Glossa
a journal of general linguistics

How children attend to events before speaking: crosslinguistic evidence from the motion domain

RESEARCH

ANN BUNGER
DIMITRIOS SKORDOS
JOHN C. TRUESWELL
ANNA PAPAFRAGOU
*Author affiliations can be found in the back matter of this article

ubiquity press

Abstract

How do children talk about the dynamic world around them? In this eyetracking study, we demonstrate language-specific patterns in the way 3- and 4-year-old speakers of English and Greek inspect motion events prior to speaking and describe such events in their native language. Across age and language groups, children were more likely to mention manners of motion than paths, but English-speaking children were more likely to provide manner information than Greek-speaking children were. Comparison of eyegaze patterns from the linguistic (description) task to eyegaze patterns observed during a nonlinguistic (memory) task with a different group of English- and Greek-speaking 3- and 4-year-olds revealed effects of language background on event inspection. These effects suggest that by the age of 3 years, children exhibit sensitivities to language-specific patterns of motion event encoding that influence the way they gather information from the visual world during the process of language production.

CORRESPONDING AUTHOR:
Ann Bunger
Department of Linguistics, Indiana University, US
acbunger@indiana.edu
Anna Papafragou
Department of Linguistics, University of Pennsylvania, US
anna4@sas.upenn.edu

KEYWORDS:
motion events; crosslinguistic differences; language production; event cognition; acquisition

TO CITE THIS ARTICLE:
Bunger, Ann, Dimitrios Skordos, John C. Trueswell and Anna Papafragou. 2021. How children attend to events before speaking: crosslinguistic evidence from the motion domain. *Glossa: a journal of general linguistics* 6(1): 28. 1–22. DOI: *https://doi.org/10.5334/gjgl.1210*

1 Introduction

During language production, we select information from our conceptual representations and encode it in linguistic strings. When we talk about events that are happening in the world around us, we usually mention only a subset of the information that is available to us. The process of selecting which information to include may be guided by the structure of our conceptual representations, by our communicative goals, and by the language that we speak (Levelt 1989; see also Papafragou & Grigoroglou 2019, for a recent review).

When describing an event in which a person moves from one location to another we tend to mention aspects of the event that are central to making sense of the motion—for example, *where* the person is going (into the cafe, across the pool, up Mount Everest) and *how* they are moving (walking, swimming, climbing)—and to omit other information (e.g., what the person is wearing). Path (*where*) and Manner (*how*) information correspond to conceptual components of motion events that are understood from a very early age. Evidence for the conceptual basicness of these components comes from the fact that some of the earliest words that children tend to use describe paths ("up") and manners ("dance") of motion (Fenson et al. 1994). Furthermore, the same meanings are encoded by the early gestures produced by deaf home signers who have not been exposed to a conventional language model (Zheng & Goldin-Meadow 2002). Additionally, there is experimental evidence that young children can discriminate both paths and manners of motion in nonlinguistic tasks by the time they are 14 months old (Pulverman et al. 2003; Pruden et al. 2004; Pulverman & Golinkoff 2004; Pulverman et al. 2006; Pruden et al. 2008; Pulverman et al. 2008; Göksun et al. 2017). In one study, when 14- to 17-month-old children learning different languages were habituated to an animated character that moved with a particular manner (e.g., rotating along a horizontal axis) along a particular path (e.g., over a stationary shape), their responses to following stimuli indicated that they were sensitive to both the manner in which the character moved and the path that was followed (Pulverman et al. 2008).

Despite these commonalities, the way motion event information is encoded in language is subject to robust language-specific factors. According to an influential proposal, most languages tend to fall into one of two typologically distinct classes that can be distinguished by where information about the core conceptual component of Path is encoded (e.g., Talmy 1975; 1985; 1991). Speakers of so-called satellite-framed languages, like English and German, tend to describe motion events by encoding information about manner of motion in the main verb of a sentence and path information in mostly non-verb elements (satellites; cf. Slobin & Hoiting 1994; Slobin 1996a). Consider the English sentence in (1): the verb "sailed" provides information about the manner of motion and the prepositional phrase "into the harbor" provides information about the path. In contrast, speakers of so-called verb-framed languages, like Modern Greek and French, often describe motion events by encoding information about the path of motion in the main verb and manner information (if it is included) in satellites, especially when describing motion events that involve boundary crossing (Aske 1989; Slobin & Hoiting 1994; Papafragou, Massey & Gleitman 2003; Hickmann & Hendriks 2006; Selimis & Katis 2010; Özçalişkan 2013; Georgakopoulos, Hörtl, & Sioupi 2019). The Greek sentence in (2) provides an example of this pattern: the verb "bike" ('entered') provides information about the path and the prepositional phrase "me to skafos tu" ('with his boat') provides information about how the man got there.

(1) A man sailed into the harbor.

(2) Enas anthropos bike sto limani me toh skafos tu.
A.NOM human.NOM entered in-the.ACC harbor.ACC with the.ACC boat.ACC his.
'A person entered the harbor with his boat'

As several commentators have noted, the verb-framed vs. satellite-framed distinction is not an absolute dichotomy but allows for degrees of convergence on a single pattern depending on lexical, morphosyntactic and even pragmatic aspects of event encoding (Skopeteas 2008; Beavers et al. 2010; among others). Within the class of verb-framed languages, there is considerable variation in how frequently path and manner information is encoded during production and how this information is distributed across the sentence (e.g., Slobin 2004; Soroli

& Verkerk 2017). Similarly, within individual languages in the typology, there is considerable variation in attested patterns of motion encoding: for instance, Greek sometimes exhibits mixed preferences by encoding manner information in the verb and/or packaging path information in prepositional phrases or particles (Talmy 2000; Selimis & Katis 2010; Soroli 2012). Nevertheless, crosslinguistic differences in motion event encoding predicted by the verb-framed vs. satellite-framed divide have been documented extensively in adults (e.g., Talmy 1975; 1985; Aske 1989; Talmy 1991; Slobin & Hoiting 1994; Slobin 1996a; Naigles et al. 1998; Papafragou et al. 2006; Özçalişkan 2013), and are known to emerge early during development. Papafragou and Selimis (2010b) demonstrated that 5-year-old Greek- and English-speaking children have already begun to prioritize manner and path elements in motion event descriptions like adult speakers of their target language, with English-speaking children more likely than Greek-speaking children to describe motion events with verbs that encode manner of motion and Greek-speaking children more likely than English-speaking children to use path verbs (cf. Papafragou et al. 2002; 2006; Papafragou & Selimis 2010a for a similar finding in older children). Even earlier effects of language environment on motion event description have been demonstrated in experimental studies comparing 3-year-old speakers of the verb-framed languages Turkish (Özçalişkan & Slobin 2000; Allen et al. 2007; Özyürek et al. 2008), French (Hickmann & Hendriks 2006; Hickmann et al. 2009; Hickmann et al. 2018) and German (Hickmann et al. 2018) to age-matched English speakers, as well as in studies comparing the spontaneous speech of 2-year-old Korean- and English-speaking children (Choi & Bowerman 1991). Crosslinguistic differences in the encoding of event components have been documented in the verb learning patterns of speakers as young as 3 years of age (Maguire et al. 2010; Skordos & Papafragou 2014). Across many (though not all; Slobin 2004; Soroli & Verkerk 2017) of these studies, speakers of satellite-framed languages were overall more likely than speakers of verb-framed languages to mention manners of motion: in one study (Papafragou et al. 2006), Greek speakers added manner of motion modifiers when the manners were novel or unexpected (cf. *A man went up the stairs running*) but English speakers encoded manners of motion in the verb regardless of typicality.

These language-specific biases in event description are reflected in systematized differences in the way adult speakers of typologically different languages select motion information to talk about during speech planning. Tracking speaker eyegaze during event viewing provides a window onto this process of information gathering as it unfolds over time (e.g., Griffin & Bock 2000; Bock et al. 2004; Griffin 2004; Meyer 2004; Papafragou et al. 2008; Trueswell & Papafragou 2010). As speakers inspect static or dynamic events with the intention to describe them, their patterns of eyegaze reveal the visual elements that they inspect, as well as the time course along which event information is gathered. These eyegaze patterns can, in turn, be linked to the content and form of the event descriptions that speakers eventually produce, providing insight into the mapping between conceptual and linguistic event representations. Critically, patterns of event inspection observed during the planning stages of language production (message planning, lexical selection, grammatical encoding) differ from those observed when people are engaged in nonlinguistic tasks.

In general, it is known that when adults view an event while planning to talk about what they see, they direct their attention very quickly to components of the scene that they plan to talk about, usually in the order that they plan to mention them (Griffin & Bock 2000). Papafragou and colleagues (2008) demonstrated, moreover, that while adult speakers of English and Greek were engaged in the process of describing motion events, they exhibited language-specific differences in event inspection that reflected differences in motion event description in these languages. Specifically, they found that when describing bounded motion events (i.e., motion that involves a goal like ***Figure 1***), English speakers were more likely than Greek speakers to use manner verbs, and the opposite held for path verbs. Consistent with these linguistic choices, when planning their event descriptions, adult speakers of these two languages directed their attention very early to event components that they planned to encode in the verb of their sentence: English speakers to event elements that provided information about the manner of motion (i.e., vehicles, instruments) and Greek speakers to elements that defined the path (i.e., Ground objects). Crucially, these crosslinguistic differences in event inspection only surfaced when participants recruited linguistic resources to accomplish a task: when they were presented with a free-viewing task that did not require the use of language, adult speakers of English and Greek did not show the same crosslinguistic differences in eyegaze patterns (see also Trueswell

& Papafragou 2010). Thus, language-specific differences in the way information is gathered from motion events are driven by the process of "thinking for speaking" (Slobin 1996b; 2006), and not by fundamental differences in nonlinguistic cognition between the two language groups (see also MacDonald 2013; Norcliffe et al. 2015; Skordos et al. 2020).

Figure 1 Sample motion event of a man sailing to an island. The still frames are extracted from the animated clip and depict the beginning, midpoint and endpoint of the event.

The way experience with a particular language affects production and attention patterns in children is not as well understood. Bunger and colleagues (2012) demonstrated that, by the time they are 4 years old, English-speaking children exhibit the same fundamental link between attention allocation and linguistic output that has been observed in adults. They found that not only do English-speaking children of this age tend to mention manners of motion more often than paths when they describe motion events, but like English-speaking adults, they also tend to direct more attention to the manners of motion events while they plan those event descriptions. It is as yet unknown, however, whether children of this age will show the same crosslinguistic differences in their attention to motion events during the planning stages of language production that adults do. Recall that crosslinguistic differences in the encoding of event components have been documented in speakers as young as 3 years of age. However, very little is known about the extent to which these differences are reflected in attention patterns as young speakers plan those utterances. The study of online attention patterns during language planning provides insight into the process that speakers are going through when selecting information from the visual world to talk about. By investigating similarities and differences in the way preschool-aged speakers of different languages prepare event descriptions before verbalizing them, we begin to tease apart behaviors during language planning that are shared from those that are specific to the acquisition of a particular language.

In the current study, we ask whether children exhibit language-specific differences in attention allocation during speech planning, and if so, whether those differences are linked to what they actually say about the events. This is one of the first studies to combine eyegaze measures with crosslinguistic event description in preschool-aged children. By investigating how early in development crosslinguistic differences in event description and attention begin to emerge, we add to the growing body of knowledge about the kinds of linguistic differences that children are sensitive to. Moreover, we begin to fill gaps in our understanding of how developmental and crosslinguistic differences influence the way information is selected during the process of language production.

Specifically, we ask whether children learning English and Greek exhibit systematic crosslinguistic patterns of attention as they plan descriptions of motion events. We look for evidence of language-specific biases 1) in the information that 3- and 4-year-old speakers of English and Greek provide when they describe dynamic motion events and 2) in their patterns of event inspection as they plan those descriptions. Following the research we summarized previously, we expect to find English and Greek-speaking children to differ in their tendency to provide information about manners (typically prioritized in English descriptions of motion events) and paths (typically prioritized in Greek descriptions of motion events) of motion. By comparing eyegaze patterns across language groups in conjunction with production of event descriptions, we fill a gap in the understanding of the way children in these age and language groups gather information about motion events in real time. Specifically, we aim to investigate whether the way they direct their attention during speech planning is linked to the information they provide about a motion event (manner vs. path), as has been demonstrated for adult speakers of these languages. To the extent that preschool-aged speakers of the two languages differ in their tendency to mention manner and/or path information in their event descriptions, we also expect them to allocate their attention differently while preparing those descriptions.

In addition, we compare eyegaze patterns during motion event description by children in these age and language groups to their eyegaze while viewing the same motion events in a nonlinguistic (memory) task. Here, we expect to find that, as for adult speakers of these languages, children exhibit different patterns of attention allocation to motion events when they are engaged in the process of language production versus when they are viewing the events in preparation for a memory task. As mentioned previously, adult speakers of these languages show similar patterns of attention to motion events when viewing them in preparation for a memory task (e.g., Papafragou et al. 2008; Trueswell & Papafragou 2010). This experiment will allow us to investigate whether patterns of attention to motion event components during nonlinguistic tasks also converge for 3- and 4-year-old speakers of these languages.

2 Method

2.1 Participants

The final data sample consisted of 79 children who were learning English or Greek as their native language. Children from two age groups were included: 3-year-olds (English: $n = 19$, mean age 3;5, range 3;0–3;11; Greek: $n = 20$, mean age 3;7, range 3;2–3;11) and 4-year-olds (English: $n = 20$, mean age 4;6,[1] range 4;0–5;0; Greek: $n = 20$, mean age 4;6, range 4;0–5;0). English-speaking children were recruited through preschools in Newark, DE ($n = 31$) and Philadelphia, PA ($n = 8$). Greek-speaking children were recruited through public ($n = 5$) and private ($n = 35$) preschools in and around Ioannina, Greece. Children had no parent-reported history of visual, cognitive, or language impairments. Data from an additional 19 children were excluded from the analysis for the following reasons: unwillingness to cooperate ($n = 4$), experimenter error or equipment failure ($n = 6$), failure to calibrate ($n = 1$), production of linguistic data that were not compatible with our coding rubric ($n = 3$; see "Coding of event descriptions" for more information), or significant trackloss during stimulus viewing ($n = 5$; see "Analysis of eye movement data" for trackloss criteria). Sample size was determined on the basis of previous eye tracking studies of motion descriptions in adults (e.g., Papafragou et al. 2008; Trueswell & Papafragou 2010).

2.2 Apparatus

Stimulus presentation and data collection were carried out using either a Tobii 1750 (77 children) or a Tobii T60 (2 children) remote eyetracking system (we used two systems because of a switch in lab equipment). The T60 is an updated version of the 1750 system: both systems track binocular eyegaze using optics embedded in a 17-in TFT flat panel monitor with a display size of 33.5 (width) × 26.8 (height) cm (31.2 deg × 25.1 deg visual angle at viewing distance of 60 cm). Both systems were set to a screen resolution of 1024 × 768. In our Tobii 1750 setup, two laptop computers running the Windows XP operating system controlled the eyetracking system: one computer displayed stimuli on the 1750 monitor (via the ClearView software from Tobii Technology); the other collected data from the eyetracker at a consistent 50 Hz sampling rate (via the TET-server software from Tobii Technology). The T60 uses an embedded server to collect data at a consistent 60 Hz sampling rate. In our T60 setup, we used a laptop computer running the Windows 7 operating system to control the display of stimuli (via the Tobii Studio software from Tobii Technology). To increase timing accuracy, all laptops in both systems were disconnected from the internet. To reconcile differences in sampling frequencies across the two systems, eyegaze data were analyzed as proportions of looking to various regions of interest during 1-s windows of the test period.

2.3 Materials

Stimuli consisted of short (9-s) videos that were created by animating clip-art images. Twelve target event videos depicted motion events in which a human or animal agent used an instrument or vehicle to move toward a stationary object (see ***Figure 1*** for a sample, and Appendix A for the full list). To assess familiarity with these instruments or vehicles in English and Greek speakers,

1 The 4-year-old English-speakers in this study are the same children described in Bunger, Trueswell and Papafragou (2012). In this paper, we take a different approach to the analysis of their eye movements and event descriptions.

we asked 10 adult speakers of each language to indicate their own familiarity with each item on a 5-point scale (1 = not very familiar, 5 = very familiar). There was no difference in the average familiarity of the vehicles between language groups (average familiarity score of 4.7 for both).

Because our goal was to assess attention to motion event components, Manners and Paths of motion in target events were represented by distinct objects in the scene. A simple, contextually appropriate background was also created for each video (e.g., a body of water). All Manners of motion were associated with the instrument or vehicle used by the agent (e.g., boat, ice skates, airplane), and clipart images were constructed so that the instrument was spatially separated from the torso and face of the agent, allowing looks to these two components to be distinguished in the analysis of eyegaze data. (We did not use spontaneous motion events just as walking or jumping, where the Manner of motion region cannot be reliably separated from the agent region.) All Paths involved movement of the agent toward a goal object (e.g., island, fishing hut, cave) that determined the Path endpoint for each event. Trajectories of agent motion were never marked by visual paths like winding roads or wake trailing a boat. Goal paths were chosen for all events because they are known to be more salient than source paths in both conceptualization and description of motion events (e.g., Regier 1996; Lakusta & Landau 2005). The specific representations of Manner and Path information in these events represent a limited set of what may be conceptually or linguistically encoded as manner or path—many motion events do not include instruments, and paths typically include more than the goal of a trajectory; these choices were made to create clear regions of interest for the eyetracking analysis that included only visible items in the stimuli within any given frame and made no additional assumptions about how the viewers conceptualized each visible item.

Twelve filler event videos depicted animate agents and inanimate objects involved in events that did not include specific endpoints (e.g., flying a kite; see Appendix A for a full list). The animation in all videos lasted for 3 s, and then the final frame of the event remained visible on the screen for an additional 6 s. When the animation ended (at 3 s), participants heard a beep; aside from this beep, all videos were silent. Clipart animations were first created in Microsoft PowerPoint and then modified and exported as Audio Video Interleave (avi) files using Apple's Final Cut Pro software. When presented on the screen of either Tobii system, stimulus videos were 23.6 (width) × 16.7 (height) cm (22.2 × 15.9 deg visual angle at a viewing distance of 60 cm).

2.4 Procedure and experimental design

All children were tested in their preschools by a native speaker of their own language. During the experiment, children sat unconstrained in a car seat firmly attached to a stationary chair placed approximately 60 cm from the eyetracker screen. The experimenter adjusted the angle of the screen for each child to obtain robust views of both eyes that were centered in the tracker's field of view. Calibration was carried out using Tobii's default 5-point calibration scheme. If the calibration was incomplete (data for fewer than 4 points were captured) or was judged by the experimenter to be otherwise unacceptable, the calibration routine was repeated, with adjustments made to the position of the child or the eyetracker, as necessary. As mentioned previously, one child who failed to calibrate was excluded from the analysis.

After the calibration routine, participants were given instructions for their task. There were two experimental tasks; participants were assigned to tasks at random, based on a rotation through an experimenter-generated list. Half of the participants in each age and language group were assigned to a Linguistic task, and the other half were assigned to a Nonlinguistic task. Instructions were presented to children in their native language. In both of the tasks, participants were informed that they would be viewing "cartoons" with "people and animals doing things."[2] Children in the Linguistic task were asked to tell the experimenter "what happened in the cartoon" as soon as they heard the beep that signaled the end of the animation.

2 We chose these general instructions because "doing something" can refer to either a manner ("sailing") or a path ("entering"). See Papafragou and Selimis (2010a) for evidence that people's interpretation of this phrase in a different, more ambiguous (categorization) task can take on either path or manner nuances.

Participants in the Nonlinguistic task were asked to "watch the cartoons very carefully" because the experimenter would be asking them "some questions about them later."[3]

Participants in each task viewed the same progression of stimuli presented in a fixed semi-random order: 4-year-olds viewed the entire set of 24 items (12 targets and 12 fillers), and, because pilot testing suggested that the full set of stimuli was too long for them, 3-year-olds were presented with a subset of 16 of these items (8 targets and 8 fillers). A recentering animation in which colorful objects (e.g., stars and smiley faces) flew around the screen was shown between all stimulus items. This animation allowed the experimenters to recapture the gaze of inattentive preschoolers while at the same time avoiding directing their attention to any particular location on the screen. Participants in the Linguistic task provided their event descriptions aloud, and these sessions were audio-recorded. Participants in the Nonlinguistic task were discouraged from engaging in linguistic encoding of the events: children in this condition who began to give descriptions were reminded to "watch quietly."

2.5 Data coding and analysis

2.5.1 Coding of event descriptions

Descriptions of stimulus events collected from participants in the Linguistic condition were transcribed and coded by native speakers of the language under consideration. Event descriptions were not available for 11 of the 392 Linguistic trials: 1 trial was skipped due to experimenter error, and 10 trials did not elicit intelligible event descriptions. These trials were excluded from all analyses. For the remaining trials, descriptions of target items were assessed for mention of the Manners and Paths of motion depicted in the event. Words or phrases that referred to instruments (e.g., "boat") or the agent's manner of motion (e.g., "sailing," "floating," "driving") were coded as Manner mentions, and those that referred to either the path endpoint (e.g., "island," "beach"), the agent's trajectory of motion (e.g., "went to"), or the relationship between the agent and the path endpoint (e.g., "reached") were coded as Path mentions. In addition, to ensure that we were coding motion event components rather than just information about objects, all utterances included in the dataset mentioned motion and/or boundary crossing. For example, an utterance like (3a) would be coded as including both Manner and Path information, whereas (3b) includes only Manner information and (3c) includes only Path information.

(3) a. He went to an island in a sailboat.
b. He was sailing the boat.
c. He went to an island.

Event descriptions that did not include information about either the Manner or the Path of the associated target event were coded as "Neither." Moreover, because we were interested in children's mention of event components rather than the instrument and goal objects that depicted them, we excluded from the analysis 59 event descriptions that consisted of no more than labels for instruments ("this is a ship") or goals ("house"): 15 of 71 items from English-speaking 3-year-olds, 27 of 75 items from Greek-speaking 3-year-olds, and 17 of 119 items from Greek-speaking 4-year-olds. Across languages, 69% of these labels referred to vehicles and 7% to goal objects; the remaining 24% referred to agents or to background elements. As mentioned in "Participants," in addition to these individual trials, three additional children (all Greek-speaking 4-year-olds) were excluded from the analysis for producing a majority of event descriptions of this type.

3 These questions about the cartoons were posed during a memory task that is not described in this paper because it is not relevant to the questions under investigation. Every child participated in the memory task after he or she had completed one of the Linguistic or Nonlinguistic tasks described in the text. Children in the Nonlinguistic task were informed in advance about the "memory game" to motivate them to pay attention to the stimuli when these were first presented and were asked during the task to indicate whether each of a new set of dynamic events was the "same" or "different" as the events they had seen before. Based on prior evidence (e.g., Hagen & Kingsley 1968; Reese 1975; Hitch et al. 1991; Flavell et al. 1996; Palmer 2000; Kahn & Snedeker 2010), we thought it was unlikely that young children would use language strategically to encode stimuli in a memory task. There were no significant differences across language groups and task for memory accuracy for either event component (all p values < 0.10).

2.5.2 Analysis of eye movement data

Eye movement data were analyzed to assess the effects of language background, age, and task on encoding of motion event components. Data samples from target trials (50 per second from the Tobii 1750, 60 per second from the Tobii T60) were time-locked to the onset of the video, and analyses were performed on raw eyegaze coordinates from each sample. Trackloss was determined separately for each eye by Tobii's eyetracking software (Clearview for the 1750, Tobii Studio for the T60). Our data set includes samples for which the system is certain that it has recorded the correct coordinates for at least one eye (i.e., samples with a validity score of 0 or 1 on a scale from 0 to 4). Missing data (samples with validity >1) were counted as trackloss for a given eye. For samples with available data from both eyes, we used an average of the gaze coordinates from the two eyes. Trials with global track-loss of $>30\%$ were excluded from the analysis ($n = 22$ from the Linguistic task, $n = 20$ from the Nonlinguistic task). Four-year-old participants with more than four excluded target trials ($n = 2$) and three-year-old participants with more than three excluded target trials ($n = 3$) were replaced in the design.

To assess attention to motion event information in our stimuli, two dynamic spatial scoring regions were defined for each target video: a Manner region, which included the instrument used by the agent as the means of motion (e.g., the sailboat), and a Path endpoint region, which included the stationary path endpoint (e.g., the island). The Manner region never included the head or torso of the agent; these visual elements were included in an additional Agent scoring region that is not reported here. Trajectories were omitted from the Path region because they were never visible in our stimulus events, and previous work in our labs has demonstrated that although viewers of motion events like these do make anticipatory eye movements that project an agent's trajectory toward a visible path endpoint, they rarely fixate empty regions of space (Papafragou et al. 2008). On average, Manner regions subtended 6.90 (width) $\times$ 2.67 (height) deg visual angle, and Path regions subtended 9.22 $\times$ 8.98 deg visual angle. The size of each region for each stimulus is given in Appendix B (cf. also Bunger et al. 2012).

During the animation in our motion event videos, the instrument moved across the screen toward the path endpoint. To keep track of looks to this dynamic event component, an automated data analysis procedure was used to update the coordinates of the Manner region in the eyetracking analysis file as the event unfolded. Manner and Path regions were first defined by hand based on the position of instruments and path endpoints in the first frame of each target video. The Manner region was then repositioned by hand for each frame of the video, and the coordinates of this region in each successive frame were recorded to a file. The size of the Manner region remained constant across frames, as did the size and position of the Path region. For the analysis, an eyegaze sample was defined as being within a region of interest if its coordinates fell within the region as defined for the corresponding video frame. As instruments moved toward path endpoints near the end of events, Manner regions sometimes partially occluded Path regions. Overlap of this sort was resolved by assigning gaze to the Manner region, a step that follows directly from our choice to code looks only to items that were visible in the stimuli in a given frame. Eyegaze data are reported as the proportion of samples (averaged across subjects) for looks within these predefined regions of interest (out of all looking), averaged into blocks of 1 second. Any looks within a region were included in the analysis, regardless of duration.

2.5.3 Statistical analyses

Multilevel mixed logit modeling with crossed random variables for Subjects and Items was used to assess the reliability of trends observed in the data (cf. Baayen et al. 2008; Barr et al. 2013). Eyegaze data (proportions of samples whose coordinates match those of a given region of interest) in statistical analyses were elogit-transformed following Barr (2008). Best fitting lmer models for each analysis were chosen through stepwise comparisons of log likelihood values. Fixed factors (Language, Age, Task, Motion Component, as appropriate) were included as random slopes in Item effects structures when they did not perfectly correlate with the intercept. All p values reported for factors within analyses are vs. an empty model with no fixed effects.

3 Results and discussion

3.1 Event descriptions

Table 1 provides information about the proportion of utterances in which the preschoolers in this study mentioned the Manner or Path of our target motion events (or neither), regardless of the syntactic position in which those event components were encoded.[4] Across age groups, English-speaking children were more likely to mention either motion event component than Greek-speaking children were, and across language groups, older children were more likely to mention either motion event component than younger children were. Across language and age groups, children were more likely to provide information about Manners of motion than about Paths of motion. These trends were confirmed by multi-level modeling of categorical values at the trial-level for mention of motion information (0,1), with Language (English, Greek), Age (3yo, 4yo), and Motion component (Manner, Path) as first-level fixed factors. The best fitting model ($p < 0.001$; *Table 2*) includes main effects of Language, Age, and Motion component, as well as an interaction between Language and Motion Component. The significant interaction between Language and Motion Component is representative of the fact that English-speaking children were significantly more likely than Greek-speaking children to mention Manners ($p < 0.001$), but the two groups were equally likely to mention Paths ($p = 0.86$).[5]

	Manner	Path
English		
3-year-olds	0.71 (±0.12)	0.21 (±0.06)
4-year-olds	0.91 (±0.04)	0.33 (±0.07)
Greek		
3-year-olds	0.53 (±0.13)	0.24 (±0.10)
4-year-olds	0.50 (±0.04)	0.33 (±0.07)

Table 1 Mean proportion (±SE) of motion event descriptions produced by each group of preschoolers that included information about Manners or Paths of motion.

Effect	Estimate	S.E.	z value	Pr(>|z|)	
Intercept	1.4963	0.4262	3.511	0.000447	***
Language: Greek vs. English	−1.8871	0.5106	−3.696	0.00219	***
Age: 4yo vs. 3yo	0.7883	0.3235	2.437	0.014825	*
Motion Component: Path vs. Manner	−3.0875	0.4678	−6.600	4.1e−11	***
Language × Motion Component	2.0248	0.6193	3.269	0.001078	**

Table 2 Fixed effects from best-fitting multilevel linear model of motion component mention. Formula in R: MotionInformation ~ Language × Motion Component + Age + (Motion Component | Subject) + (1 | Item). Significance values: * p < 0.05, ** p < 0.01, *** p < 0.001.

As shown in *Table 3*, even though children sometimes combined Manner with Path information in their target event descriptions, Manner information across languages and age groups mostly appeared in the absence of Path (see (5a) for an example from English and (5b) for an example from Greek).[6] Beyond this broad pattern, English-speaking children were almost twice as likely

4 Proportions for each group in Table 1 do not add up to 1 because some utterances included information about both event components, and some included information about neither. See Table 3 for a breakdown of the data that includes these details.

5 Across age and language groups, children were more likely to encode Manner information in verbs like "sailing" rather than in subject position (e.g., "man with a boat") or in post-verbal positions (e.g., "in a boat"). (Proportion of Manner in verbs: English-speaking 3-year-olds: 0.85; English-speaking 4-year-olds: 0.75; Greek-speaking 3-year-olds: 0.60; Greek-speaking 4-year-olds: 0.71). Multilevel mixed logit modeling on categorical values at the trial level for Manner location (Verb, Elsewhere) with Language (English, Greek) and Age (3yo, 4yo) as first-level fixed factors revealed no effects of Language or Age on the location of Manner encoding, and no interaction between Language and Age.

6 Table 1 presents gross information about the proportion of event descriptions that include either Manner or Path information. In Table 3, this information has been reorganized to communicate the full semantic content of the descriptions. So, for example, the values in the "Manner Only" and "Both" columns in Table 3 add up to the values in the "Manner" column in Table 1 (and likewise for the Path columns).

as Greek-speaking children to use Manner-only descriptions. Path-only descriptions were infrequent overall.

	Manner Only	Path Only	Both	Neither
English				
3-year-olds	0.60 (±0.12)	0.10 (±0.04)	0.12 (±0.05)	0.19 (±0.10)
4-year-olds	0.63 (±0.08)	0.06 (±0.03)	0.28 (±0.07)	0.04 (±0.02)
Greek				
3-year-olds	0.38 (±0.12)	0.09 (±0.04)	0.16 (±0.08)	0.38 (±0.14)
4-year-olds	0.38 (±0.04)	0.18 (±0.05)	0.16 (±0.04)	0.29 (±0.07)

Table 3 Mean proportion (±SE) of motion descriptions of target motion events. This table is drawn from the same data presented in Table 1, reorganized in terms of how motion components are distributed across each event description.

(5) a. He sailing it.

b. Anthropos odigai to karavi.
Human.NOM drives the.ACC boat.ACC
'man drives the boat'

Overall, the language-specific pattern of omissions we observe in young learners of English and Greek is consistent with trends in adult production in the two languages: adult speakers of English tend to mention Manners of motion more often than adult speakers of Greek (Papafragou et al. 2002; 2006, among others). Unlike those prior cross-linguistic reports, however, Greek-speaking children also mentioned Manners more often than Paths. This finding is reminiscent of studies pointing out that Greek presents some variation in how frequently Manner is expressed in motion descriptions (e.g., Soroli 2012). Notice that the stimuli used in this experiment included vehicle-defined Manners of motion (e.g., steering a boat, driving a car, flying a plane). The movement of these vehicles was the only motion that occurred in the animated stimuli, and many of these vehicles were interesting themselves (parachutes, hot air balloons, sailboats, ice skates). In the next section we present eyegaze patterns that suggest that children found these vehicles particularly engaging. It is likely that this interest in the vehicles (or their dynamic motion) that defined Manners of motion in our stimuli led children in both language groups to talk about them. Critically, the English bias to talk about Manners of motion went beyond this baseline interest in features of our stimuli: despite an overall preference across language groups to talk about Manners, English-speaking children mentioned Manners of motion even more often than Greek-speaking children did.

In summary, we found both similarities and differences in the way preschool-aged speakers of English and Greek described our motion events. Not surprisingly, older children in both language groups tended to provide more motion information than younger children did (see Hickmann & Henriks 2006 and Hickmann et al. 2018 for similar developmental findings in English-, German-, and French-speaking children). Both English- and Greek-speaking children mentioned Manner information more often than Path information (cf. also Soroli 2012), a fact that may have been due to features of our dynamic stimuli. Consistent with adult trends in production, however, we found that across age groups, English-speaking children were more likely to mention Manners of motion than were Greek-speaking children. This finding indicates that, by the time they are 3 years old, children learning English and Greek are sensitive to the way adult speakers of their own language describe motion events and have already started to follow these patterns in their own language use. This conclusion is consistent with prior work on how children describe motion (e.g., Özçalişkan & Slobin 2000; Papafragou et al. 2002; Allen et al. 2007; Özyürek et al. 2008; Papafragou & Selimis 2010a) while recognizing some variation in how cross-linguistic differences are manifested (cf. also Selimis & Katis 2010; Soroli & Verkerk 2017).

3.2 Eye movements

Given these similarities and differences in the way English- and Greek-speaking preschoolers describe motion events, we next ask whether children of this age gather information from the visual world during speech planning in similar ways or in language-specific ways. We have collapsed across age groups in our assessment of these eyegaze patterns because, although older

children tended to say more about target events than younger children did, the kind of motion information they were providing did not change with development (i.e., everyone tended to talk more about Manners).

3.2.1 Attention to Manner

In a first analysis, we look at patterns of attention to Manner in our motion events for trials on which children mentioned or did not mention this specific component. We focus on Manner since mention of this component was the locus of a strong cross-linguistic difference (see ***Table 1***). To probe for eyegaze patterns that are specific to the process of language production, we compare attention to motion event components by participants who completed the Linguistic task to those who completed the Nonlinguistic task. We will discuss this data with respect to two research questions: First, do we see differences in the way children direct their attention when engaged in linguistic and nonlinguistic tasks? And second, do English- and Greek-speaking children show the same patterns of attention across these tasks?

Figure 2 depicts the attention that English- (***Figure 2A***) and Greek- (***Figure 2B***) speaking preschoolers directed to the Manner elements of our motion events. Because we are interested in the way speakers gather information in preparation for speaking, these graphs depict just 5 s of the total 9 s viewing period, including the 3 s before children were signaled to speak (by the beep) and 2 sec after this signal.

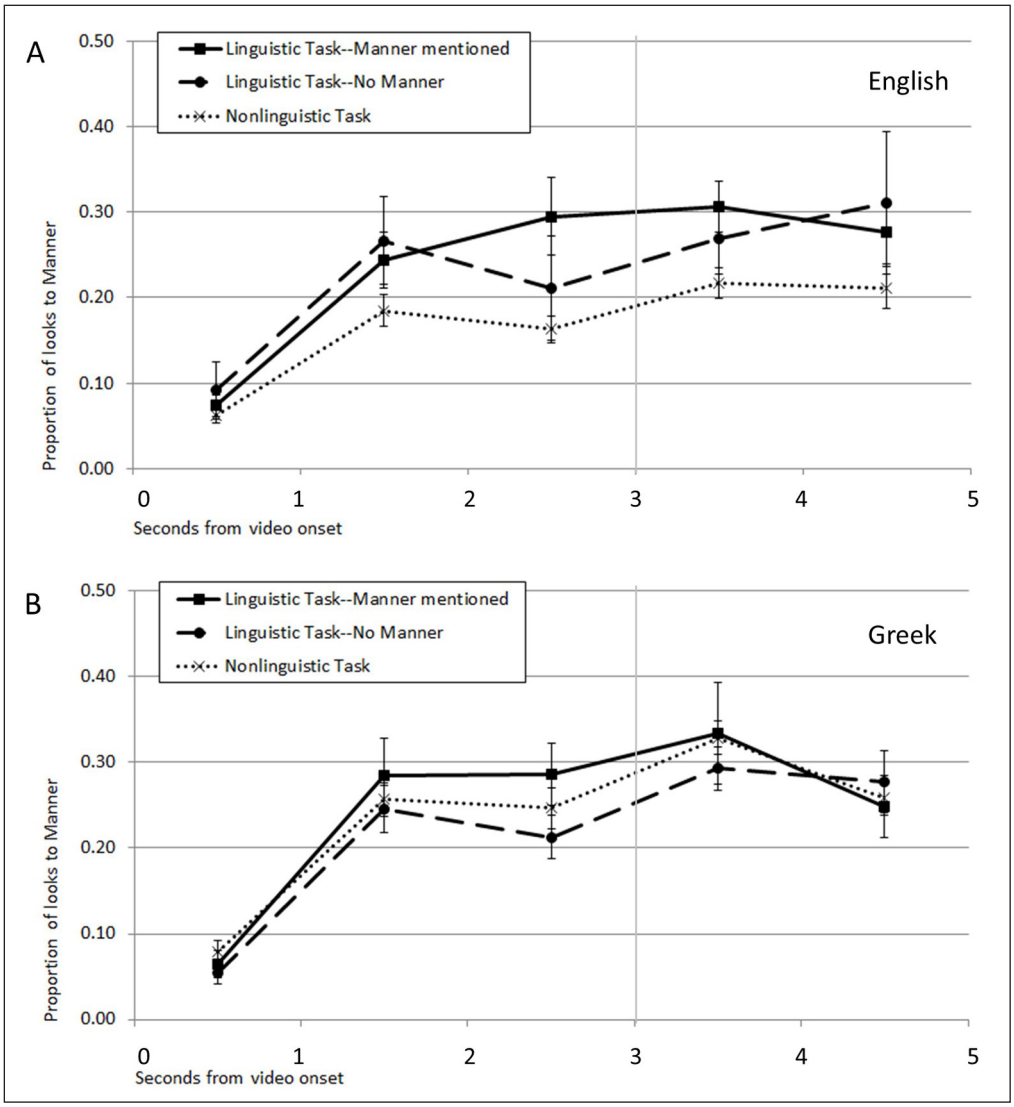

Figure 2 Proportion of looks to Manner regions of motion event stimuli by English **(A)** and Greek **(B)** speaking children in the Linguistic and Nonlinguistic tasks. Data for the Linguistic task are divided by whether or not Manner information was provided in the event description for that trial. Proportions are calculated based on looks to the entire target image. The vertical line at 3 s marks the point at which the animated motion stopped and the beep sounded.

We used multilevel mixed elogit modeling as described above to compare patterns of attention to Manner elements in our stimuli across tasks and language groups. Elogit-transformed proportions of looks to Manner regions were modeled separately within five 1-s windows beginning at stimulus onset, and Language (English, Greek) and Task (Nonlinguistic, Linguistic)

were entered as first-level fixed factors. We will refer to these analysis windows by their start and end times: thus, we analyzed data for the following five 1-s blocks of time: the 0–1 s window, the 1–2 s window, the 2–3 s window, the 3–4 s window, and the 4–5 s window. Data for the Linguistic task were assessed separately for trials on which Manner had (*Table 4*) and had not (*Table 5*) been mentioned.

Effect	Estimate	S.E.	t-value
1–2 s window			
Intercept	−2.48	0.30	−8.30
Task: Nonlinguistic vs. Linguistic	0.42	0.18	2.26
2–3 s window			
Intercept	−2.76	0.39	−7.00
Language: Greek vs. English	0.69	0.26	2.65
Task: Nonlinguistic vs. Linguistic	1.02	0.28	3.62
Task × Language	−0.88	0.43	−2.05
3–4 s window			
Intercept	−2.04	.29	−7.02
Language: Greek vs. English	0.67	0.22	3.09

Table 4 Fixed effects from best-fitting multilevel linear model of attention to Manner regions by time window for the Nonlinguistic task vs. Linguistic task trials with Manner mention. Models are provided only for windows in which an empty model with no fixed factors did not provide the best fit. The models presented are the best fitting models for each time window; when effects or interactions do not appear, it is because adding them to the models did not reliably improve the fit. Formulas in R: 1–2 s: MannerLooks ~ Task + (1 | Subject) + (Language | Item); 2–3 s: MannerLooks ~ Language × Task + (1 | Subject) + (1 | Item); 3–4 s: MannerLooks ~ Language + (1 | Subject) + (1 | Item).

Effect	Estimate	S.E.	t-value
2–3 s window			
Intercept	−2.73	0.39	−6.99
Language: Greek vs. English	0.54	0.20	2.70
3–4 s window			
Intercept	−2.73	0.39	−6.99
Language: Greek vs. English	0.54	0.20	2.70
4–5 s window			
Intercept	−2.37	0.29	−8.17
Language: Greek vs. English	0.45	0.20	2.26

Table 5 Fixed effects from best-fitting multilevel linear model of attention to Manner regions by time window for the Nonlinguistic task vs. Linguistic trials with no Manner mention. Models are provided only for windows in which an empty model with no fixed factors did not provide the best fit. The models presented are the best fitting models for each time window; when effects or interactions do not appear, it is because adding them to the models did not reliably improve the fit. Formula in R for all models: MannerLooks ~ Language + (1 | Subject) + (1 | Item).

When assessing attention to Manner regions for trials in the Linguistic task on which participants *did* mention Manner information (*Table 4*), we found effects of both Task and Language. A significant effect of Task was found for the 1–2 s analysis window ($p < 0.001$), such that during this time period children in the Linguistic task who went on to mention Manner information directed more attention to Manner regions than children in the Nonlinguistic task, regardless of language background. A significant interaction between Task and Language was found for the 2–3 s analysis window ($p < 0.05$): In this analysis window, only English-speaking children who went on to mention Manners showed a significant increase in attention to Manners in the Linguistic task vs. the Nonlinguistic task ($p < 0.05$). Additionally, Greek-speaking children in the Nonlinguistic task directed more attention to Manners during this window compared to English-speaking children ($p < 0.05$). Finally, a significant effect of Language was found for the 3–4 s analysis window ($p < 0.01$), such that Greek children directed significantly more attention to Manner regions than English-speaking children did. No effects of Task or Language, or interactions between them, were found for the other two analysis windows.[7]

7 Although visual examination of the eyegaze data presented in Figure 2 may suggest that there should be task effects for Greek-speaking children in the 2–3 s analysis window, this pattern does not reach statistical significance in the elogit-transformed data on which the analyses were performed, perhaps because variance in the 2-3s window is greater than that in the 1-2s window.

When assessing attention to Manner regions for trials in the Linguistic task on which participants *did not* mention Manner information (***Table 5***), a significant effect of Language was found for the 2–3 s, 3–4 s, and 4–5 s analysis windows (all $p < 0.05$). In all cases, Greek-speaking children were directing more attention to Manner regions than were English-speaking children. No effects of or interactions with Task were found in these windows, and no effects of Task or Language, or interactions between them, were found for the other two analysis windows.

To return to the questions we set out at the beginning of this section, these data do show differences in the way children direct their attention to dynamic motion events when engaged in linguistic and nonlinguistic tasks, with both similarities and differences across language groups. First, we found that by the time they are 3 years old, children, like adults, have begun to direct their eyegaze during the process of language production in ways that are linked to what they are planning to talk about. Specifically, our results demonstrate that when they were planning event descriptions that included Manner information, preschool-aged speakers of both languages devoted more attention to Manners of motion in the visual world than they did in a Nonlinguistic task. This increase in attention to Manners while planning event descriptions that included Manner information is consistent with a strategy in which children directed more attention to event components that they were planning to talk about. Critically, the increase in attention to Manners that we observed began within the second second of event viewing, i.e., as children were planning their event descriptions and before they had actually begun to produce them. There was no equivalent increase in attention to Manner information for trials on which children in the Linguistic task did not mention Manners in their motion event descriptions.

Moreover, we found that this increase in attention to Manner regions in the Linguistic task was less consistent for Greek-speaking children than it was for English-speaking children. This finding may be due to the fact that Greek-speaking children were already directing a considerable amount of attention to Manner regions, as demonstrated by their high level of attention to Manner regions even in the Nonlinguistic task. That is, if the interest that Greek-speaking children exhibited in Manner regions was already near ceiling, the process of planning to talk about those Manners might not have been able to boost attention to them beyond the baseline preference. If this is true, then it leaves open the possibility that Greek-speaking children in the Linguistic task were led to mention Manner information more than in prior reports (e.g., Papafragou et al. 2003, 2006) because those event elements were salient to them. We return to this finding in the General Discussion.

3.2.2 Attention to Manner over Path

In our first analysis, we focused on whether Manner was mentioned or omitted in event descriptions. In our second analysis, we pursue a more specific link between utterance content and attention allocation: we ask whether the relative attention allocated to Manner over Path within the Linguistic task changed depending on whether children encoded Manner exclusively or not. For this analysis, we compare trials on which children offered only Manner information (by far the most prevalent option in both languages, and more prevalent in English than in Greek) to trials for which they offered combinations of Manner and Path information (see ***Table 3***). As in our previous analysis, we expect to see differences in the way children gather information from our motion events that are consistent with differences in the linguistic encoding biases in each language.

Figure 3 depicts the way English- (***Figure 3A***) and Greek- (***Figure 3B***) speaking preschoolers directed their attention to motion event components in our events, split by the context in which Manner information was given (i.e., alone, or in conjunction with Path information). As before, the graphs depict just the 3 s before children were signaled to speak (by the beep) and 2 s after this signal. We used multilevel mixed elogit modeling as described above to compare patterns of attention to motion event components in our stimuli across language groups and types of event descriptions. Difference scores were calculated for each trial in the Linguistic task on which Manner information had been mentioned, whether alone or in conjunction with Path information by subtracting elogit-transformed proportions of looks to Path regions from elogit-transformed proportions of looks to Manner regions in five 1-s windows beginning at stimulus onset. As in our description of ***Figure 2***, we will refer to these analysis windows by their start and end times: thus, we analyzed data for the following five 1-s blocks of time: the 0–1 s window, the 1–2 s window, the 2–3 s window, the 3–4 s window, and the 4–5 s window. Difference scores

were modeled separately within each 1-s analysis window, with Language (English, Greek) and Motion Information (Manner, Manner + Path) entered as first-level fixed factors.

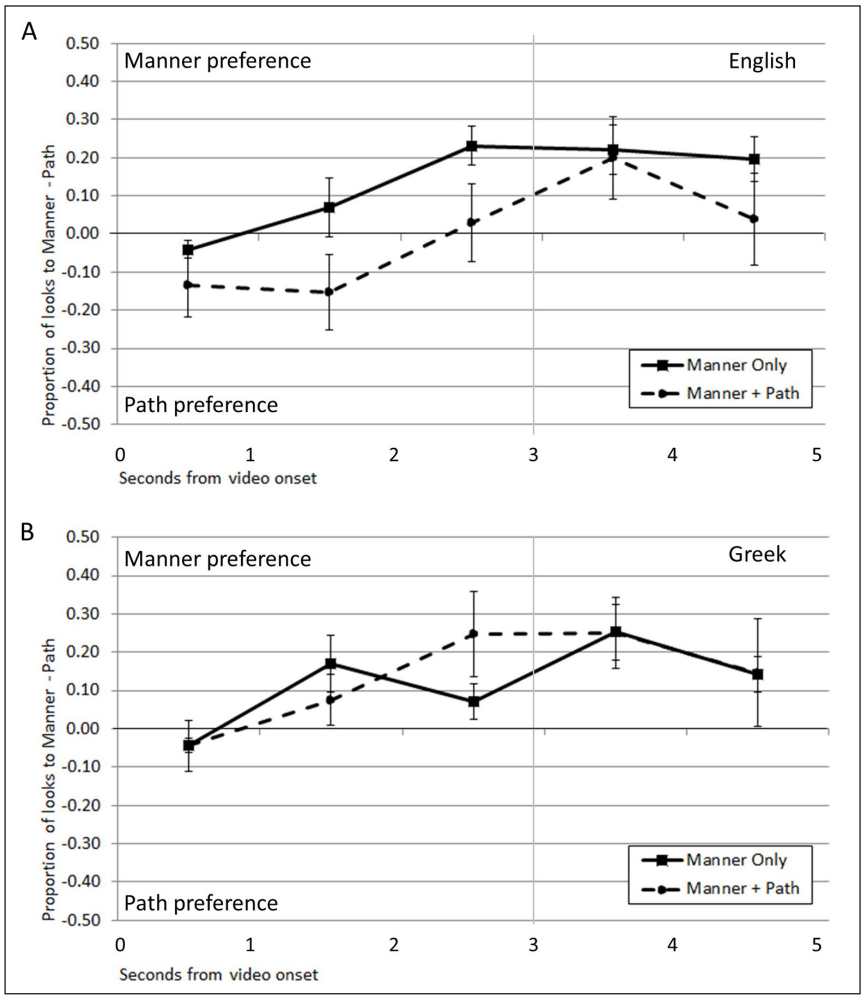

Figure 3 Average proportion of looks to motion event regions by English **(A)** and Greek **(B)** speaking children in the Linguistic task for trials on which Manner information was given. Data are divided by whether Manner information was given alone or in conjunction with Path information. Proportions are calculated based on looks to the entire target image. The vertical line at 3 s marks the point at which the animated motion stopped and the beep sounded. Positive difference scores indicate a preference to look at Manner information; negative difference scores indicate a preference to look at Path information.

For this analysis, we found effects of both Language and Motion Information on attention to motion event components (*Table 6*). A significant effect of Language was found for the 1–2 s analysis window ($p < 0.05$): Greek-speaking children demonstrated a stronger preference for Manner regions over Path regions than English-speaking children did, regardless of the type of event description they were preparing. Additionally, a significant interaction between Language and Motion Information was found for the 2–3 s ($p < 0.01$) analysis window. In this analysis window, when English-speaking children mentioned only Manner information, they showed a preference to look at Manners that was significantly greater than that of Greek-speaking children who mentioned only Manners ($p < 0.01$). In addition, when English-speaking children mentioned both Manner and Path information, their preference for Manner regions was significantly lower than that shown by English-speaking children who mentioned only Manner information ($p < 0.05$); this pattern did not hold for Greek-speaking children. No effects of Motion Information or Language, or interactions between them, were found for the analysis windows not described.

Effect	Estimate	S.E.	t-value
1–2 s window			
Intercept	0.21	0.63	0.33
Language: Greek vs. English	1.15	0.55	2.08
2–3 s window			
Intercept	3.43	0.98	3.51
Motion Information: Manner vs. Manner + Path	−1.67	0.64	−2.61
Language: Greek vs. English	−4.26	1.35	−3.16
Motion Information × Language Interaction	2.85	1.00	2.86

Table 6 Fixed effects from best-fitting multilevel linear model of attention to motion event regions by time window for trials on which Manner information was given. Models are provided only for windows in which an empty model with no fixed factors did not provide the best fit. The models presented are the best fitting models for each time window; when effects or interactions do not appear, it is because adding them to the models did not reliably improve the fit. Formulas in R: 1–2 s window: Eyegaze ~ Language + (1 | Subject) + (1 | Item), 2–3 s window Eyegaze ~ Language * Motion Information + (1 | Subject) + (1 | Item).

This pattern of results reveals significant similarities in the relative attention that children paid to Manner and Path information in our motion events as they planned event descriptions but also two language-specific differences. In early stages of event apprehension and sentence planning (second analysis window), Greek-speaking children demonstrated a preference to look at Manner regions that exceeded that shown by English-speaking children regardless of whether learners planned to mention Manner alone or a combination of Manner and Path information. As mentioned previously, this overall preference for Manners in Greek learners may be related to apparent interest in the dynamic vehicles depicted in our stimuli. In later stages of event apprehension and sentence planning (third analysis window), English-speaking children who mentioned only Manner information allocated more attention to Manner regions compared to Greek-speaking children who mentioned Manners exclusively. Furthermore, English-speaking children were more likely to shift their attention toward Path regions when planning to mention both Manner and Path compared to cases in which only Manner was mentioned, unlike their Greek-speaking peers who overall attended primarily to Manner. This pattern shows a tighter coupling between sentence content and attention allocation in English compared to Greek learners that reflects the very stable bias in motion encoding observed in adult English speakers. The presence of a more diffuse pattern in Greek learners is the result of an overall bias to attend to Manner (perhaps also coupled with some flexibility in motion lexicalization preferences; Selimis & Katis 2010; Soroli & Verkerk 2017).

4 General discussion

In this study, we used a combination of linguistic and online methods to investigate the way preschool-aged speakers of English and Greek inspect and describe dynamic motion events. Our goal in investigating the way children of this age describe motion events was to investigate how early children begin to exhibit the kind of language-specific biases in the encoding of motion event information that have been previously reported for adult speakers of these languages. Specifically, we asked whether preschoolers' tendency to mention Manner and Path information when describing motion events mirrors that of adult speakers of their language. Additionally, we used eyetracking to carry out a novel investigation of the way "thinking for speaking" operates across young speakers of different languages. More specifically, we asked whether preschool-aged speakers of English and Greek, like adult speakers of these languages (e.g., Papafragou et al. 2008), exhibit language-specific patterns of event inspection when they are involved in the process of selecting motion information to talk about.

Our assessment of children's event descriptions confirms that by the time they are 3 years old, children learning English and Greek are already beginning to show differences in the way they prioritize motion event information in event descriptions. Specifically, we saw that, consistent with adult patterns, English-speaking preschoolers mentioned the Manners of our target motion events more often than Greek-speaking preschoolers did. Additionally, we found that older children tended to provide more information about our motion events than younger children did. These findings are consistent with crosslinguistic data on motion event encoding in very young children (e.g., Özçalişkan & Slobin 2000; Papafragou et al. 2002; Özyürek et al. 2008; Papafragou & Selimis 2010a), and are reminiscent of recent crosslinguistic work on the description and inspection of complex causative events (Bunger et al. 2016). We also found that children from both language groups were more likely to provide Manner information about our motion events than Path information. This finding points to the fact that motion-typological patterns apply with some flexibility within individual languages, and is consistent with evidence that Greek speakers sometimes uses satellite-framed motion encoding (Soroli & Verkerk 2017; cf. also Selimis & Katis 2010). As mentioned previously, we think that this crosslinguistic preference to talk about Manners had to do with properties of our stimuli that included interesting, dynamic, and visually salient means of transportation such as parachutes, sailboats, planes and hot air balloons (see Appendix A).[8] The amount of attention that children directed to Manner regions of our stimuli supports this conclusion: across language groups, children devoted a considerable amount of attention to Manner regions even when they were not preparing descriptions of the events.

8 Future work in this area could focus on types of gait rather than vehicle-driven manners of motion to go beyond such interesting manners of motion. Soroli and Hickmann (2010) describe one way that areas of attention to gait and path might be distinguished.

Additionally, we found that young speakers of English and Greek demonstrated subtle differences in the way they inspected motion events while preparing to talk about them. Specifically, English-speaking children demonstrated a high attentional preference for the Manner regions of our events while they were preparing event descriptions that included Manner information. When they were preparing event descriptions that also included Path information, they shifted this attentional preference in the direction of the Path regions. Greek speaking children, on the other hand, did not demonstrate differences in their attention to motion regions as they planned different kinds of motion event descriptions: regardless of the motion information they planned to mention, they demonstrated a consistent preference to attend to the Manners of those events. We have suggested that both of these patterns are consistent with a sensitivity toward adult-like patterns of motion event description in each language. English speakers are more likely to gather Manner information from a motion event while planning to talk about it (most likely in a verb), and so when they also plan to mention Path information, their relative preference for Manner over Path decreases. Greek speakers, on the other hand, are less likely to mention Manner information (even though they sometimes present a more balanced typological pattern with respect to the description of motion events; Selimis & Katis 2010; Soroli 2012), and Greek-speaking children in this study did not show differences in their relative preference for Manner and Path elements as they planned event descriptions with these different kinds of information.

Finally, we found that young speakers of English and Greek, like adult speakers of their languages, direct their attention to motion events in different ways when they are preparing to talk about them compared to their attention patterns while inspecting the same events in a nonlinguistic task. Specifically, we found that children from both language groups increased their attention to the Manners of our motion events when they were planning sentences that included Manner information compared to the attention they paid to these regions when viewing them in preparation for a memory task. Previous work has demonstrated this pattern of thinking for speaking in 4-year-old speakers of English (Bunger et al. 2012); here we extend it not only to younger speakers but also to speakers of a language that is typologically different from English. Again, these findings suggest that the emerging linguistic biases that we saw in the preschoolers in this study have already begun to have online effects on the way they gather information from motion events.

It remains to be determined *why* Greek-speaking children showed a stronger preference for Manners in the nonlinguistic task than English-speaking children did.[9] One possibility is that this preference is related to the way Greek speakers tend to encode motion information in language. Papafragou and colleagues (2008) reported that, when they were asked to remember a motion event in a Nonlinguistic task, adult speakers towards the end of each trial directed more attention to motion components that were not encoded in verbs in their language: English speakers to Path endpoints and Greek speakers to Manners of motion. Later work (Trueswell & Papafragou 2010) suggested that this eyegaze pattern was due to a potential for covert linguistic encoding during the Nonlinguistic task: when adult speakers had access to linguistic processing resources, they (silently) encoded information in language to support memory. It is, therefore, conceivable that Greek-speaking children covertly encoded Manners using language in the current nonlinguistic task in preparation for the upcoming memory test. This possibility is not fully satisfying: as argued earlier, covert linguistic encoding is unlikely in children of this age because of limitations on the use of language to support memory before the age of 6 or 7 years (e.g., Hitch et al. 1991; Palmer 2000; Kahn & Snedeker 2010; see footnote 3). An alternative possibility is that, despite the finding that our Manners were equally familiar to adult speakers of the two languages (see Materials section), there may still have been a difference in the degree of exposure to various vehicles between the two groups of children. To probe this further and to increase the generalizability of these findings, future studies in this area could include a wider range of manners and paths of motion, as well as additional languages that clearly show verb-framed vs. satellite-framed preferences.

9 Notice that, although Greek learners directed their attention to Manner more than English learners did, they included Manner expressions less often than English learners did in their descriptions of motion events. This suggests that what learners say in production is only partly determined by the nonlinguistic salience of components of a scene or event (for a similar point, see Bunger et al. 2012).

5 Conclusion

This study breaks new ground by demonstrating that children as young as 3 years of age demonstrate crosslinguistic patterns of eyegaze during the process of language production that are consistent with Slobin's thinking for speaking hypothesis (Slobin 1996b; 2006). As reviewed in the Introduction, previous studies have demonstrated that adult speakers of English and Greek prioritize Manner and Path information differently when describing motion events, and that they also direct their attention to motion events in language-specific and task-specific ways. Here, we expand our understanding of the development of these cross-linguistic patterns by demonstrating that 3- and 4-year-old speakers of these languages have already begun to prioritize information about motion events like adult speakers of their languages do when describing such events, that they demonstrate language-specific patterns of event inspection as they plan those descriptions (by directing their attention by and large to things that they plan to talk about), and that they direct their attention to motion events depending on whether the task is linguistic (language production) or nonlinguistic (memory). Together these results enrich our knowledge of how sentence production proceeds in speakers of different languages (Levelt 1989) and illustrate the rich processes that allow children to transform their thoughts into utterances.

Appendix A

A1. List of motion events that were used as target stimuli

1. An alien drives a car to the mouth of a cave
2. A man in a hot air balloon lands on top of a building
3. A man in a sailboat lands on an island
4. A man paddles a canoe to a dock
5. A woman on a magic carpet lands on the moon
6. A man drives a motorcycle to a carwash
7. A man parachutes from the sky and lands on a tree
8. A man lands an airplane on a platform
9. A boy roller skates to a soccer net
10. A girl rides a scooter to the mouth of a cave
11. A duck ice skates to a fishing hut
12. A skier skis to a finish line

A2. List of motion events that were used as filler stimuli

1. A frog takes one hop away from a tree
2. A couple dances in a circle next to a tree
3. A person pushes a ball across a pool table
4. A child rakes some leaves
5. A turtle swims across the screen carrying an apple on its back
6. A bird flies across the sky carrying a package by a string from its feet
7. A person closes the lid of a box
8. A child pulls a kite down from the sky toward him
9. An elf walks across the screen
10. A basketball player tosses a ball up and catches it
11. A person carries a full trash can across the screen away from a house
12. A child pushes a snowball down a hill

Appendix B

Size of Manner and Path regions for each target event, given in degrees of visual angle at a viewing distance of 60 cm (cf. Table A1 in Bunger et al. 2012).

Target event	Manner region (deg)	Path region (deg)
An alien drives a car into the mouth of a cave	car	cave
	6.31 × 2.26	12.02 × 11.21
A man in a hot air balloon lands on top of a building	hot air balloon	building
	2.69 × 2.59	14.05 × 10.02
A man in a sailboat lands on an island	sailboat	island
	8.67 × 2.40	8.92 × 3.86
A man paddles a canoe to a dock	canoe	dock
	7.12 × 2.86	9.17 × 6.12
A woman on a magic carpet lands on the moon	magic carpet	moon
	8.17 × 2.53	5.62 × 5.39
A man drives a motorcycle into a carwash	motorcycle	carwash
	8.67 × 3.66	12.08 × 15.54
A man parachutes from the sky and lands on a tree	parachute	tree
	4.50 × 2.20	5.43 × 5.05
A man lands an airplane on a platform	plane	platform
	9.11 × 3.53	7.99 × 3.26
A boy roller skates into a soccer net	roller skates	soccer net
	5.87 × 3.59	9.97 × 10.15
A girl rides a scooter into the mouth of a cave	scooter	cave
	6.93 × 2.06	13.87 × 14.95
A duck ice skates into a fishing hut	ice skates	fishing hut
	5.74 × 2.79	6.74 × 12.53
A skier skis through a finish line	skis	finish line
	9.04 × 1.53	4.81 × 9.69

Abbreviations

ACC = accusative, NOM = nominative.

Funding information

This work was supported in part by the Eunice Kennedy Shriver National Institute of Child Health and Human Development under grant #R01HD055498.

Competing interests

The authors have no competing interests to declare.

Author affiliations

Ann Bunger *orcid.org/0000-0002-6080-3927*
Department of Linguistics, Indiana University, US

Dimitrios Skordos orcid.org/0000-0001-9056-5141
SLLLC, University of Calgary, CA
John C. Trueswell
Department of Psychology, University of Pennsylvania, US
Anna Papafragou orcid.org/0000-0001-5435-1058
Department of Linguistics, University of Pennsylvania, US

References

Allen, Shanley, Asli Özyürek, Sotaro Kita, Amanda Brown, Reyhan Furman, Tomoko Ishizuka & Mihoko 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. DOI: *https://doi.org/10.1016/j.cognition.2005.12.006*

Aske, Jon. 1989. Path predicates in English and Spanish. *Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society* 15. 1–14. DOI: *https://doi.org/10.3765/bls.v15i0.1753*

Baayen, R. H., D. J. Davidson & D. M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59. 390–412. DOI: *https://doi.org/10.1016/j.jml.2007.12.005*

Barr, Dale J. 2008. Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language* 59. 457–474. DOI: *https://doi.org/10.1016/j.jml.2007.09.002*

Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structures for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68. 255–278. DOI: *https://doi.org/10.1016/j.jml.2012.11.001*

Beavers, John, Beth Levin & Shiao Wei Tham. 2010. The typology of motion events revisited. *Journal of Linguistics* 46. 331–377. DOI: *https://doi.org/10.1017/S0022226709990272*

Bock, Kathryn, David E. Irwin & Douglas J. Davidson. 2004. Putting first things first. In John M. Henderson & Fernanda Ferreira (eds.), *The interface of language, vision, and action: Eye movements and the visual world*, 224–250. New York: Psychology Press.

Bunger, Ann, Dimitrios Skordos, John C. Trueswell & Ann Papafragou. 2016. How adults and children encode causative events cross-linguistically: Implications for language production and attention. *Language, Cognition and Neuroscience* 31. 1015–1037. DOI: *https://doi.org/10.1080/23273798.2016.1175649*

Bunger, Ann, John C. Trueswell & Anna Papafragou. 2012. The relation between event apprehension and utterance formulation in children: Evidence from linguistic omissions. *Cognition* 122. 135–149. DOI: *https://doi.org/10.1016/j.cognition.2011.10.002*

Choi, Soonja & Melissa Bowerman. 1991. Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition* 41. 83–121. DOI: *https://doi.org/10.1016/0010-0277(91)90033-Z*

Fenson, Larry, Philip S. Dale, J. Steven Reznick, Elizabeth Bates, Donna J. Thal, Stephen J. Pethick, Michael Tomasello, Carolyn B. Mervis & Joan Stiles. 1994. Variability in early communicative development. *Monographs of the Society for Research in Child Development* 59 (5, Serial No. 242). DOI: *https://doi.org/10.2307/1166093*

Flavell, John H., David R. Beach, & Jack M. Chinsky. 1966. Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development* 37(2). 283–299. DOI: *https://doi.org/10.2307/1126804*

Georgakopoulos, Thanasis, Holden Härtl & Athina Sioupi. 2019. Goal realization: An empirically based comparison between English, German and Greek. *Languages in Contrast* 19(2). 280–309. DOI: *https://doi.org/10.1075/lic.17010.geo*

Göksun, Tilbe, Kathy Hirsh-Pasek & Roberta Michnick Golinkoff. 2017. Trading spaces: Carving up events for learning language. *Perspectives on Psychological Science* 5. 33–42. DOI: *https://doi.org/10.1177/1745691609356783*

Griffin, Zenzi M. 2004. Why look? Reasons for eye movements related to language production. In John M. Henderson & Fernanda Ferreira (eds.), *The interface of language, vision, and action: Eye movements and the visual world*, 192–222. New York: Psychology Press.

Griffin, Zenzi M. & Kathryn Bock. 2000. What the eyes say about speaking. *Psychological Science* 11. 274–279. DOI: *https://doi.org/10.1111/1467-9280.00255*

Hagen, John W. & Phillip R. Kingsley 1968. Labeling effects in short-term memory. *Child Development* 39. 113–121. DOI: *https://doi.org/10.2307/1127363*

Hickmann, Maya, & Henriëtte Hendriks. 2006. Static and dynamic location in French and in English. *First Language* 26. 103–105. DOI: *https://doi.org/10.1177/0142723706060743*

Hickmann, Maya, Henriëtte Hendriks, Anne-Katharina Harr & Philippe Bonnet. 2018. Caused motion across child languages: a comparison of English, German, and French. *Journal of Child Language* 45. 1247–1274. DOI: *https://doi.org/10.1017/S0305000918000168*

Hickmann, Maya, Pierre Taranne & Philippe Bonnet, 2009. Motion in first language acquisition: Manner and Path in French and English child language. *Journal of Child Language* 36. 705–741. DOI: *https://doi.org/10.1017/S0305000908009215*

Hitch, Graham J., M. Sebastian Halliday, Alma M. Schaafstal & Thomas M. Heffernan. 1991. Speech, "inner speech," and the development of short-term memory: Effects of picture-labeling on recall. *Journal of Experimental Child Psychology* 51. 220–234. DOI: *https://doi.org/10.1016/0022-0965(91)90033-O*

Khan, Manizeh & Jesse Snedeker. 2010. Spontaneous implicit naming of visual objects. *Poster presented at the 35th Annual Boston University Conference on Language Development*, Boston, MA, November 5–7, 2010.

Lakusta, Laura & Barbara Landau. 2005. Starting at the end: The importance of goals in spatial language. *Cognition* 96. 1–33. DOI: *https://doi.org/10.1016/j.cognition.2004.03.009*

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

MacDonald, Maryellen C. 2013. How language production shapes language form and comprehension. *Frontiers in Psychology* 4. 1–16. DOI: *https://doi.org/10.3389/fpsyg.2013.00226*

Maguire, Mandy J., Kathy Hirsh-Pasek, Roberta Mitchnick Golinkoff, Mutsumi Imai, Etsuko Haryu, Sandra Vanegas, Hiroyuki Okada, Rachel Pulverman & Brenda 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. DOI: *https://doi.org/10.1016/j.cognition.2009.10.002*

Meyer, Antje S. 2004. The use of eye tracking in studies of sentence generation. In John M. Henderson & Fernanda Ferreira (eds.), *The interface of language, vision, and action: Eye movements and the visual world*, 173–190. New York: Psychology Press.

Naigles, Letitia R., Ann R. Eisenberg, Edward T. Kako, Melissa Highter & Nancy McGraw. 1998. Speaking of motion: Verb use in English and Spanish. *Language and Cognitive Processes* 13. 521–549. DOI: *https://doi.org/10.1080/016909698386429*

Norcliffe, Elisabeth, Agnieszka E. Konopka, Penelope Brown & Stephen C. Levinson. 2015. Word order affects the time course of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience* 30. 1187–1208. DOI: *https://doi.org/10.1080/23273798.2015.1006238*

Özçalişkan, Şeyda. 2013. Ways of crossing a spatial boundary in typologically distinct languages. *Applied Psycholinguistics* 36. 485–508. DOI: *https://doi.org/10.1017/S0142716413000325*

Özçalişkan, Şeyda, & Dan Issac Slobin. 2000. *Climb up* vs. *ascend climbing*: Lexicalization choices in expressing motion events with manner and path components. In S. Catherine Howell, Sarah A. Fish & Thea Keith-Lucas (eds.), *Proceedings of the 24th Annual Boston University Conference on Language Development*, 558–570. Somerville, MA: Cascadilla Press.

Özyürek, Asli, Sotaro Kita, Shanley Allen, Amanda Brown, Reyhan Furman & Tomoko Ishizuka. 2008. Development of cross-linguistic variation in speech and gesture: Motion events in English and Turkish. *Developmental Psychology* 44. 1040–1054. DOI: *https://doi.org/10.1037/0012-1649.44.4.1040*

Palmer, Sue. 2000. Working memory: A developmental study of phonological recoding. *Memory* 8. 179–193. DOI: *https://doi.org/10.1080/096582100387597*

Papafragou, Anna, Christine Massey & Lila Gleitman. 2002. Shake, rattle, 'n' roll: The representation of motion in thought and language. *Cognition* 84. 189–219. DOI: *https://doi.org/10.1016/S0010-0277(02)00046-X*

Papafragou, Anna, Christine Massey & Lila Gleitman. 2006. When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition* 98. B75–B87. DOI: *https://doi.org/10.1016/j.cognition.2005.05.005*

Papafragou, Anna, Justin Hulbert & John C. Trueswell. 2008. Does language guide event perception? Evidence from eye movements. *Cognition* 108. 155–184. DOI: *https://doi.org/10.1016/j.cognition.2008.02.007*

Papafragou, Anna, & Myrto Grigoroglou. 2019. The role of conceptualization in language production: Evidence from event encoding. *Language, Cognition and Neuroscience* 34. 1117–1128. DOI: *https://doi.org/10.1080/23273798.2019.1589540*

Papafragou, Anna & Stathis Selimis. 2010a. Event categorisation and language: A cross-linguistic study of motion. *Language and Cognitive Processes* 25. 224–260. DOI: *https://doi.org/10.1080/01690960903017000*

Papafragou, Anna & Stathis Selimis. 2010b. Lexical and structural biases in the acquisition of motion verbs. *Language Learning and Development* 6. 87–115. DOI: *https://doi.org/10.1080/15475440903352781*

Pruden, Shannon M., Kathy Hirsh-Pasek, Mandy J. Maguire & Meredith A. Meyer. 2004. Foundations of verb learning: Infants form categories of path and manner in motion events. In Alejna Brugos, Linnea Micciulla & Christine E. Smith (eds.), *Proceedings of the 28th annual Boston University Conference on Language Development*, 461–472. Somerville, MA: Cascadilla Press.

Pruden, Shannon M., Kathy Hirsh-Pasek & Roberta Michnick Golinkoff. 2008. Current events: How infants parse the world events for language. In Thomas F. Shipley & Jeffrey M. Zacks (eds.), Understanding events: From perception to action, 160–192. New York: Oxford University Press. DOI: *https://doi.org/10.1093/acprof:oso/9780195188370.003.0008*

Pulverman, Rachel, Jennifer L. Sootsman, Roberta Michnick Golinkoff & Kathy Hirsh-Pasek. 2003. The role of lexical knowledge in nonlinguistic event processing: English-speaking infants' attention to manner and path. In Barbara Beachley, Amanda Brown & Frances Conlin (eds.), *Proceedings of the 27th Annual Boston University Conference on Language Development*, 662–673. Somerville, MA: Cascadilla Press.

Pulverman, Rachel, Kathy Hirsh-Pasek, Roberta Michnick Golinkoff, Shannon Pruden & Sarah J. Salkind. 2006. Conceptual foundations for verb learning: Celebrating the event. In Kathryn A. Hirsh-Pasek & Roberta Michnick Golinkoff (eds.), *Action meets word: How children learn verbs*, 134–159. New York: Oxford University Press. DOI: *https://doi.org/10.1093/acprof:oso/9780195170009.003.0006*

Pulverman, Rachel & Roberta Michnick Golinkoff. 2004. Seven-month-olds' attention to potential verb referents in nonlinguistic events. In Alejna Brugos, Linnea Micciulla & Christine E. Smith (eds.), *Proceedings of the 28th annual Boston University Conference on Language Development*, 473–480. Somerville, MA: Cascadilla Press.

Pulverman, Rachel, Roberta Michnick Golinkoff, Kathy Hirsh-Pasek & Jennifer Sootsman Buresh. 2008. Infants discriminate manners and paths in non-linguistic dynamic events. *Cognition* 108. 825–830. DOI: *https://doi.org/10.1016/j.cognition.2008.04.009*

Reese, Hayne W. 1975. Verbal effects in children's visual recognition memory. *Child Development* 46(2). 400–407. DOI: *https://doi.org/10.2307/1128134*

Regier, Terry. 1996. *The human semantic potential: Spatial language and constrained connectionism.* Cambridge, MA: MIT Press. DOI: *https://doi.org/10.7551/mitpress/3608.001.0001*

Selimis, Stathis, & Demetra Katis. 2010. Motion Descriptions in English and Greek: A Cross-Typological Developmental Study of Conversations and Narratives. *Linguistik Online* 42: 57–76. DOI: *https://doi.org/10.13092/lo.42.421*

Skopeteas, Stavros. 2008. Encoding spatial relations: Language typology and diachronic change in Greek. STUF - *Language Typology and Universals* 61. 54–66. DOI: *https://doi.org/10.1524/stuf.2008.0006*

Skordos, Dimitrios, Ann Bunger, Catherine Richards, Stathis Selimis, John Trueswell, & Anna Papafragou. (2020). Motion verbs and memory for motion events. *Cognitive Neuropsychology* 37(5–6). 254–270. DOI: *https://doi.org/10.1080/02643294.2019.1685480*

Skordos, Dimitrios & Anna Papafragou. 2014. Lexical, syntactic, and semantic-geometric factors in the acquisition of motion predicates. *Developmental Psychology* 50. 1985–1998. DOI: *https://doi.org/10.1037/a0036970*

Slobin, Dan I. 1996a. Two ways to travel: Verbs of motion in English and Spanish. In Masayoshi Shibatani & Sandra Thompson (eds.), *Grammatical constructions: Their form and meaning*, 195–219. Oxford: Clarendon Press.

Slobin, Dan I. 1996b. From 'thought and language' to 'thinking for speaking'. In John J. Gumperz & Stephen C. Levinson (eds.), *Rethinking linguistic relativity*, 70–96. New York: Cambridge University Press.

Slobin, Dan. 2004. The many ways to search for a frog: Linguistic typology and the expression of motion events. In Sven Strömqvist & Ludo Verhoeven (eds.), *Relating events in narrative: Typological and contextual perspectives*, 219–257. Mahwah, NJ: Lawrence Erlbaum Associates.

Slobin, Dan I. 2006. What makes manner of motion salient? Explorations in linguistic typology, discourse, and cognition. In Maya Hickman & Stéphane Robert (eds.), *Space in languages: Linguistic systems and cognitive categories*, 59–81. Amsterdam: John Benjamins. DOI: *https://doi.org/10.1075/tsl.66.05slo*

Slobin, Dan I. & Nini Hoiting. 1994. Reference to movement in spoken and signed languages: Typological considerations. *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society: General Session Dedicated to the Contributions of Charles J. Fillmore* 20. 487–505. DOI: *https://doi.org/10.3765/bls.v20i1.1466*

Soroli, Eva. 2012. Variation in spatial language and cognition: exploring visuo-spatial thinking and speaking cross-linguistically. *Cognitive Processing* 13. 333–337. DOI: *https://doi.org/10.1007/s10339-012-0494-4*

Soroli, Eva & Maya Hickmann. 2010. Spatial language and cognition in French and English: Some evidence from eye-movements. In Giovanna Marotta, Alessandro Lenci, Linda Meini & Francesco Rovai (eds.), *Space in language: Proceedings of the Pisa International Conference*, 581–597. Firenze, Italy: Edizioni ETS.

Soroli, E. & Verkerk A. 2017. Motion events in Greek. *Cognitextes – Revue de l'Association Française de Linguistique Cognitive*, 1–54. DOI: *https://doi.org/10.4000/cognitextes.889*

Talmy, Leonard. 1975. Semantics and syntax of motion. In John P. Kimball (ed.), *Syntax and semantics*, vol. 4, 118–238. New York: Academic Press.

Talmy, Leonard. 1985. Lexicalization patterns: Semantic structure in lexical forms. In Timothy Shopen (ed.), *Language typology and syntactic description. Vol. 3: Grammatical categories and the lexicon*, 225–282. Cambridge: Cambridge University Press.

Talmy, Leonard. 1991. Path to realization: A typology of event conflation. *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on The Grammar of Event Structure* 17. 480–519. DOI: *https://doi.org/10.3765/bls.v17i0.1620*

Trueswell, John C. & Anna Papafragou. 2010. Perceiving and remembering events cross-linguistically: Evidence from dual-task paradigms. *Journal of Memory and Language* 63. 64–82. DOI: *https://doi.org/10.1016/j.jml.2010.02.006*

Zheng, Mingyu & Susan Goldin-Meadow. 2002. Thought before language: How deaf and hearing children express motion events across cultures. *Cognition* 85. 145–175. DOI: *https://doi.org/10.1016/S0010-0277(02)00105-1*

TO CITE THIS ARTICLE:
Bunger, Ann, Dimitrios Skordos, John C. Trueswell and Anna Papafragou. 2021. How children attend to events before speaking: crosslinguistic evidence from the motion domain. *Glossa: a journal of general linguistics* 6(1): 28. 1–22. DOI: *https://doi.org/10.5334/gjgl.1210*

Submitted: 03 February 2020
Accepted: 02 February 2021
Published: 24 March 2021

Glossa: a journal of general linguistics is a peer-reviewed open access journal published by Ubiquity Press.