12   The Representation of Events in Language and Cognition

Anna Papafragou

12.1   Introduction

The world around us is full of constantly unfolding dynamic events: as I look out the window from my study, a car is honking, birds are flying, two friends are crossing the street, and the sun is moving behind the clouds. To successfully interact in the world, we must track and interpret a rich array of events to relate them to others. How do we accomplish this?

According to most theories, the way we talk about the world is limited by basic, probably universal, constraints on how we conceptualize space, objects, and events. For instance, models of language production assume that the act of speaking begins with event conceptualization (the preverbal apprehension of the broad details of an event), followed by further processes of information selection and linguistic formulation that culminate in speech execution (Levelt 1989; Bock, Irwin, and Davidson 2004; cf. also Lashley 1951; Paul 1970; Wundt 1970). Similarly, models of language acquisition generally assume that learners are equipped with a set of concepts for representing events and their components (even though the size of that set is under dispute), such that acquiring language includes mapping incoming speech stimuli onto this set of concepts (Gleitman 1990; Jackendoff 1996; Miller and Johnson-Laird 1976; Pinker 1989).

This view on how event representations make contact with language faces two challenges. First, the underlying perceptual and conceptual representations of events are elusive. Existing work on event recognition has addressed the perception of biological motion (Blake and Shiffrar 2007; Giese and Poggio 2003; Johansson 1973; Lange and Lappe 2006), the segmentation of events from dynamic displays (Newtson 1973, 1976; Zacks, Tversky, and Iyer 2001), and more recently the neural bases of human action representations (Kable and Chatterjee 2006; Tranel et al. 2003). However, little work has been done on the conceptual structure of events per se, and on the relationship between the linguistic and nonlinguistic representation of event components (see Shipley and Zacks 2008 and the infant studies reviewed in Baillargeon et al. 2011;

Göksun, Hirsh-Pasek and Golinkoff 2010; and Spelke, Phillips, and Woodward 1995 for relevant evidence). Some studies have suggested that information about individual event components (e.g., the person being affected by an action) or relationships between event components that determine whether an event is coherent can be extracted rapidly by human viewers (see, respectively, Griffin and Bock 2000; Dobel et al. 2007; cf. also Dobel et al. 2011; Webb, Knott, and MacAskill 2010). Nevertheless, these studies as a whole have focused on a limited number of events and event components and thus leave the basic building blocks and processes underlying event representation still underspecified.

Furthermore, there have been few explicit attempts to connect the process of building event representations to mechanisms of language production. In the first study of this kind, Griffin and Bock (2000) recorded speakers' direction of gaze as they visually inspected and described static line drawings of simple actions (e.g., a woman shooting a man). Analysis of the eye movements in relation to people's linguistic choices led to the conclusion that there exists an initial rapid event-apprehension stage that is temporally dissociable from any sentence-generation stage. But further eye-tracking studies using picture description tasks have shown that event apprehension and linguistic formulation overlap temporally to a greater extent than suggested by this first study (see Gleitman et al. 2007).

A second challenge for theories attempting to account for how event cognition interfaces with language is that, across natural languages, the lexical-structural representation of events varies considerably. For instance, languages have different means for encoding space, motion, number, objects, and informational access to events (see Bowerman and Levinson 2001; Gentner and Goldin-Meadow 2003; Gleitman and Papafragou, in press, for reviews, and later sections of this chapter for examples). These crosslinguistic differences raise the question whether the underlying perceptual and conceptual event representations might also vary in the minds of speakers of different languages. This possibility has been famously espoused in the past by Benjamin Lee Whorf (Whorf 1956) and has been recently revived in several neo-Whorfian theories of how perception and cognition are connected to language (e.g., Boroditsky 2001; Levinson 2003). Proponents of these theories argue that "habitual or obligatory categorical distinctions made in one's language result in language-specific categorical distortions in objective perceptual tasks," with language-specific categories being used "on-line" during perception (Winawer et al. 2007, 7783). Similarly, other researchers propose that "experience, including experience with language, can influence the perceptual system such that it is more or less attuned to particular features in the environment" (Majid et al. 2004, 113), perhaps through the selective direction of attention (Smith et al. 2002). If language can reorganize the perceptual and conceptual representation of events, then it is no longer true that (universal) event categories provide the building blocks for event encoding cross-linguistically: the causal relation can also

flow in the opposite direction, with different languages having stable and permanent restructuring effects on the event concepts held by their speakers (Levinson 1996, 2003).

In this chapter, I present the results from a multipronged research collaboration between my lab and John Trueswell's lab that combines on-line, developmental, and crosslinguistic methods to address these issues. The series of studies I discuss ask two types of questions. First, how do people perceive events? Is there concrete evidence for a link between event perception and the lexical-structural representation of events in language? Second, do crosslinguistic differences in event encoding affect event apprehension? In particular, do speakers of different languages process events in distinct ways? I take up each of these types of question in turn, introduce key findings from our studies, and sketch some general implications for the interface between event representation and language.

12.2   How Are Events Perceived?

Studies of event perception need to begin with the question of what the constituents of events might be, and how these constituents contribute to event individuation and identification. Consider the simple event of a man chasing a toddler. A straightforward way of characterizing this event includes distinguishing between the role of the man, who is the causer of the action, and that of the toddler, who is the entity directly affected by that action. Inspired by linguistic analyses of events in sentences (e.g., Fillmore 1968; Chomsky 1981), one could call the first entity the *agent* of the event and the second entity the *patient*. In language, agents and patients are among several possible *thematic roles*, or relations that a sentential constituent (typically a noun phrase such as *a man* or *a toddler*) has with a main verb (Chomsky 1981; Jackendoff 1990)—see the sentence in (1a) and its representation in (1b):

(1a) A man is chasing a toddler.

(1b) [A man]$_\text{AGENT}$ is chasing [a toddler]$_\text{PATIENT}$.

It seems reasonable to assume that agents, patients, and related roles have cognitive corollaries that can be used to capture the who-did-what-to-whom structure of event representations. The idea that there is a tight relation between the cognitive and linguistic representation of events and event components is not new. Several commentators have proposed that linguistic roles such as agents and patients map relatively directly onto underlying conceptual event roles, at least in the most typical cases (see Baker 1997; Dowty 1991; Jackendoff 1990; Levin and Pinker 1991). Others have proposed that linguistic events could be individuated based on thematic roles, such that the representation of an event may have only one agent, one patient, and so forth (Carlson 1998); similar ideas can be used to characterize nonlinguistic event

El

individuation. However, the extent to which nonlinguistic event construal is truly sensitive to components such as agents and patients as well as the processes underlying the extraction of such components remain unspecified.

In one of the first studies to investigate these issues, Hafri, Papafragou, and Trueswell (2013) asked whether the conceptual roles of agent and patient (and the events they comprise) could be rapidly extracted from visual stimuli.[1] The researchers displayed naturalistic photographs of a wide range of two-participant events (e.g., a girl punching a boy) for either 37 ms or 73 ms followed by a mask. People were then asked to identify the category of the event ("Did you see 'punching'?"), the role of agent ("Is the girl performing the action?"), the role of patient ("Is the boy being acted upon?"), or a combination of event category and roles ("The girl is punching the boy, true or false?"). It was found that people successfully recognized both event categories and event roles (and combinations of the two) even from the briefest visual displays. Further experiments revealed that the rapid recognition of event roles made use of certain physical features (e.g., outstretched extremities) that can be linked to features such as causality that typically characterize (linguistic) agenthood (e.g., Dowty 1991). Specifically, under short display durations, subjects recognized the roles of event participants less accurately when patients possessed agent-like features such as outstretched extremities (this difference disappeared with longer display durations).

These findings show that people are able to extract event role and event type information on the basis of briefly presented visual stimuli. They also show that the perceptual-conceptual representation of events is sensitive to roles such as agents and patients that are relevant for the description of events in language. Finally, the fact that event roles depend for rapid identification on certain physical features that are linked to their causal role in the event is consistent with proposals to treat linguistic thematic roles more generally as "convenient mnemonics" for prominent structural configurations of conceptual structure (Jackendoff 1990, 47).

Can these conclusions be extended to more complex events containing more participants? Consider causative events such as those in figure 12.1, in which a human agent uses an instrument to move an object toward a goal. The event in figure 12.1 would typically be described with the sentence in (2a), which has the underlying thematic-role structure in (2b):

(2a) A woman is hitting a ball into a basket with a tennis racquet.

(2b) [A woman]$_{AGENT}$ is hitting [a ball]$_{PATIENT}$ [into a basket]$_{GOAL}$ [with a tennis racquet]$_{INSTRUMENT}$.

Within linguistic theory, the thematic roles involved in (2b) would be ranked differently with respect to each other, with agents considered more prominent than patients,

---

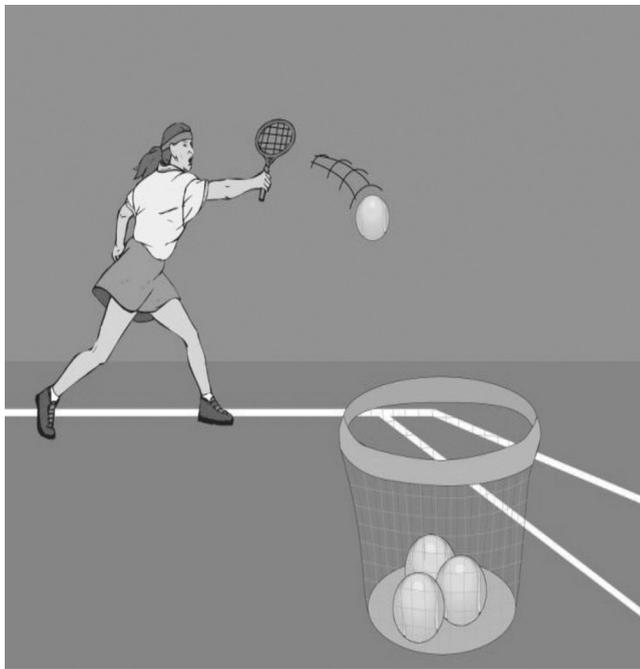1. All experiments in this section were conducted with English-speaking adults.

Figure 12.1
Sample stimulus from the study by Wilson et al. (2011).

and patients considered more prominent than goals; instruments would be considered least prominent. This ranking of thematic roles, otherwise known as *thematic hierarchy* (Baker 1997; Jackendoff 1990), determines the linguistic behavior of different thematic roles, affecting, for example, how each role gets mapped onto specific syntactic positions in the sentence (e.g., how agents get mapped onto sentential subjects, etc.). For present purposes, it is important to point out that the thematic hierarchy has been claimed to be based on prominence relations between event roles in the nonlinguistic representation of events (see, e.g., Levin and Rappaport Hovav 2005; Baker 1997). Such claims have been difficult to evaluate, however, because of lack of evidence about how different event components are represented nonlinguistically.

In a series of experiments, we (Wilson, Papafragou, and Trueswell 2011; Wilson et al. 2011) tested how people process causative events with the goal of discovering whether there is an internal hierarchy of event components and whether that hierarchy is shared between language and perception/cognition. We used pictures of events with clearly visually defined, distinct areas corresponding to the event roles of agent, patient, goal, and instrument. (Figure 12.1 depicts an actual stimulus item.) One group of people completed a linguistic task, in which they viewed the events and

were asked to describe them. A different group of people participated in a nonlinguistic task. These participants were told that they would see pictures depicting an action or event and were asked to identify certain event components by quickly looking at them and then pressing a computer key (see Griffin and Bock 2000 and especially Hafri, Papafragou, and Trueswell 2011 for similar methods). Depending on the condition they were assigned to, participants in the nonlinguistic task were told to look as quickly as possible at "the person or animal who was performing the action" (agent condition), "the object directly affected by the action" (patient condition), "the goal or destination of the action" (goal condition), or "the tool or body part used to make the action" (instrument condition). Of interest was whether the frequency with which people mentioned components of events such as agents, patients, goals, and instruments in the linguistic task would conform to the thematic hierarchy, with some components being mentioned more reliably than others. Also of interest was whether the speed of identification of individual event components in the nonlinguistic task would reveal a similar asymmetry.[2]

This dual expectation was confirmed in our data. The linguistic task revealed a difference among event components that was consistent with the thematic hierarchy: agents were mentioned most often in people's descriptions, followed by patients and goals; instruments were mentioned least often. This difference was mirrored in the eye-gaze data of the nonlinguistic task: finding an agent in a causative event (as measured by participants' eye movements to the corresponding region) occurred more quickly than finding a patient, goal, or instrument, and instruments were identified more slowly than the other event components. Further analyses ensured that this difference in speed of identification could not be due to a difference in size between the physical entities representing the different event components in our stimuli but could be confidently attributed to the role each entity played in the event.

Summarizing, the set of studies just reviewed breaks new ground by introducing a novel approach to the study of event perception and its relation to language. A major conclusion from these studies is that event perception delivers representations with internal structure (including event roles such as agents, patients, goals, and instruments). Furthermore, this structure is homologous to linguistic event structure, as evidenced by the fact that the salience of roles in complex events is similar across perception/cognition and language. Thus our results offer concrete support for the much-cited but little-tested view that the language of events builds on underlying

2. To maintain the naturalness of the events, the agents in our studies were always animate entities, and the other three components were inanimate entities. Even though we mention agents for completeness, our main focus was the asymmetry between the three components that did not differ in animacy.

El

nonlinguistic event concepts: as our data show, linguistic agents, patients, and so on are relational notions at the interface with conceptual structure (cf. also Jackendoff 1990; Levin and Rappaport Hovav 2005; Pinker 1989).

12.3   Do Crosslinguistic Differences in Event Encoding Affect Event Cognition?

The studies reviewed so far have dealt with event representations and the way these representations map broadly onto linguistic strings. I now turn to a separate issue: Is it possible that one's representation of events in the world depends on the language one uses to describe such events? And if so, is the representation of events different across speakers of different languages?

To address this issue, my collaborators and I focused on motion events, in which an agent moves spontaneously in a certain manner (e.g., rolling, crawling), following a specific trajectory, or path (e.g., into a cave, through a forest; Miller and Johnson-Laird 1976; Talmy 1985). Motion is an ideal domain for studying aspects of event cognition crosslinguistically. First, the representation of motion and space is a fundamental human cognitive ability (Carlson and van der Zee 2005; Hayward and Tarr 1995; Newcombe and Huttenlocher 2003; Stiles-Davis, Kritchevsky, and Bellugi 1988; and many others). Some of the spatial-mechanical conceptual machinery for representing motion is already available early in life: we know that infants in the first year of life detect changes in the path and manner of motion and find the invariant path and manner in actions (e.g., Pulverman et al. 2008; Pruden et al. 2013).

Second, and most pertinent for present purposes, despite being rooted in a shared conceptual typology, the linguistic encoding of motion and space is characterized by intense typological variability. For instance, the ways the motion primitives of manner and path are lexicalized in spatial vocabularies and the ways these primitives are conflated into sentential structure vary considerably crosslinguistically (Talmy 1975, 1985, 2000). Manner languages[3] (e.g., English, German, Russian, and Mandarin Chinese) typically encode manner in the verb (cf. English *skip, run, hop, jog*) and path in a variety of other devices such as particles (*out*), adpositions (*into the room*), or verb prefixes (e.g., German *raus-,* "out"; cf. *raus-rennen,* "run out"). Path languages (e.g., Modern Greek, Romance, Turkish, Japanese, and Hebrew) typically encode path in the verb (cf. Greek *vjeno,* "exit"; *beno,* "enter"; *diashizo,* "cross"), and manner in gerunds, adpositions, or

---

3. We adopt the distinction between path and manner languages introduced in Papafragou, Massey, and Gleitman (2002), rather than the original distinction between V-framed and S-framed languages in Talmy (1985) for ease of reference. The former distinction reflects broader lexical and morpho-syntactic asymmetries in the groups of languages involved and is therefore not specific to motion event encoding (see Beavers, Levin, and Tham 2010).

El

adverbials (*trehontas,* "running"; *me ta podia* "on foot"). Often, manner languages possess large and varied manner verb vocabularies, unlike path languages. The manner/path language distinction should not be taken to imply that the relevant languages lack certain kinds of verbs altogether. But the most characteristic (i.e., colloquial, frequent, and pervasive) way of describing motion in these languages involves manner and path verbs, respectively.

To illustrate these differences, the sentences in (3) and (4) describe the same event in English and in Greek. In (3), the manner of motion (*driving*) is encoded in the main verb, and the path appears as a prepositional phrase (*into the soccer net*). In (4), the same event is encoded in two different ways in Greek, none of which map directly onto their English counterpart: in (4a), the path of motion is encoded in the main verb, and the manner of motion is encoded in an optional prepositional phrase; in (4b), the event is encoded in two separate clauses, the first one containing a manner verb and the second a path verb. Of the two alternatives in (4), (4a) is more natural than (4b):

(3) A boy is driving a bike into a tent.

(4a) Ena agori beni se mia skini (me mia mihani).

"A boy enters in a tent (with a bike)."

(4b) Ena agori odigi mia mihani ke beni se mia skini.

"A boy is-driving a bike and is-entering in a tent,"

The typological differences between path and manner languages are reflected in the way speakers of different languages habitually talk about motion (e.g., Özçalişkan and Slobin 1999, 2003; Papafragou, Massey, and Gleitman 2002, 2006). These differences are already in place as early as three years of age: young children quickly grasp and follow language-specific patterns for the expression of motion, such that, for instance, English learners encode path concepts in prepositions such as *in* and *out*, while Turkish learners encode them in verbs with meanings such as *enter* and *exit* (Allen et al. 2007; Bowerman and Choi 2003; Slobin 1996, 2003; cf. Naigles et al. 1998; Papafragou, Massey, and Gleitman 2002, 2006). Furthermore, these statistical tendencies in motion descriptions affect conjectures about the meanings of novel verbs: when presented with a novel intransitive motion verb, English-speaking children and adults tend to interpret it as a manner verb, whereas Greek-speaking children and adults tend to construe it as a path verb (for similar data from other languages, see Papafragou and Selimis 2010b; Skordos and Papafragou 2010; cf. Maguire et al. 2010; Naigles and Terrazas 1998).

Could such lexicalization preferences affect perceptual and conceptual motion event encoding? Some commentators have proposed that the manner of motion for speakers

El

of manner languages is a "salient and differentiated conceptual field" compared with speakers of path languages, with potential implications for how manner of motion is perceived or attended to on-line and remembered (Slobin 2004; cf. Bowerman and Choi 2003). However, an initial series of studies with both adults and children found no evidence supporting this idea. For instance, speakers of English and Greek were equally likely to remember path and manner details of motion events, despite the differences in the way the two groups described these events (Papafragou, Massey, and Gleitman 2002). Similarly, the categorization of motion proceeded identically in speakers of the two languages (Papafragou, Massey, and Gleitman 2002; Papafragou and Selimis 2010a; see also Gennari et al. 2002).

Could linguistic encoding affect the dynamics of motion perception itself? The studies just reviewed involved slower, more complex cognitive processes, and thus leave open the possibility that on-line measures of event processing might capture early effects of language on attention. To test this possibility, we (Papafragou, Hulbert, and Trueswell 2008) conducted an eye tracking study in which adult speakers of Greek and English viewed a set of short animated motion events (e.g., a boy driving his bike into a tent). The stimuli were constructed so that linguistically relevant manner and path information could be easily defined as distinct regions spatially separated from each other: manners corresponded to the vehicle that propelled the moving agent (e.g., the bike), and paths corresponded to the object that served as the endpoint of the motion trajectory (e.g., the tent). Events unfolded for three seconds, at which point a beep was heard, and the last clip from the event remained on the screen for another two seconds. Half of the participants were asked to describe the events after hearing the beep (linguistic task). The other half were asked to watch the events carefully in preparation for a memory test (nonlinguistic task).

Of interest was to compare the way participants viewed these events. If language has deep, enduring effects on event perception, gaze patterns should differ cross-linguistically when people inspect motion events regardless of whether the task involves producing a linguistic description or not. Thus in both the linguistic and the nonlinguistic task, English speakers should look earlier and more often at manner of motion regions than Greek speakers, whereas Greek speakers should prioritize the path endpoint more.

As expected, the linguistic description data confirmed the typological difference between path and manner languages. For English speakers, 78 percent of all sentences describing an event of motion to an endpoint contained a manner verb, compared with only 32 percent for Greek speakers. Early eye movements in the linguistic task reflected this cross-linguistic difference: Greek speakers were more likely than English speakers to fixate the path endpoint first (the tent in the earlier example) rather than the manner of motion region (the bike), probably as a result of planning the informational content of the main motion verb in their sentence.

After about a second and a half, Greek speakers turned their attention to manner, while English speakers focused on the path endpoint, presumably as a result of the preparation of relevant postverbal modifiers in each language. These eye-movement patterns were repeated after the beep while people were describing aloud the events. This pattern is in accord with other eye-movement production studies, which were done within a single language, where participants' preparation to describe regions of a scene was preceded by fixations on these regions (e.g., Altmann and Kamide 2004; Griffin and Bock 2000). These results offer the first cross-linguistic evidence for the hypothesis that language production mobilizes attention: where languages differ from each other in how they encode event structure, this difference shapes how speakers visually interrogate events during speech planning (cf. also Bunger et al. 2013 for evidence that such shifts of attention during speech planning are present already in three- to four-year-old children).

Crucially, however, in the nonlinguistic task, where people were asked to simply study (but not describe) the events, attention allocation as the events unfolded was strikingly *similar* for both language groups: overall, people prioritized looks to the path endpoint and inspected the manner of motion slightly later. Thus, when inspecting the world freely, people were alike in how they perceived events, regardless of their native language. This finding argues against the hypothesis that linguistic encoding has deep and persistent effects on the way events are perceptually experienced but is entirely expected under accounts that emphasize universal aspects of event perception and cognition.

One striking finding from the nonlinguistic task introduced complexity into this picture: late in each trial, after the event had unfolded and its last frame remained on the screen, English speakers spent more time inspecting the path endpoint (e.g., the tent) rather than the manner of motion (e.g., the bike) compared with the Greek speakers, who tended toward the opposite pattern. We (Papafragou, Hulbert, and Trueswell 2008) call this a "reverse-Whorfian" finding, since each language group seemed to attend to aspects of the event that would not typically be encoded in a verb (path endpoint in English, manner in Greek).

How could this reverse-Whorfian pattern be explained? According to a first hypothesis, this effect reflects the on-line (albeit covert) use of language to help participants remember the core event and to direct attention to additional details in the scene in preparation for the memory test (see Antes 1974; Antes and Kristjanson 1993; and Loftus 1972 for evidence that participants typically begin by studying the fundamental aspects of a scene, and then increasingly fixate the "details"). Here, it appears that what was considered a detail was linguistically determined: for Greek speakers, the manner of motion was a detail of the event, because it was not preferentially encoded in the main verb of the event description, and for English speakers, the path endpoint

El

was a detail. According to a second hypothesis, the reverse-Whorfian pattern is due to conceptual reorganization, that is, to differences in postperceptual event processing between English and Greek speakers caused by years of language use. On this hypothesis, attention was drawn late to "details" of events not because speakers used on-line (covert) verbalization as a way to remember event details but because language has created deep, stable differences in how salient manners and paths are in the minds of English versus Greek speakers.

The method and data in Papafragou, Hulbert, and Trueswell (2008) could not distinguish between the two possible explanations of the reverse-Whorfian effect, but a further study (Trueswell and Papafragou 2010) tested these alternatives directly. This study presented a new group of English and Greek speakers with motion events that were very similar to the ones in Papafragou, Hulbert, and Trueswell (2008). Participants were placed in one of three conditions. In the No Interference condition, people were asked to watch and remember the events for a later memory test, much like in the original study. In the Linguistic Interference condition, people were given the same instructions as in the No Interference condition but were additionally presented with a secondary task: they heard a series of numbers before each event (e.g., "35, 20, 76") and were asked to repeat these numbers back to the experimenter as long as the stimulus was displayed. In the Nonlinguistic Interference condition, people were given the same basic set of instructions but a different secondary task: before each event, the experimenter demonstrated a series of rhythmic taps on the table, and participants had to reproduce the taps continuously until the stimulus finished.

The logic of the experiment was as follows. If the reverse-Whorfian effect reflected the on-line use of language to recode the motion stimuli, a secondary task that blocked on-line access to the linguistic code (as in the Linguistic Interference condition) should make the effect disappear; by contrast, a secondary task that did not affect the availability of the verbal code (as in the Nonlinguistic Interference condition) should allow the effect to surface. Alternatively, if the reverse-Whorfian pattern was due to a deeper conceptual reorganization of event representations driven by language-specific forces, it should persist whether or not covert verbalization was permitted within the experimental setup (i.e., in both the Linguistic and the Nonlinguistic Interference conditions).

Results from the study supported the on-line use hypothesis over the conceptual reorganization hypothesis. Specifically, the reverse-Whorfian effect appeared strongly in the Nonlinguistic Interference condition, yielding differences in attention allocation between English and Greek speakers. However, this effect disappeared in the Linguistic Interference condition. In this condition, as in the No Interference condition, attention allocation was very similar for English and Greek speakers. These results replicate and extend the findings from the nonlinguistic task of Papafragou, Hulbert, and Trueswell

(2008)[4] and are consistent with prior evidence for the independence of motion cognition from linguistic-encoding patterns (Papafragou, Hulbert, and Trueswell 2008; Papafragou, Massey, and Gleitman 2002, 2006; Gennari et al. 2002; cf. also Malt et al. 2008).

The studies summarized in this section point to a few key conclusions about the relationship between crosslinguistic event encoding and event representation. First, the fact that languages differ in how they encode events has implications about the way speakers of different languages inspect events in the world before they talk about what they see. Second, such differences in attention do not percolate into situations that do not involve communication: basic patterns in event perception are independent from one's native language. Third, language can be used on-line as a tool to support memory. This implicit use of the verbal code may give rise to crosslinguistic differences in attention allocation as event details are apprehended and committed to memory (what we have called a reverse-Whorfian effect). Such on-line effects of language are distinct from effects of permanent conceptual reorganization, since they can be disrupted through processes that block access to the verbal code (e.g., Linguistic Interference). Taken together, these conclusions paint a nuanced picture of how language makes contact with event representations. In this picture, rather than permanently shaping the mechanisms supporting event perception and cognition (as proposed by recent neo-Whorfian accounts), language offers an alternative, optionally recruited system of representing, organizing, and tracking events.

4. Unlike in Papafragou, Hulbert, and Trueswell (2008), we did not observe the reverse-Whorfian effect in the No Interference condition. We suspect that the reason has to do with methodological differences between the two studies (see Trueswell and Papafragou 2010 for discussion). In our prior study, during the nonlinguistic task, participants were given examples of event changes as practice for the memory test (e.g., a woman eating an apple would turn into a woman eating a pie). Participants were also told that the memory test would involve identifying whether static images were taken from the motion events they had earlier seen. As a result, participants knew that changes relevant for the memory test would be quite subtle and that the format of the items in the memory phase would be different from the one in the presentation phase (static vs. dynamic clips). By contrast, the No Interference condition in Trueswell and Papafragou (2010) contained no practice items, and the memory items had the same format as the presentation items (dynamic events). Both of these features may have convinced participants in the earlier but not the later study that the memory task would be difficult and might have encouraged spontaneous adoption of a linguistic recoding strategy so that event details could be kept in memory. This assumption is supported by the fact that when cognitive load was increased in the Trueswell and Papafragou study (as in the Nonlinguistic Interference condition), the reverse-Whorfian effect emerged clearly and early (e.g., for English speakers, it was present after only about 800 or 900 ms from the beginning of the event).

12.4   Events and Language: Final Thoughts

The studies discussed in this chapter present a novel approach to studying the nature and building blocks of perceptual and conceptual event representations and the processes that contribute to event apprehension (especially event category and event role recognition). Additionally, these studies paint a subtle picture of how events and language interact, with language both building on event representations and also potentially affecting on-line event cognition—though not reorganizing basic perceptual-cognitive processes (for related perspectives, see Dessalegn and Landau 2008; Frank et al. 2008).

The present approach opens up several questions for further work. For instance, what is the exact nature of event representations in the first moments of observing an event? The level of specificity of these representations will presumably depend on both local features (e.g., facial expressions or eye gaze) and more global features of the scene (e.g., spatial layout of participants; Oliva 2005) in ways that remain to be specified. Relatedly, how might the recognition of an event category (e.g., punching, filming) interact with recognition of event roles? Furthermore, what are the implications of event role hierarchies for the way children learn to segment and name events? Infant work suggests that event components such as agents and patients are available to humans at a very young age (e.g., Goldin-Meadow 1985; Golinkoff 1975; Golinkoff and Kerr 1978; Gordon 2003). It remains to be seen how nonlinguistic event representations affect the acquisition of different types of event predicates (e.g., verbs or prepositions). Finally, what are the circumstances under which language is recruited in nonlinguistic event processing? Language may be particularly useful for some tasks but not others (Potter 1979), and this is bound to affect the ways in which verbal encoding is brought to bear on the representation of events.

Acknowledgments

El

References

Allen, S., A. Özyürek, S. Kita, A. Brown, R. Furman, T. Ishizuka, and M. Fujii. 2007. Language-specific and universal influences in children's syntactic packaging of manner and path: A comparison of English, Japanese, and Turkish. *Cognition* 102:16–48.

Altmann, G., and Y. Kamide. 2004. Now you see it, now you don't: Mediating the mapping between language and the visual world. In *The Interface between Language, Vision and Action: Eye Movements and the Visual World*, ed. J. Henderson and F. Ferreira, 347–385. New York,; Hove, UK: Psychology Press.

Antes, J. R. 1974. The time course of picture viewing. *Journal of Experimental Psychology* 103 (1): 62–70.

Antes, J. R., and A. Kristjanson. 1993. Effects of capacity demands on picture viewing. *Perception & Psychophysics* 54:808–813.

Baillargeon, R., J. Li, Y. Gertner, and D. Wu. 2011. How do infants reason about physical events? In *The Wiley-Blackwell Handbook of Childhood Cognitive Development*, 2nd ed., ed. U. Goswami, 11–48. Oxford: Blackwell.

Baker, M. C. 1997. Thematic roles and syntactic structure. In *Handbook of Generative Syntax*, ed. L. Haegeman, 73–137. Dordrecht: Kluwer.

Beavers, J., B. Levin, and S. W. Tham. 2010. The typology of motion events revisited. *Journal of Linguistics* 46:331–377.

Blake, R., and M. Shiffrar. 2007. Perception of human motion. *Annual Review of Psychology* 58:47–73.

Bock, K., D. Irwin, and D. Davidson. 2004. Putting first things first. In *The Interface between Language, Vision and Action: Eye Movements and the Visual World*, ed. J. Henderson and F. Ferreira, 249–317. New York; Hove, UK: Psychology Press.

Boroditsky, L. 2001. Does language shape thought? Mandarin and English speakers' conception of time. *Cognitive Psychology* 43:1–22.

Bowerman, M., and S. Choi. 2003. Space under construction: Language-specific spatial categorization in first language acquisition. In *Language in Mind: Advances in the Study of Language and Thought*, ed. D. Gentner and S. Goldin-Meadow, 387–427. Cambridge, MA: MIT Press.

Bowerman, M., and S. Levinson, eds. 2001. *Language Acquisition and Conceptual Development*. Cambridge: Cambridge University Press.

Bunger, A., D. Skordos, J. Trueswell, and A. Papafragou. 2013. Cross-linguistic differences in event description and inspection in preschool speakers of English and Greek. Talk delivered at the 87th Annual Meeting of the Linguistic Society of America, Boston, January 3–6.

Carlson, G. 1998. Thematic roles and the individuation of events. In *Events and Grammar*, ed. S. Rothstein, 35–51. Dordrecht: Kluwer.

El

Carlson, L., and E. van der Zee, eds. 2005. *Functional Features in Language and Space: Insights from Perception, Categorization and Development.* Oxford: Oxford University Press.

Chomsky, N. 1981. *Lectures on Government and Binding.* Dordrecht: Foris.

Dessalegn, B., and B. Landau. 2008. More than meets the eye: The role of language in binding and maintaining feature conjunctions. *Psychological Science* 19 (2): 189–195.

Dobel, C., R. Glanemann, H. Kreysa, P. Zwitserlood, and S. Eisenbeiss. 2011. Visual encoding of coherent and non-coherent scenes. In *Event Representation in Language: Encoding Events at the Language Cognition Interface*, ed. J. Bohnemeyer and E. Pederson, 189–215. Cambridge: Cambridge University Press.

Dobel, C., H. Gumnior, J. Bölte, and P. Zwitserlood. 2007. Describing scenes hardly seen. *Acta Psychologica* 12:129–143.

Dowty, D. 1991. Thematic proto-roles and argument selection. *Language* 67:547–619.

Fillmore, C. J. 1968. The case for case. In *Universals in Linguistic Theory*, ed. E. Bach and R. T. Harms, 1–88. New York: Holt, Rinehart and Winston.

Frank, M. C., D. L. Everett, E. Fedorenko, and E. Gibson. 2008. Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition* 108:819–824.

Gennari, S., S. Sloman, B. Malt, and T. Fitch. 2002. Motion events in language and cognition. *Cognition* 83:49–79.

Gentner, D., and S. Goldin-Meadow, eds. 2003. *Language and Mind.* Cambridge, MA: MIT Press.

Giese, M. A., and T. Poggio. 2003. Neural mechanisms for the recognition of biological movements. *Nature Reviews: Neuroscience* 4:179–192.

Gleitman, L. R. 1990. The structural sources of verb learning. *Language Acquisition* 1:1–63.

Gleitman, L., D. January, R. Nappa, and J. Trueswell. 2007. On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language* 57:544–569.

Gleitman, L., and A. Papafragou. In press. New perspectives on language and thought. In *Cambridge Handbook of Thinking and Reasoning*, ed. K. Holyoak and B. Morrison. Cambridge: Cambridge University Press.

Göksun, T., K. Hirsh-Pasek, and R. M. Golinkoff. 2010. Trading spaces: Carving up events for learning language. *Perspectives on Psychological Science* 5:33–42.

Goldin-Meadow, S. 1985. Language development under atypical learning conditions: Replication and implications of a study of deaf children of hearing parents. In *Children's Language*, vol. 5, ed. K. Nelson, 197–245. Hillsdale, NJ: Erlbaum.

Golinkoff, R. M. 1975. Semantic development in infants: The concepts of agent and recipient. *Merrill-Palmer Quarterly* 21:181–193.

El

Golinkoff, R. M., and J. L. Kerr. 1978. Infants' perception of semantically defined action role changes in filmed events. *Merrill-Palmer Quarterly* 24:53–61.

Gordon, P. 2003. The origin of argument structure in infant event representations. In *Proceedings of the 26th Boston University Conference on Language Development*, ed. B. Skarabela, S. Fish, and A. H.-J.Do, 189–198. Somerville, MA: Cascadilla Press.

Griffin, Z., and K. Bock. 2000. What the eyes say about speaking. *Psychological Science* 11:274–279.

Hafri, A., A. Papafragou, and J. Trueswell. 2011. The gist of events: Implications for language production and scene description. Poster presented at the CUNY Conference on Human Sentence Processing, Stanford University, March 24–26.

Hafri, A., A. Papafragou, and J. C. Trueswell. 2013. Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experimental Psychology: General* 142:880–905.

Hayward, W. G., and M. J. Tarr. 1995. Spatial language and spatial representation. *Cognition* 55:39–84.

Jackendoff, R. 1990. *Semantic Structures*. Cambridge, MA: MIT Press.

Jackendoff, R. 1996. The architecture of the linguistic-spatial interface. In *Language and Space*, ed. P. Bloom, M. Peterson, L. Nadel, and M. Garrett, 1–30. Cambridge, MA: MIT Press.

Johansson, G. 1973. Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics* 14:201–211.

Kable, J. W., and A. Chatterjee. 2006. Specificity of action representations in the lateral occipito-temporal cortex. *Journal of Cognitive Neuroscience* 18:1498–1517.

Lange, J., and M. Lappe. 2006. A model of biological motion perception from configural form cues. *Journal of Neuroscience* 26:2894–2906.

Lashley, K. S. 1951. The problem of serial order in behavior. In *Cerebral Mechanisms in Behavior*, ed. L. A. Jeffress, 112–136. New York: Wiley.

Levelt, W. 1989. *Speaking*. Cambridge, MA: MIT Press.

Levin, B., and Pinker, S. 1991. Introduction. In Lexical and conceptual semantics, special issue, *Cognition* 41:1–7.

Levin, B., and M. Rappaport Hovav. 2005. *Argument Realization*. Cambridge: Cambridge University Press.

Levinson, S. C. 1996. Frames of reference and Molyneux's question: Crosslinguistic evidence. In *Language and Space*, ed. P. Bloom, M. Peterson, L. Nadel, and M. Garrett, 109–170. Cambridge, MA: MIT Press.

Levinson, S. C. 2003. *Space in Language and Cognition*. Cambridge: Cambridge University Press.

Loftus, G. 1972. Eye fixations and recognition memory for pictures. *Cognitive Psychology* 3:525–551.

Maguire, M., K. Hirsh-Pasek, R. M. Golinkoff, M. Imai, E. Haryu, S. Vengas, H. Okada, R. Pulverman, and B. Sanchez-Davis. 2010. A developmental shift from similar to language specific strategies in verb acquisition: A comparison of English, Spanish, and Japanese. *Cognition* 114:299–319.

Majid, A., M. Bowerman, S. Kita, D. B. Haun, and S. C. Levinson. 2004. Can language restructure cognition? The case for space. *Trends in Cognitive Sciences* 8:108–114.

Malt, B. C., S. Gennari, M. Imai, E. Ameel, N. Tsuda, and A. Majid. 2008. Talking about walking: Biomechanics and the language of locomotion. *Psychological Science* 19:232–240.

Miller, G., and P. Johnson-Laird. 1976. *Language and Perception.* Cambridge, MA: Harvard University Press.

Naigles, L., A. Eisenberg, E. Kako, M. Highter, and N. McGraw. 1998. Speaking of motion: Verb use in English and Spanish. *Language and Cognitive Processes* 13:521–549.

Naigles, L. R., and P. Terrazas. 1998. Motion-verb generalizations in English and Spanish: Influences of language and syntax. *Psychological Science* 9:363–369.

Newcombe, N., and J. Huttenlocher. 2003. *Making Space: The Development of Spatial Representation and Reasoning.* Cambridge, MA: MIT Press.

Newtson, D. 1973. Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology* 28:28–38.

Newtson, D. 1976. The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology* 12:436–450.

Oliva, A. 2005. Gist of the scene. In *Neurobiology of Attention*, ed. L. Itti, G. Rees, and J. K. Tsotsos, 251–257. San Diego: Elsevier Academic Press.

Özçalışkan, Ş., and D. I. Slobin. 1999. Learning how to search for the frog: Expression of manner of motion in English, Spanish, and Turkish. In *Proceedings of the 23rd Annual Boston University Conference on Language Development,* ed. A. Greenhill, H. Littlefield, and C. Tano, 541–552. Somerville, MA: Cascadilla Press.

Özçalışkan, Ş., and D. I. Slobin. 2003. Codability effects on the expression of manner of motion in Turkish and English. In *Studies in Turkish Linguistics*, ed. A. S. Özsoy, D. Akar, M. Nakipoğlu-Demiralp, E. Erguvanlı-Taylan, and A. Aksu-Koç, 259–270. Istanbul: Boğaziçi University Press.

Papafragou, A., J. Hulbert, and J. C. Trueswell. 2008. Does language guide event perception? Evidence from eye movements. *Cognition* 108:155–184.

Papafragou, A., C. Massey, and L. Gleitman. 2002. Shake, rattle, 'n' roll: The representation of motion in thought and language. *Cognition* 84:189–219.

El

Papafragou, A., C. Massey, and L. R. Gleitman. 2006. When English proposes what Greek presupposes: The linguistic encoding of motion events. *Cognition* 98:B75–B87.

Papafragou, A., and S. Selimis. 2010a. Event categorisation and language: A cross-linguistic study of motion. *Language and Cognitive Processes* 25:224–260.

Papafragou, A., and S. Selimis. 2010b. Lexical and structural biases in the acquisition of motion verbs. *Language Learning and Development* 6:87–115.

Paul, H. [1886] 1970. The sentence as the expression of the combination of several ideas. In *Language and Psychology: Historical Aspects of Psycholinguistics*, ed. A. L. Blumenthal, 34–37. New York: Wiley.

Pinker, S. 1989. *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.

Potter, M. C. 1979. Mundane symbolism: The relations among objects, names, and ideas. In *Symbolic Functioning in Childhood*, ed. N. R. Smith and M. B. Franklin, 41–65. Hillsdale, NJ: Erlbaum.

Pruden, S. M., T. Göksun, S. Roseberry, K. Hirsh-Pasek, and R. M. Golinkoff. 2013. Infant categorization of path relations during dynamic events. *Child Development* 84:331–345.

Pulverman, R., R. M. Golinkoff, K. Hirsh-Pasek, and J. Sootsman Buresh. 2008. Manners matter: Infants' attention to manner and path in non-linguistic dynamic events. *Cognition* 108:825–830.

Shipley, T. F., and J. M. Zacks, eds. 2008. *Understanding Events: From Perception to Action*. New York: Oxford University Press.

Skordos, D., and A. Papafragou. 2010. Extracting paths and manners: Linguistic and conceptual biases in the acquisition of spatial language. In *Proceedings from the 34th Annual Boston University Conference on Language Development,* ed. K. Franich, K. M. Iserman, and L. L. Keil. Somerville, MA: Cascadilla Press.

Slobin, D. 1996. From "thought and language" to "thinking for speaking." In *Rethinking Linguistic Relativity*, ed. J. Gumperz and S. Levinson, 70–96. New York: Cambridge University Press.

Slobin, D. I. 2003. Language and thought online: Cognitive consequences of linguistic relativity. In *Language in Mind: Advances in the Study of Language and Thought*, ed. D. Gentner and S. Goldin-Meadow, 157–191. Cambridge, MA: MIT Press.

Slobin, D. I. 2004. The many ways to search for a frog: Linguistic typology and the expression of motion events. In *Relating Events in Narrative: Typological and Contextual Perspectives*, ed. S. Strömqvist and L. Verhoeven, 219–257. Mahwah, NJ: Erlbaum.

Smith, L., S. Jones, B. Landau, L. Gershkoff-Stowe, and L. Samuelson. 2002. Object name learning provides on-the-job training for attention. *Psychological Science* 13:13–19.

El

Spelke, E. S., A. T. Phillips, and A. L. Woodward. 1995. Infants' knowledge of object motion and human action. In *Causal Cognition: A Multidisciplinary Debate*, ed. D. Sperber, D. Premack, and A. Premack, 44–78. New York: Oxford University Press.

Stiles-Davis, J., M. Kritchevsky, and U. Bellugi, eds. 1988. *Spatial Cognition: Brain Bases and Development*. Hillsdale, NJ: Erlbaum.

Talmy, L. 1975. Semantics and syntax of motion. In *Syntax and Semantics*, vol. 4, ed. J. P. Kimball, 181–238. New York: Academic Press.

Talmy, L. 1985. Lexicalization patterns: Semantic structure in lexical forms. In *Grammatical Categories and the Lexicon*, vol. 3, ed. T. Shopen, 57–149. Cambridge: Cambridge University Press.

Talmy, L. 2000. *Typology and Process in Concept Structuring*. Vol. 2, Toward a Cognitive Semantics. Cambridge, MA: MIT Press.

Tranel, D., D. Kemmerer, R. Adolphs, H. Damasio, and A. R. Damasio. 2003. Neural correlates of conceptual knowledge for actions. *Cognitive Neuropsychology* 20:409–432.

Trueswell, J., and A. Papafragou. 2010. Perceiving and remembering events cross-linguistically: Evidence from dual-task paradigms. *Journal of Memory and Language* 63:64–82.

Webb, A., A. Knott, and M. R. MacAskill. 2010. Eye movements during transitive action observation have sequential structure. *Acta Psychologica* 133:51–56.

Whorf, B. L. 1956. *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf*. Ed. J. B. Carroll. Cambridge, MA: MIT Press.

Wilson, F., A. Papafragou, A. Bunger, and J. Trueswell. 2011. Rapid extraction of event participants in caused motion events. In *Proceedings from the 33rd Annual Meeting of the Cognitive Science Society*, ed. L. Carlson, C. Hoelscher, and T. F. Shipley. Hillsdale, NJ: Erlbaum.

Wilson, F., A. Papafragou, and J. Trueswell. 2011. Mapping event participants to linguistic structure. Poster presented at Structuring the Argument: A Multidisciplinary Workshop on the Mental Representation of Verbal Argument Structure, Paris, September 5–7.

Winawer, J., N. Witthoft, M. C. Frank, L. Wu, and L. Boroditsky. 2007. Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences of the United States of America* 104:7780–7785.

Wundt, W. [1900] 1970. The psychology of the sentence. In *Language and Psychology: Historical Aspects of Psycholinguistics,* ed. A. L. Blumenthal, 20–31. New York: Wiley

Zacks, J. M., B. Tversky, and G. Iyer. 2001. Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General* 130:29–58.

El

El