was removed to reveal one-object outcomes or two-object outcomes.

Xu and Carey found that: (1) In the property/kind condition, 10-month-old infants did not look longer at the unexpected outcome of one object than the expected outcome of two objects, but 12-month olds did. This suggests that infants younger than 1 year of age were unable to use the property/kind differences to determine that more than one object existed behind the screen; (2) 10- and 12-month-old infants did look longer at the unexpected outcome of one object in the spatiotemporal condition, suggesting that they are able to use spatiotemporal information to individuate objects and determine that two distinct individuals existed behind the screen. Moreover, Xu and Carey (1996) found a correlation between 12-month-olds' comprehension of the words for the familiar objects used in the experiments and their success in individuating these objects in the property/kind condition. Xu and Carey (1996) examined the data from 10- and 12-month-olds in the experiments and the parental reports on the infants' comprehension of the words for the objects used in the task - “ball”, “bottle”, “cup”, “book”. They contrasted the data from these two age groups and found that only 43% of the 10-month-olds were reported to comprehend at least 2 of those words, whereas 85% of the 12-month-olds were reported to understand at least 2 of those words. Thus, there seems to be a relationship between the capacity to individuate objects in the property/kind condition and the comprehension of the first words, such as cup, bottle, ball, book.

## **An alternative**

We suggest that the capacity to represent sortals may be part of an evolutionarily ancient adaptation, which predates the emergence of a human linguistic representational capacity, in which

case the notion of object kinds would not be distinctly human. This question may be addressed by doing similar experiments with nonlinguistic creatures, such as nonhuman primates. If these creatures use sortal/kind information to establish representations of distinct objects, then language is not required for representing sortal concepts. If, on the other hand, these nonlinguistic creatures failed to establish distinct objects on the basis of sortal information, then perhaps the ability to represent sortals depends on having a language and the ability to represent sortals is uniquely human.

To test these hypotheses, Uller, Carey, Xu and Hauser (1997) conducted a study with nonhuman primates to see whether they represent sortal concepts in the absence of language. This research adapted the experimental paradigm used with infants in Xu and Carey (1996), with the goal to determine whether nonhuman primates would rely primarily on spatiotemporal cues in individuating objects, or if they would possess the ability to individuate objects using property/kind information. Several previous studies have successfully replicated other human infant preferential looking and violation of expectancy experiments on nonhuman primate populations. These studies have already established that the use of this method with nonhuman primate participants yields interpretable results (e.g., Hauser, MacNeilage & Ware, 1996; Uller, Carey & Hauser, 2001).

### **The Uller et al. (1997) study**

In adapting the infant paradigm of violation of expectancy for use with a primate species, several changes in the experimental arrangement had to be made. First, a new apparatus needed to be devised and constructed. Given the environmental conditions of the field studies, this was necessary so that the apparatus was small enough to be transported onto the island by boat, light enough to be carried along from 7 am to 5 pm during the days of testing and sturdy enough to resist different sorts of accidents, such as rain, or during transportation on boat or while being carried on the island. Second, differently from the infant studies where the stimuli are generally toys, here we used edible items (pretesting had determined that monkeys were more willing to participate in the task if the stimuli looked like food). The stimuli found most appropriate, according to the availability of the local provision, were bright orange carrots and a yellow squash that was cut into slices roughly the same volume as the carrots.

Different groups of monkeys were tested on either one of two test conditions only, a

property/kind condition and a spatiotemporal condition. In the property/kind condition, the monkeys saw two nonidentical objects come into view, and then disappear behind a screen, one at a time. Specifically, there was temporal discontinuity between the emergences of one object and the other object. In the spatiotemporal condition, the monkeys saw two nonidentical objects come into view, both at the same time, and then disappear behind a screen. Unlike the multiple presentations to infants in the Xu and Carey (1996) design, here only two pairs of familiarizations and one pair of test trials, with one pair of objects, were used. The experiments were conducted during a period of 5 days with a semi free-ranging population of rhesus monkeys (*Macaca mulatta*) living on the 15-hectare island of Cayo Santiago, Puerto Rico (see Rawlins & Kessler, 1987 for a description of the island and history of research on this population). The population consisted of approximately 900 individuals, divided roughly into five to six social groups. Cayo Santiago is inhabited only by these monkeys; there are no natural predators. Personnel of the Caribbean Primate Research Center provide food (Purina monkey chow) every day. The monkeys are extremely well habituated to the presence of human observers, and individual monkeys can be easily recognized from unique chest tattoos and ear notches. Maternal kinship, age and sex are available from a long-term data base. While attempts were made to test both males and females, the females tended to be far more easily distracted than the males. Therefore, the collected data come from adult male participants.

Participants were selected based on their age as well as the apparent favorability of experimental conditions at the time (e.g., low density of participants near focal, orientation of the participant relative to the experimenter, degree of the participant's engagement with other activities). The monkeys were tested when they were resting and either alone or in small social groups. When a participant was located, we set up the testing apparatus and the video camera and then proceeded to run the individual experimental sessions. Data from 48 adult rhesus monkeys were used in this study. One hundred and three animals were excluded from the study due to aborted or repeated sessions when the animals showed no interest in the proceedings. The most common factor causing distraction was the initiation of social interactions by nearby individuals. Sessions were thus aborted if any of the following occurred: (1) Monkeys moved or failed to look steadily towards the stage area at any point during familiarization or test trials for at least 1 s.; (2) The time elapsed between successive trials within a session exceeded 120 seconds.

The individual sessions were conducted using a white foam core box, measuring about 50



cm x 25 cm x 35 cm. The box had a platform base and a back cover, but no sides. A screen with a hidden tray attached to the back of it covered the front when it was in place. The bottom edge of the screen fit into a groove in the platform base. Two objects were used as stimuli: a bright orange carrot, measuring about 22 cm long and 3 cm in diameter in its fattest end, and a slice of bright yellow squash, roughly 20 cm long (curved) and 2 ½ cm thick. Both the carrot and the piece of squash were substituted by another identical carrot and another identical piece of squash on the third day of experimentation on the island. The original objects became rotten due to heat, humidity and manipulation.

Forty-eight monkeys were assigned to each of the two conditions, the Property/Kind condition (PK) or the Spatiotemporal condition (ST), both modeled on Xu and Carey (1996) with a few differences. First, there was no baseline condition. Second, the number of times the objects emerged from behind the screen was less in both introductory and familiarization trials. Third, the objects used were unfamiliar to the monkeys. Fourth, the pair of objects utilized in all of the test trials was the same. Each condition consisted of two sections: pretest familiarization trials and test trials. The pretest trials served to familiarize the monkeys with the objects and the apparatus. The monkeys learned that there were objects behind the screen. These pretest trials did not provide any information as to the number of objects that would be present when the screen was removed. The two test trials involved one of the two outcomes, namely, the expected outcome (two objects) or the unexpected outcome (one object).

Each participant was shown a series of four pretest familiarization emergences, one test trial, two additional familiarization emergences, and the second test trial.

A familiarization trial consisted of the experimenter showing the monkey one object emerging from behind the screen to the left or right of the experimenter, who sat directly behind the testing box, aligned with its center. Once the monkey had looked at the object for at least 2 s, it was brought back behind the screen. Immediately thereafter, the second object was brought out from the other side of the screen. After the monkey had looked at this object for at least 2 s, it was again brought back behind the screen. In the third familiarization trial, the experimenter brought the first object out from behind the side of the screen where she had first brought it out, laid it on the ground, and left it there for the monkey to look at it for approximately 5 s before bringing it back behind the screen. In the fourth familiarization trial, the experimenter brought the second object out from behind the side of the screen where she had previously brought it out, placed it on



the ground, and let the monkey look at it for 5 seconds before bringing it back behind the screen.

Immediately following the first four familiarization trials, the first test trial was presented. The screen was lifted up and placed behind the box, revealing either one object (the unexpected event) or two objects (the expected event). To produce the unexpected event, the experimenter placed one of the objects in the hidden tray attached to the back of the screen during the third or fourth familiarization trial. The trial in which this occurred depended on whether the one-object outcome was the last object seen or the one before last seen by the monkey. The participant was allowed to look at the display for a maximum of 10 s, after which the experimenter replaced the screen in front of the box.

Participants were tested individually. When a participant was located, we set up the apparatus from a distance of approximately 1.5 - 2 meters (depending on the terrain) between participant and testing box/experimenter and then proceeded to run the individual experimental sessions. The video camera, testing apparatus and the experimental participant were lined up in a straight line so that the video record provided an unambiguous, head on view of the participants' eyes. In video taping a trial, we attempted to fill as much of the recorded image with the participant's head as possible. The start of each familiarization and test trial was announced by the experimenter, for the benefit of the offline observers as well as of the observer videotaping the session. The experimental sessions were recorded on a video camera model Panasonic IQ404. Analog records of the videotapes were digitized onto a Macintosh Quadra 950 using the Radius VideoVision board. Frame-by-frame quantification (30 frames/second) of the total amount of time looking at the display (out of 10 seconds) was then scored using the Adobe Premiere (version 4.2) software. The experimenter and the two live observers separately coded the videotapes of the experimental sessions. They were blind to the experimental conditions presented. A correlation analysis on the looking times of the first four monkeys indicated that inter-observer reliability was 99.5%.

Statistical analyses of the data revealed surprising results. In the PK condition, 15 out of 24 monkeys looked longer when there was only one object in the box than when there were two objects, while in ST condition, 21 out of 24 monkeys looked longer at the outcome of one object than at the outcome of two objects. In addition, and most importantly, monkeys in both conditions looked significantly longer at the unexpected outcome of one object than at the expected outcome of two objects.

## Discussion

The main result of Uller et al. (1997) provides significant implications for the strong versions of a Language of Thought hypothesis. The animals seemed sensitive to the presence of two distinct individuals, or the numerical identity of the set of objects presented to them regardless of whether they were seen together or one-at-a-time, namely, whether the animals were given property/kind information or spatiotemporal information about the existence of object A and object B. Minimally, one has to concede that the evidence provided by Uller et al.'s (1997) results is consistent with the possibility that the monkeys use property/kind differences between objects to individuate such objects as sortals in the absence of language. If this finding holds up, we would have to conclude that the ability to represent sortals (kinds) is not uniquely human but rather an evolutionarily ancient adaptation, and that it is unlikely that infants would need language to construct sortal concepts.

A word of caution. It could have been possible for the monkeys to use solely property differences, as opposed to kind differences, to individuate the objects. For example, it is possible that they could have represented the carrot as a bright, orange, and carrot-shaped object and the piece of squash as a yellow, slimy, and squash-shaped object, and they know that things that are bright and orange do not turn into things that are yellow and slimy, therefore there must be two things behind the screen. So the monkeys could have succeeded at this task without the need to represent the two objects (sortals) as instances of the kinds carrot and squash. It could also be that the monkeys used a mass quantificational system that allowed them to represent 'amount of stuff' without necessarily representing the carrot and the squash as distinct individuals or sortals. For example, the monkeys could have encoded what was behind the screen as 'a certain amount of carrot stuff' and 'a certain amount of squash stuff', and when the screen was removed, the monkeys expected to see some carrot stuff and some squash stuff without representing each as an individual or sortal.

Overall, we take the results of Uller et al. (1997) to represent evidence in support to the idea that there is thought without language. A nonlinguistic animal was able to represent distinct individuals, individual objects (sortals), and tell numerical identity, without the need to see the objects at the same time, that is, in the absence of spatiotemporal information, solely in the present of property/kind information. While it is still open as an alternative to entertain the possibility of

different systems – e.g., perhaps humans represent the world differently to nonhuman animals, specifically as a function of a language-based cognitive architecture -, or the animals used property only to individuate the objects – regardless, the questions addressed by this research are interesting, the evidence is relevant to questions of linguistic architecture, and of what makes us human. It remains to be seen where future research in this area will go and what exciting findings it will bring.

## REFERENCES

- Fodor, J. (1975). *The Language of Thought*. Cambridge, MA: MIT Press.
- Hauser, M., MacNeilage, P. & Ware, M. (1996). Representations in primates: numerical or perceptual? *Proceedings of the National Academy of Sciences*, 93, 1514-1517.
- Quine, W. V. O. (1960). *Word and Object*. Cambridge, MA: MIT Press.
- Rawlins, R. & Kessler, M. (1987). *The Cayo Santiago Macaques*. New York, NY: SUNY Press.
- Simon, T., Hespos, S. & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development*, 10, 253-269.
- Spelke, E. & Kestenbaum, R. (1986). Les origines du concept d' objet. *Psychologie Française*, 31, 67-72.
- Spelke, E., Kestenbaum, R., Simons, D. J. & Wein, D. (1995). Spatiotemporal continuity, smoothness of motion and object identity in infancy. *British Journal of Developmental Psychology*, 13, 113-142.
- Uller, C., Hauser, M. & Carey, S. (2001). Spontaneous representation of number in cotton-top tamarins. *Journal of Comparative Psychology*, 115, 1-10.
- Uller, C., Carey, S., Hauser, M., Xu, F. (1997). Is language needed for constructing sortal concepts? A study with nonhuman primates. In E. Hughes et al. (Eds). *Boston University Conference on Language Development Proceedings*, 21, 665-677. Somerville, MA: Cascadilla Press.
- Uller, C., Carey, S., Huntley-Fenner, G. & Klatt, L. (1999). What representations might underlie infant numerical knowledge.
- Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 358, 749-750.
- Xu, F. & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, 30, 111-153.