Perceiving and remembering events cross-linguistically: Evidence from dual-task paradigms

John C. Trueswell [a,*], Anna Papafragou [b]

[a] Department of Psychology, University of Pennsylvania, 3401 Walnut Street C-Wing, Philadelphia, PA 19104, United States
[b] University of Delaware, United States

ARTICLE INFO

ABSTRACT

What role does language play during attention allocation in perceiving and remembering events? We recorded adults' eye movements as they studied animated motion events for a later recognition task. We compared native speakers of two languages that use different means of expressing motion (Greek and English). In Experiment 1, eye movements revealed that, when event encoding was made difficult by requiring a concurrent task that did not involve language (tapping), participants spent extra time studying what their language treats as the details of the event. This 'linguistic encoding' effect was eliminated both when event encoding was made easier (no concurrent task) and when the concurrent task required the use of language (counting aloud). In Experiment 2, under conditions of a delayed concurrent task of counting aloud, participants used language covertly just prior to engaging in the additional task. Together, the results indicate that language can be optionally recruited for encoding events, especially under conditions of high cognitive load. Yet, these effects are malleable and flexible and do not appear to shape core biases in event perception and memory.

© 2010 Elsevier Inc. All rights reserved.

When we inspect a picture, or almost any visual scene, our eyes rapidly dart from person to person, place to place, object to object. Research into understanding the dynamics of scene perception indicates that these eye movements, although partially driven by bottom-up visual factors, reflect goal-directed categorization processes; the entities, events and states of affairs are placed into task-relevant categories, designed to achieve the person's immediate and longer-term goals (Yarbus, 1967; cf. Griffin & Bock, 2000; Henderson, 2003, 2007; Mennie, Hayhoe, & Sullivan, 2007; Pelz & Canosa, 2001). The present paper follows this line of research, by asking: to what extent are task-relevant categorization processes that occur during scene perception necessarily or predominantly linguistic in nature? And if language is a typical currency for representing the world, does this fact influence what we attend to even

when we are not using linguistic representations as a means of encoding?

Although there have been explorations into the role language plays in spatial reasoning and problem solving (e.g., Levinson, 2003; Li & Gleitman, 2002) and the memory and categorization of scenes/events (e.g., Gennari, Sloman, Malt, & Fitch, 2002; Papafragou, Massey, & Gleitman, 2002; Papafragou & Selimis, 2010), much less is known about whether and/or how linguistic knowledge plays a role in the dynamics of scene perception itself and the categorization processes that comprise it. For example, some potentially positive, albeit indirect, evidence that language does play a role in scene perception can be found in the early work of Loftus (1972), as well as from more recent work reported by Antes and Kristjanson (1993). These particular lines of research were not designed to test the relationship between language and scene perception (they instead explore the relationship between scene encoding and scene memory generally), yet both papers included experiments that could be interpreted as having some

* Corresponding author.
  E-mail address: trueswel@psych.upenn.edu (J.C. Trueswell).

bearing on the issue of language. In these experiments, eye movements were recorded while participants studied static images of simple scenes, such as a farm, an urban neighborhood, etc. Participants were asked to view these images for a short time in preparation for a memory test, while either engaged in a concurrent task that involves language (e.g., counting backward by three) or not engaged in a concurrent task. Both Loftus (1972) and Antes and Kristjanson (1993) report that participants' eye movements and picture memory were negatively affected by this linguistic interference. Fixations became longer and, although insufficient details were provided about the specific gaze patterns of participants, the authors of both papers suggested that viewers spent less time inspecting important details of the images. Nevertheless, even when controlling for how often people looked at objects in these images, memory for the objects was impaired by linguistic interference, consistent with non-eyetracking studies of picture memory that also used interference tasks involving language during or after picture encoding (e.g., Allport, Antonis, & Reynolds, 1972; Antes & Kristjanson, 1993; Rowe & Rogers, 1975; Wolfe, Horowitz, & Michod, 2007). Such findings are only suggestive of a role for language however, largely because the studies were not designed to test the hypothesis that language is involved in perceiving and remembering scenes (e.g., direct comparisons with non-Linguistic interference tasks were not done).

The present study addresses the role of language in scene perception more directly, asking how and when language is recruited when people view depictions of simple events. These questions are particularly relevant in the context of a resurgence of interest in the relationship between language and perception/cognition (see Boroditsky, 2001; Gumperz & Levinson, 1996; Levinson, 2003; Whorf, 1956; cf. also papers in Bowerman & Levinson, 2001; Gentner & Goldin-Meadow, 2003). Until quite recently, most psycholinguists and cognitive psychologists believed, as part of their background assumptions for their theories of perception and language use, that linguistic encoding of the world was an optional process done either for purposes of communication (e.g., see Levelt, 1989) or as an aid for short- and long-term memory (e.g., see Miller, 1956; Paivio, 1971). In most other situations, lexical and syntactic characterizations of the world were not routinely deployed during perception, as these characterizations were believed to have little or no bearing on our physical interactions with the world.

Recently however, a number of researchers have questioned these background assumptions. For instance, it has been claimed that linguistic representations may actually be used as a matter of course, perhaps even automatically, in a range of non-linguistic perceptual tasks, including on-line perceptual categorization of the features of objects and objects themselves (e.g., Gilbert, Regier, Kay, & Ivry, 2008; Winawer, Witthoft, Frank, Wu, & Boroditsky, 2007; cf. Roberson, Davidoff, & Braisby, 1999). In particular, when discussing color categorization, Winawer and colleagues state that "language-specific categorical representations play an online role in simple perceptual tasks that one would tend to think of as being primarily sensory" and that these linguistic representations are "brought online spon-taneously during even rather simple perceptual discriminations" (Winawer, Witthoft, Frank, Wu, & Boroditsky, 2007, p. 7784). Similar conclusions are drawn by Gilbert et al. (2008) who extend these conclusions beyond color categories to object categories such as dog and cat. They argue that the use of lexical information during object categorization is not a strategic process but rather "language affects discrimination on-line through the activation of lexical codes." Still others have argued that specific syntactic mechanisms are necessary for integrating feature/object information with location information (e.g., Hermer-Vazquez, Spelke, & Katsnelson, 1999). The processes of interest to these researchers (i.e., categorization of features/objects plus the integration with location information) are computations believed to be central to scene perception (e.g., Henderson, 2003). Thus these accounts predict that scene perception should routinely include the computation of linguistic information either because of automatic processes (such as spreading activation) or because the integrative encoding of the world requires the use of linguistic mechanisms.

Others have taken a more nuanced perspective on these issues. For instance, Slobin (2003, 2006) maintains the standard assumption about linguistic encoding being optional. However, he argues that it is a mistake to think that linguistic encoding of the world during perception is a rare or exceptional case for humans in their everyday lives. For instance, he notes that learning a first language most likely requires children to attend to conceptual contrasts more commonly encoded in their native tongue (e.g., language-specific contrasts relevant for encoding events) and not the contrasts less commonly encoded (Slobin, 2003). Slobin also notes that humans often anticipate a need to describe events later to other individuals, perhaps producing supplementary linguistic encoding of the events even when not currently talking about them. Slobin proposes that this pervasiveness of 'thinking for speaking' may in fact lead individuals to be especially attuned to conceptual contrasts made in their language, modulating attention to aspects of the world that are relevant to their particular language, even when linguistic encoding of the world is not occurring (Slobin, 2006). This latter conclusion suggests that language-specific effects on attention can have stable global consequences on our perceptual interrogation of the world.

These various positions have not always been clearly separated in the literature (nor are they easily separable in practice). Empirical studies have only just begun to investigate the role of language on event perception systematically (see the next section) but, so far, such studies have not attempted to distinguish between these theoretical options about the robustness of the role of language. In the present work, we address this gap by exploring the interface between language and perception/memory under different interference conditions and among speakers of different languages.

We take as our test bed the perception and memory of simple motion events (e.g., a man skating over to a hockey net) for which there are strong cross-linguistic differences in how they are routinely described (Talmy, 1975, and next section). Our own past work shows that differences also