Cognition 170 (2018) 9–24

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

Competing perspectives on frames of reference in language and thought

Peggy Li[a], Linda Abarbanell[b]

[a] Harvard University, United States
[b] San Diego State University, United States

ARTICLE INFO

Keywords:
Linguistic relativity
Frames of reference
Tseltal Mayan
Hai||om

ABSTRACT

A study found that Dutch-speaking children who prefer an egocentric (left/right) reference frame when describing spatial relationships, and Hai||om-speaking children who use a geocentric (north/south) frame had difficulty recreating small-scale spatial arrays using their language-incongruent system (Haun, Rapold, Janzen, & Levinson, 2011). In five experiments, we reconciled these results with another study showing that English (egocentric) and Tseltal Mayan (geocentric) speakers can flexibly use both systems (Abarbanell, 2010; Li, Abarbanell, Gleitman, & Papafragou, 2011). In replicating and extending Haun et al. (Experiment 1), English- but not Tseltal-speaking children could use their language-incongruent system when the instructions used their non-preferred frame of reference. Perseveration due to task order may explain the discrepancies between present English- and previous Dutch-speaking children, while not understanding task instructions using left/right language may explain why present Tseltal- and previous Hai||om-speaking children had difficulty with their language-incongruent systems. In support, Tseltal-speaking children could use an egocentric system when the instructions were conveyed without left/right language (Experiments 2–4), and many did not know left/right language (Experiment 5). These findings help reconcile seemingly conflicting sets of results and suggest that task constraints, rather than language, determine which system is easier to use (Experiment 2 vs. 3).

1. Introduction

The possibility that speakers may experience and even perceive the world along language-specific lines is central to debates regarding the plasticity of human cognition. How much of our cognitive make-up is determined by our biology and how much is shaped over ontogenetic development by cultural systems such as those that may be expressed and acquired through language? This debate is not only central to the cognitive sciences, but has perennially fascinated the public with the possibility that core aspects of our humanity may radically differ depending on the environment in which one is raised. As a New York Times article explains, if your language habitually "forces" you to attend to certain experiences, it is only "natural" that one develops related habits of mind that shape one's "experiences, perceptions, associations, feelings, memories and orientations in the world" (Deutscher, 2010, August, 29, p. MM42).

Natural, perhaps, but how strong is the empirical evidence supporting such intuitions? The present paper revisits evidence from one of the most often cited cases in support of linguistic relativity concerning the language of space; specifically, the coordinate systems or frames of reference (FoR) speakers use to talk about locations and directions. While several studies have yielded seemingly conflicting results (Haun, Rapold, Janzen, & Levinson, 2011; Li et al., 2011), we will endeavor to make sense of the conflict with a set of new studies. To do so, we will first present an overview of the arguments for linguistic relativity in this domain, highlighting how different sets of data have been used to argue for two competing accounts: a *linguistic enculturation* account arguing for an effect of language on nonlinguistic representations via practice effects (e.g., Haun et al., 2011; Levinson, 2003; Levinson et al., 2002; Slobin, 2003) and a *pragmatic inference* account arguing for a weaker effect of language on how speakers' interpret certain ambiguous or underspecified tasks (Gleitman & Papafragou, 2013; Li & Gleitman, 2002; Li et al., 2011). We then present the new set of studies aimed at reconciling discrepancies between previous reports. Finally, we end with some thoughts about cognitive diversity and flexibility in this domain and speculate on where language may and may not be predicted to have an effect.

Spatial frames of reference has been a fruitful area to study linguistic relativity. At first glance, we might expect something as basic to our survival as spatial localization to have a strong biological basis and therefore to be encoded in more or less the same way across languages (e.g., Landau & Jackendoff, 1993; Talmy, 1983). It turns out, however, that evolution has endowed us with multiple solutions to navigate the world, leaving language communities to pick and choose the coordinate system(s) speakers primarily use to encode directions and locations of objects. These systems, known as linguistic frames of reference, are

E-mail addresses: pegs@wjh.harvard.edu (P. Li), labarbanell@mail.sdsu.edu (L. Abarbanell).

http://dx.doi.org/10.1016/j.cognition.2017.09.005
Received 12 November 2015; Received in revised form 8 September 2017; Accepted 8 September 2017

used to specify the direction of an entity, generally called the 'figure', with respect to another entity, or 'ground' (Talmy, 1983). In English and other standard European languages, speakers primarily use the perspective of a viewer to establish a directional system (e.g., left/right). Environment-derived terms like 'north' and 'south' are primarily restricted to large-scale or map space. In some other languages, however, like Tseltal Mayan (Chiapas, Mexico) speakers use fixed aspects of their environment to describe directions and locations even when describing items in small-scale space such as the arrangement of items on a table top. Such cross-linguistic differences have been argued to affect how speakers interpret, store and retrieve spatial information across modalities, resulting in greater facility using language-congruent strategies and a dispreference for language-incongruent ones (e.g., Haun, Rapold, Call, Jenzen, & Levinson, 2006; Haun et al., 2011; Levinson, 1996, 2003; Pederson et al., 1998). As Haun et al. (2011) put it, "if a certain strategy is culturally required (for example through language use), and thus heavily practiced leading to a preferred default cognitive strategy, there is reason to expect performance using that strategy to be better" (p. 73).

Testing these predictions, however, is complicated by the fact that there are many ways to partition and classify these variations in FoRs (Shusterman & Li, 2016a, 2016b). A natural division, based on the physical world in which we live, is between *object-centered* FoRs and *geocentric* FoRs (Gallistel, 2002; Wang, 2012). That is, between FoRs derived from objects or entities that move around on earth versus FoRs that are anchored to and invariant of earth. On the other hand, cognitive psychologists generally distinguish between *egocentric* and *allocentric* frames. The singling out of the ego as a special kind of object makes sense given that it is expected to have a privileged role in cognition: it is the perspective from which we take in information and plan our movement in the world (Gallistel, 1990; Wang, 2012; Wang & Spelke, 2002). Allocentric frames includes everything else, from the facets of other (non-egocentric) movable objects, including other animate beings, to fixed features of the environment. Within this broad category, researchers consider how the reliability of an entity and its features affects its likelihood of being chosen as a reference frame (Gallistel, 1990). In yet a third way of classifying FoRs, researchers documenting spatial language generally distinguish three different reference frames, based not only on the perspective, but on the extent to which the perspective is projected. Using Levinson's (1996, 2003) terminology, an *intrinsic* system is a binary relation between a figure and ground, with the figure located at some feature of the ground (e.g., "the cup is at the mouth of the pitcher"). In contrast, a *relative* system is a ternary relation between figure, ground, and some independent viewer's perspective that is projected onto the ground (e.g., "the cup is to the left of the pitcher"). Lastly, in an *absolute* system, the figure is located with respect to the ground by projecting an environmentally fixed asymmetry (e.g., "the cup is to the north of the pitcher").

While there has been debate in the literature on exactly which classification best maps onto testable predictions about spatial cognition (see Shusterman & Li, 2016a, 2016b for how the classifications relate to each other), most of the experimental evidence in this domain has focused on speakers of languages that habitually take an egocentric-relative versus a geocentric-absolute perspective when describing figure-ground relations, with researchers making use of the properties of each system under rotation. For example, in the animals-in-a-row task, participants are shown a row or an arrangement of toy animals, and then asked to recreate the "same" array after turning 90° or 180° to face a second table (see Fig. 1 for an example of the table set-up in the present study). In a large and influential body of work, studies with over 20 languages revealed a robust and striking correlation: speakers of languages like English that habitually use an egocentric system, often rotated the animals along with their bodies (Fig. 1a), while speakers of languages like Tseltal often maintained the orientation of the animals with respect to the environment (Fig. 1b). These results have been used to argue that the habitual use of one linguistic system versus another

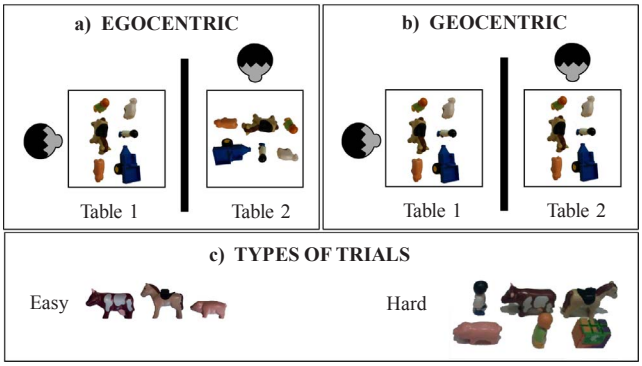

Fig. 1. Experimental set-up of animals-in-a-row task, where participants remember an array of animals at Table 1 and have to recreate it at Table 2. (a) Shows the egocentric solution and (b) the geocentric solution. (c) Shows sample arrays from the present studies, consisting of easy (3 figurines) or hard (6 figurines) arrays that participants have to remember.

yields practice effects, making speakers better at reasoning in the system congruent with their language (Brown & Levinson, 1993b; Levinson, 1996, 2003; Majid, Bowerman, Kita, Haun, & Levinson, 2004; Pederson et al., 1998), sometimes at the expense of the system that is less frequently used in their language (Haun et al., 2006, 2011).

Not all researchers assume these results by themselves are conclusive in demonstrating that language use shapes the availability of frames of reference in everyday spatial reasoning (Li & Gleitman, 2002; Li et al., 2011; Newcombe & Huttenlocher, 2000; Pinker, 2007). They take issue in particular with the open-ended structure of the tasks. Using Brown & Levinson, 1993a, 1993b's "chip task" as an example, Li et al. (2011) explained how pragmatic inference is implicitly involved when deciding what is meant by the command to find or make the "same" array. In the chips task, participants were shown a card with two colored dots (e.g., a red dot left/south of a green dot) at the first table. After turning to face a different direction and a second table, participants saw four copies of the original card, each in a different orientation. They were told to "find the 'same' card." To find the same card, they had to have reasoned that "same" did not mean the card they saw at the first table, nor did it mean all four copies at the second table. Instead, participants had to have reasoned that "same" must refer to a unique card in a specific orientation. But which orientation should one pick? Here, how one's linguistic community generally talks about and responds to queries about locations and directions, rather than one's own cognitive preference or predilection to reason in a particular frame, may serve as a probabilistic basis for inferring which orientation the experimenter meant by the "same." Conceivably, participants' expectations of, or adherence to linguistic conventions could similarly be driving the response pattern for other such open-ended "find the same" or "make the same" tasks or tasks that require guessing the intent of the experimenter (e.g., Haun et al., 2006; Majid et al., 2004).

There are therefore two competing accounts of the covariation between linguistic preferences in frame of reference use and task performance on these open-ended tasks. According to the linguistic enculturation account, the habitual use of one system versus another in one's language affects the availability of and ease of using the corresponding nonlinguistic spatial representations. As Haun et al. (2011) put it, "if a certain strategy is culturally required (for example through language use), and thus heavily practiced…, there is reason to expect performance using that strategy to be better (Haun et al., 2006)." In contrast, under the pragmatic inference account, language may exert a weaker effect by influencing how speakers interpret ambiguous or open-ended representations, without necessarily altering the availability of each system or the ease with which they are used to solve different types of spatial problems in daily life.

To adjudicate between these two accounts, Li et al. (2011) and Abarbanell (2010) tested English- and Tseltal-speaking adults and