Cognition 122 (2012) 135–149

Contents lists available at SciVerse ScienceDirect

Cognition

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

The relation between event apprehension and utterance formulation in children: Evidence from linguistic omissions

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

[a] Department of Psychology, University of Delaware, 108 Wolf Hall, Newark, DE 19716, USA
[b] Department of Psychology, University of Pennsylvania, 3401 Walnut Street, Philadelphia, PA 19104, USA

ARTICLE INFO

Article history:
Received 17 February 2010
Revised 23 September 2011
Accepted 2 October 2011
Available online 4 November 2011

Keywords:
Event cognition
Children
Speech planning
Motion
Space
Spatial language

ABSTRACT

The relation between event apprehension and utterance formulation was examined in children and adults. English-speaking adults and 4-year-olds viewed motion events while their eye movements were monitored. Half of the participants in each age group described each event (Linguistic task), whereas the other half studied the events for an upcoming memory test (Nonlinguistic task). All participants then completed a memory test in which they identified changes to manners of motion and path endpoints in target events. In the Nonlinguistic task, eye movements and memory responses revealed striking similarities across age groups. Adults and preschoolers attended to manner and path endpoints with similar timing, and in the memory test both successfully detected manner and path changes at similar rates. Substantial differences in production emerged between age groups in the Linguistic task: whereas adults usually mentioned both manners and paths in their event descriptions, preschoolers tended to omit one event component or the other. However, eyegaze patterns remained equivalent across the two age groups, with both children and adults allocating more attention to event components that they planned to talk about. Children in the Linguistic task were at chance in the memory test, whereas adults actually showed a memory benefit as compared to the Nonlinguistic task. We conclude that developmental differences in the description of motion events are not due to pure attentional differences between adults and children, but leave open the possibility that they stem from limitations that are solely linguistic in nature or that arise at the interface of attention and language production.

1. Introduction

The ultimate goal in studies of language production is to gain an understanding of the processes that carry an idea from message formation to linguistic formulation. This area of research is complicated by the fact that the object most readily available for study is the output of language production, i.e., utterances, but the representations and processes that contribute to this output are more difficult to observe. Recent studies have demonstrated that in adults, on-line tracking of attention allocation provides a powerful window into the process of utterance formulation. Studying the time course from apprehension of some visual stimulus to grammatical encoding of some components of that stimulus has allowed researchers to infer the content of underlying representations as they are built and to observe the way that adults transform those conceptual representations into linguistic representations (e.g., Bock, Irwin, Davidson, & Levelt, 2003; Gleitman, January, Nappa, & Trueswell, 2007; Griffin & Bock, 2000; Papafragou, Hulbert, & Trueswell, 2008; Trueswell & Papafragou, 2010). However, the way that language production relates to attention allocation and event representation in young children has not yet been studied

* Corresponding author. Tel.: +1 302 831 8827; fax: +1 302 831 3645.
E-mail addresses: bunger@udel.edu (A. Bunger), trueswel@psych.upenn.edu (J.C. Trueswell), papafragou@psych.udel.edu (A. Papafragou).

0010-0277/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2011.10.002

systematically. In this paper, we describe a novel experimental method that sheds new light on this process, offering an opportunity to better understand how children accomplish the mapping between thinking and speaking.

When adults view a static or dynamic depiction of an event while planning to talk about what they see, they very quickly direct their attention to components of the scene that they plan to talk about, usually in the order that they plan to speak about them (e.g., Bock et al., 2003; Gleitman et al., 2007; Griffin & Bock, 2000). Griffin and Bock (2000) tracked eyegaze as adults viewed static line drawings of simple actions that could be described with transitive sentences (e.g., a mouse spraying a turtle with a water pistol). Analysis of eye movements in relation to language production revealed the influence of linguistic formulation on allocation of attention to the participants in the depicted events. Regardless of whether speakers produced active or passive descriptions of the images, they turned their attention to the event participant that they would encode as the subject of their sentence up to a full second *before* they began their utterance, followed by looks to the event participant they planned to encode as the object. For example, speakers who were preparing an active event description like "The mouse is spraying the turtle" directed their attention to the mouse approximately one second before they began to speak, followed by looks to the turtle. Gleitman and colleagues (2007) replicated this tight coupling of attention allocation and word order, demonstrating moreover that even the earliest looks to event participants can be used to predict which of them will be mentioned first in a description of the event.

Bock and colleagues (2003) have argued that patterns of attention allocation during language planning are informative not only about the process of linguistic formulation, but also about the way that speakers form a representation of (conceptualize) the relationships between components of a visual stimulus during apprehension. In their study, speakers of English and Dutch who were asked to report times presented on analog and digital clocks using either absolute (e.g., "four thirty") or relative (e.g., "half-past four") linguistic formulations directed their attention to the region of the clock that would be relevant to the first part of their report (hours for absolute reports and minutes for relative reports) within a few hundred milliseconds after onset of the visual stimulus. Gleitman and colleagues (2007) showed, moreover, that the starting point of event apprehension can be manipulated by nonlinguistic attention capture mechanisms, which give rise, in turn, to differences in the order in which event components are mentioned in linguistic output.

Taken together, these studies demonstrate clear relations between patterns of attention allocation and language planning in adults. Other studies have shown that the attention that adults allocate to different components of a scene in preparation for speaking varies by language background (Papafragou et al., 2008). This suggests that crosslinguistic differences in the encoding of event components run at least as deep as linguistic formulation. However, adults do not exhibit these language-specific patterns of scene inspection while performing nonlinguistic tasks (e.g., when asked to inspect a scene in preparation for a memory task; Papafragou et al., 2008) or tasks in which access to language has been blocked during event viewing (Trueswell & Papafragou, 2010).

Very little is known about how children allocate their attention as they view scenes, either while they are planning a description of a scene or while they are simply inspecting the scene. Does speech planning direct attentional resources in children in the same ways that it does in adults? The development of remote eye-tracking systems has made it easier to capture the relations between the way children apprehend the world and the way they formulate utterances to describe their experience of that world. In this study, we take advantage of these technological advances to look more closely at the relation between language production and the ongoing inspection of event scenes in preschool-aged children. Specifically, we examine children's attention allocation (via eyetracking) while they describe short animated events. As we argue below, a fuller picture of what children understand about events as well as the attentional mechanisms that support language planning can be gained by examining children's visual inspection of events when they are, and are not, engaged in the task of describing them. We focus on children's apprehension and description of simple motion events, which provide a particularly suitable domain in which to probe for developmental differences.

Following Talmy (1985), we define a motion event as one in which a particular Figure experiences a change in location with respect to some Ground object. The details of a motion event may be filled in by specifying the Manner in which the Figure moves (e.g., roll, bounce, slide) and the trajectory, or Path, that the Figure takes in relation to the Ground object (e.g., exit, circle, down, up). In the case of a bounded motion event, the Ground object may define either the Origin (often referred to as the Source) or the Endpoint (often referred to as the Goal) of motion. In addition, motion events may involve some Instrument (e.g., a car) and/or Cause (e.g., a catapult) that determines either the manner or path of motion. When describing a motion event, speakers make choices about which of these event components they want to include in their utterance and how they want to package those components in the linguistic representation. There are well-known typological trends in the encoding of motion event components (e.g., Slobin, 1996; Talmy, 1985). Adult speakers of so-called *satellite-framed* languages like English and Chinese tend to conflate motion with manner in the main verb of a sentence, and to either express path in a non-verb position or to omit it altogether. This pattern is demonstrated in sentence (1) below, in which the verb "flew" describes the bird's manner of motion and the optional post-verbal prepositional phrase "to its nest" describes its path of motion. In contrast, speakers of *verb-framed* languages like Spanish, Greek, and Turkish are more likely to conflate motion with path in the main verb and to encode manner in a non-verb position (if at all). In sentence (2), the verb "entered" describes the path of the bird's motion and the post-verbal modifier "flying" describes its manner of motion.

(1) English

The bird	flew	(to its nest).
FIGURE	MANNER VERB	PATH MODIFIER

(2) Modern Greek

To puli	bike	sti folia tu	(petontas).
The bird	entered	into its nest	flying
FIGURE	PATH VERB	PATH MODIFIER	MANNER MODIFIER

These language-specific patterns begin to emerge in early childhood, according to some reports by as early as 3 years of age (Allen et al., 2007; Naigles, Eisenberg, Kako, Highter, & McGraw, 1998; Özyürek et al., 2008; Papafragou, Massey, & Gleitman, 2006; Papafragou & Selimis, 2010; Slobin, 1996; Zheng & Goldin-Meadow, 2002). There are, however, substantial developmental differences in the description of motion events. In particular, children's descriptions tend to be less informative than those provided by adults. Developmental studies of motion event description find differences between adults and children in the frequency with which event components are encoded in linguistic representations (Allen et al., 2007; Özyürek et al., 2008; Papafragou et al., 2006; Papafragou & Selimis, 2010). Papafragou and Selimis (2010) report, for example, that 4- and 5-year-old English-speaking children are far less likely than adults to include *both* manner and path in their descriptions of dynamic motion events. In their study, adults mentioned both the manner and the path of motion events in 72% of their event descriptions, whereas children mentioned both event components in only 29% of their descriptions. Özyürek and colleagues (2008) report a similar pattern of omissions: although English-speaking 3- and 5-year olds and adults showed a similar syntactic distribution of manner and path elements in their motion event descriptions, adults mentioned both event components in 84% of their event descriptions compared to 58% for 5-year-olds and 49% for 3-year-olds. Papafragou and colleagues (2006) found that this trend toward omission extends to children as old as 8 years of age: in their study English-speaking 8-year-olds encoded both manner and path in only 50% of motion event descriptions as compared to around 80% for adults.

The developmental literature suggests that linguistic gaps in children's utterances about motion are unlikely to arise as a result of developmental gaps in conceptual sophistication (cf. Gentner, 1982; Huttenlocher, Smiley, & Charney, 1983; Johnston & Slobin, 1978; Piaget, 1955). On the contrary, children from a range of cultural and linguistic backgrounds both perceive and describe motion event components from an early age (cf. Lakusta, Wagner, O'Hearn & Landau, 2007; Mandler, 2004; Pruden, Hirsh-Pasek, & Golinkoff, 2008). Children appear to be able to discriminate both manners and paths of motion in nonlinguistic tasks by 14 months of age (Pulverman & Golinkoff, 2004; Pulverman, Golinkoff, Hirsh-Pasek, & Buresh, 2008; Pulverman, Hirsh-Pasek, Golinkoff, Pruden, & Salkind, 2006; Pulverman, Sootsman, Golinkoff, & Hirsh-Pasek, 2003). Pulverman et al. (2008) report, for example, that after being habituated to an event in which an animated character moved with a particular manner along a particular path, 14- to 17-month-old children dishabituated to changes in either the manner or the path of motion. This finding held across language backgrounds: dishabituation to each event component was equivalent in children learning English (a satellite-framed language) and Spanish (a verb-framed language). Young children also seem to be able to form categories based on motion event components across varied motion events: paths of motion by 10 months and manners of motion by 13 months of age (Pruden, Hirsh-Pasek, Maguire, & Meyer, 2004; Pruden et al., 2008). In addition, expressions that specify manners and paths of motion events appear very early in children's vocabularies and, in the case of deaf home signers, in their gestural repertoire (e.g., Fenson, Dale, Reznick, & Bates, 1994; Naigles et al., 1998; Zheng & Goldin-Meadow, 2002).

Given that children appear to have a clear conceptual grasp of motion event components from a very early age, the question remains why they do not mention these components as often as adults do when describing motion events. One possibility is that children, unlike adults, lack the attentional capacity to simultaneously encode the complete set of components that make up a complex event—i.e., limitations in the event apprehension process, rather than in the linguistic formulation process, are resulting in this developmental difference in what children and adults talk about. There is independent evidence that the cognitive mechanisms that regulate attentional control develop over time, and that young children are more limited than adults in, for example, their ability to divide their attention between multiple locations in a visual array (e.g., Johnson, 1995; Scerif et al., 2005). If a general attentional deficit underlies children's omissions, then it may be the case that children do not mention a particular event component simply because they have not paid attention to it. That is, perhaps children sometimes attend to manner and sometimes to path, based on the individual properties of a given event, and whichever event component they have attended to is the one they end up mentioning. Adults, in contrast, can adequately divide/allocate their attention among all relevant event components, and so are more likely to mention more of those components.

Alternatively, it may be the case that children's omission of manner and path information is driven by constraints introduced by the task of language production itself. That is, children may attend to motion events and conceptually encode them with the same sophistication as adults, but linguistic output limitations—related to cognitive or linguistic resources or specific task demands—make it difficult for them to actually talk about everything they have encoded, resulting in the dropping of information from their event descriptions. These limitations could arise for a number of reasons, including developmental differences in lexical and/or syntactic accessibility associated with the typical ways of describing components, or even the working memory demands associated with planning certain structures.

In the current study, we adjudicate between these two general explanations for child omissions by assessing the relations between event apprehension and event description in preschool-aged English-speaking children. Pairing eyetracking with a language production task allows us to ask whether age-related discrepancies in linguistic output stem from developmental differences in attention to motion events or whether they arise as part of the process of encoding those events in language. In particular, we investigate children's allocation of attention to particular components of a dynamic motion event and the encoding of those components in a verbal description of the event. We ask whether differences in children's utterances relative to adults can be attributed to differences in attentional patterns. In addition, we ask whether, like adults, children show a tight relationship between the allocation of attentional resources during utterance formulation and the content of the resulting utterance. To our knowledge, this is one of the first studies to compare the attentional signatures of speech planning in children and adults.

More specifically, we compare eye movements of English-speaking 4-year-old children and adults viewing dynamic motion events while manipulating viewing conditions to present a Linguistic task (in which participants viewed and described motion events) and a Nonlinguistic task (in which participants freely viewed motion events in preparation for a memory task). Examining the eyegaze data allows us to assess the event components that participants in each age group actually attend to, whether those components make it into their event descriptions or not. Eye movement patterns in the Nonlinguistic task should reveal the way that children and adults assemble an event representation to be used for committing the event to memory (Papafragou et al., 2008), and eye movement patterns in the Linguistic task are expected to be informative about the process of apprehending an event while formulating an utterance describing the event (see Griffin and Bock (2000) and Papafragou et al. (2008) for discussion of adult data).

If children's omissions of manners and paths from their event descriptions are due to an attentional difference between children and adults, we would expect to also observe this difference in their eye movements as the event unfolds, regardless of the task presented to them. That is, children in the Linguistic task should pay less attention than adults to the same event components that they tend to omit, and crucially, children in the Nonlinguistic task should show an equivalently restricted pattern of attention allocation. Alternatively, if the omissions arise as a result of limitations introduced by language planning, *any age-related differences in event apprehension should only surface in the Linguistic task*. That is, if children apprehend and conceive of the motion events as adults do, allocation of attention to manners and paths should be quite similar for children and adults in the Nonlinguistic task. However, in the Linguistic task, differences in attention allocation might arise as children grapple with competing resources.

A memory task presented to all participants at the end of the experiment provides an independent measure of event conceptualization, on the assumption that memory for the details of particular event components is revealing of the structure of underlying event representations. If children differ from adults in the way they apprehend motion events (especially manners and paths of motion), this difference should surface in memory performance in the Nonlinguistic task: children should have a poorer memory for the same event components that they did not attend to as much as adults. Alternatively, if children apprehend motion events as adults do, memory for manner and path elements should be comparable between the two age groups.

2. Method

2.1. Participants

The final sample included 20 4-year-old children (mean age 4;6 years;months, range 4;1–5;0) and 20 adults. Data from eight additional preschool-aged participants were excluded from the analysis for the following reasons: equipment failure ($n = 4$), unwillingness to cooperate ($n = 1$), experimenter error ($n = 1$), inability to calibrate ($n = 1$), or significant trackloss during stimulus viewing ($n = 1$; see Section 2.5 for trackloss criteria). All participants were native monolingual speakers of English. The children's parents were recruited from preschools in Newark, DE and Philadelphia, PA. Adult participants were students at either the University of Delaware or the University of Pennsylvania and received either $8 as compensation or course credit for participation. Children had no parent-reported history of visual, cognitive, or language impairments. Before participating, adults provided written consent and children provided verbal consent along with written consent from a parent or guardian, as approved by the Institutional Review Board (IRB).

2.2. Apparatus

A Tobii 1750 remote eyetracking system was used for stimulus presentation and data collection. This system performs binocular tracking using optics embedded in a flat panel monitor with a display size of 33.5 (width) × 26.75 (height) cm (31.2° × 26.9° visual angle at viewing distance of 60 cm). Two laptop computers running the Windows XP operating system were used to control the system: one displayed stimuli on the eyetracker screen (via the ClearView software from Tobii Technology) and the other collected eyegaze data from the eyetracker (via the TET-server software developed by Tobii Technology). Both laptops were disconnected from the internet to increase timing accuracy. Data sampling rate was a consistent 50 Hz, and the spatial resolution of the eyetracker is approximately 0.5–1° visual angle, including corrections for head drift and small head movements. At a 60-cm viewing distance, the Tobii 1750 has a tolerance to head motion of about 30 × 16 × 20 cm. The system recovers from a complete eyetracking failure in <100 ms.

2.3. Materials

Stimuli consisted of short videos created by animating clip-art images. Target items consisted of 12 motion events in which a human or animal agent moved with the assistance of an instrument to reach a stationary path endpoint. Still frames from the beginning, middle, and end of one of

Fig. 1. Still frames taken from the beginning (A), middle (B), and end (C) of a target event. In this target event, a boy (the agent) on roller skates (the instrument) moves across the screen and into a soccer net (the path endpoint). These still frames show the beginning, middle, and end of a 3-s animation.

the target events can be seen in Fig. 1: in the animation that these still images are taken from, the clipart boy on roller skates moved across the scene space from left to right, ending in the soccer net. Manners and paths in each target event were represented by distinct visual elements. In every event, the manner of motion was associated with an instrument or vehicle (e.g., roller skates, car, hot air balloon) and the agent rode on or in this instrument toward a visible endpoint that determined the path (e.g., soccer net, cave, building). A full list of target events can be found in Appendix A, Table A1. In four target events, subparts of the instruments were animated to achieve the illusion of motion: in separate events, an airplane propeller and the wheels of a car, motorcycle, and roller skates were made to rotate. Other than these exceptions, target events included no motion other than gross movement of the agent and instrument (as a unit) toward the path endpoint. No subpart of the agents (e.g., hands, arms, heads, legs, feet) moved independently from the agent-instrument unit.

Clipart images were constructed such that the instrument was spatially separated from the torso and face of the agent, allowing looks to these two regions to be distinguished in the analysis of eyegaze data. In keeping with the goal of creating clear regions of interest for the eyetracking analysis, our stimuli did not include agent-driven manners of motion like running or jumping. The identification of manner in events like running and jumping requires global appreciation of the agent's motion (e.g., people can discriminate such events from simple point-light displays; Johansson, 1973), and it is not clear how a Manner region that is distinct from the agent itself would have been identified in videos of these kinds of events. In addition, trajectories of agent motion in our stimuli were never marked by visual paths like winding roads or jet streams. A simple background was created for each video that was contextually appropriate (e.g., a sidewalk in a neighborhood).

Two modified versions of each target video were created for the memory test. In the Manner-change version (Fig. 2A), the agent's manner of motion was modified by changing the instrument or vehicle used to approach the path endpoint, e.g., substituting a skateboard for roller skates. In the Path-change version (Fig. 2B), the object denoting the path endpoint changed. Path changes always resulted in a linguistically relevant difference in the relation between the agent and the path landmark, e.g., instead of skating *into* the soccer net, the boy skated *to* a gazebo. Independent testing was carried out to assess the degree of visual difference in Manner- vs. Path-change versions of each target video. Ten monolingual English-speaking adults viewed pairs of videos consisting of a target event and one changed version (24 pairs of videos in all) and were asked to rate on a 5-point scale (1 = not very different; 5 = very different) how "visually different" the pairs were. Ratings of visual difference between target and change videos revealed no significant difference for Manner vs. Path changes (average ratings for Manner-change vs. target: 2.81 and Path-change vs. target: 2.97; *t*-test on item means: $t(11) = 0.6$, $p = 0.56$).

Twelve filler videos were also constructed. Filler videos depicted animate agents and inanimate objects involved in events that did not include specific endpoints (e.g., flying a kite).

In all videos, movement lasted for 3 s, at which time a beep was heard. Except for this beep, videos were silent. After the movement ended in a video (and the beep sounded), the final frame of the event remained visible for an additional amount of time: 2 s for adults and 6 s for preschoolers. Pilot testing of preschoolers with adult-length (5-s) videos revealed that the children needed more time to complete their event descriptions than adults did. Videos presented to preschoolers were lengthened to avoid a situation in which participants were asked to produce descriptions of events that were no longer visible. Clipart animations were first created in Microsoft PowerPoint and then converted to Audio Video Interleave (avi) files using a conversion program. Timing was verified in video editing software. To lengthen videos for presentation to preschoolers, adult-length avi files were imported into Apple's Final Cut Pro software, and the still image that appeared at the end of the video was extended for an additional 4 s. When presented on the Tobii screen, stimulus videos were 20.9 (width) $\times$ 16.7 (height) cm for adults ($19.8^\circ \times 15.9^\circ$ visual angle at a viewing distance of 60 cm) and 23.6 $\times$ 16.7 cm for preschoolers ($22.2^\circ \times 15.9^\circ$).[1]

2.4. Procedure and experimental design

Each participant was run individually at his/her university campus or preschool. For preschoolers, two

[1] Modification of stimuli in Final Cut Pro resulted in the export of videos with a different aspect ratio from those created using PowerPoint. The resulting horizontal stretch does not produce a noticeable distortion of the clipart images.

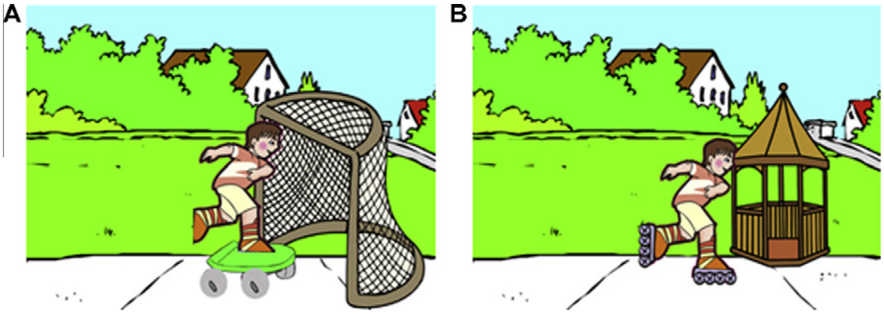

Fig. 2. Still frames taken from the end of test events depicting changes to Manner (A) and Path (B) elements. These frames depict changes to the target event shown in Fig. 1. In the Manner-change version (A), the boy moves into the soccer net on a skateboard instead of on roller skates; in the Path-change version (B), the boy moves on roller skates to a gazebo instead of into a soccer net. Each of these still frames is taken from the end of a 3-s animation.

experimenters were present during the session—one to manage the Tobii system and control the display of stimuli, and another to interact with the child and record data during the memory task. For adults, only one experimenter was present. Participants were seated approximately 60 cm from the Tobii screen, and the experimenter adjusted the angle of the screen as necessary to obtain robust views of both eyes, centered in the tracker's field of view. Preschoolers sat unconstrained in a car seat firmly attached to an office chair to limit wiggling during the session.

The Tobii ClearView default 5-point calibration scheme was used for participants in both age groups. During this procedure, a blue circle of size "medium" moved around the screen while changing size, looming to full size ($\sim$2.5 cm diameter) in the center of the screen and in each of the four corners (with the center of the circle inset $\sim$4.5 cm toward the center of the screen at each corner). Preschoolers were told that the calibration routine was a "blue dot game," and were instructed to "sit very still and chase the dot with their eyes." Adults were simply asked to watch the blue dot as it moved around the screen. Immediately after the calibration routine, experimenters judged calibration quality by examining the ClearView calibration plot for (1) the number of calibration points for which eyegaze data was collected and (2) the spread of data around those points. If the calibration data did not meet the default criteria of the ClearView Software or if it was incomplete (fewer than 5 points for adults, fewer than 4 points for preschoolers), the calibration routine was repeated. Adults typically required only one calibration; one preschooler who failed to calibrate was excluded from the analysis.

Participants were then given instructions specific to their task. Participants were randomly assigned to one of two tasks based on a rotation through an experimenter-generated list: half of the participants in each age group were assigned to a Linguistic task, and the other half to a Nonlinguistic task. Stimulus presentation in both tasks followed the same progression: during an initial encoding phase, participants viewed a sequence of 24 animated clips (12 target events and 12 fillers), and after viewing this set of events, they proceeded directly to the memory test phase, in which they viewed a second set of 24 clips (12 target-change events and 12 fillers). Instructions for children and adults were the same, with minor age-appropriate changes in wording. In both of the tasks, participants were informed before the encoding phase that they would be viewing "short animated clips" ("cartoons" for preschoolers) with "people and animals doing things." Participants in the Linguistic task were asked to "describe what happened in the clip" ("tell me what happened in the cartoon") as soon as the beep sounded and the clip froze on the screen. These participants were not informed in advance about the memory test. Participants in the Nonlinguistic task were asked to remember the clips presented during encoding in preparation for a memory task ("a memory game").[2] This instruction was given to ensure that Nonlinguistic participants were motivated to pay attention to the videos during the encoding phase of the experiment. Linguistic encoding of events by participants in the Nonlinguistic task was discouraged: participants in this condition who began to give descriptions were asked to "watch quietly." When the memory test began, participants in both conditions were asked to say whether these new clips were "the same" or "different" from the ones they had seen before.

In both conditions, participants viewed one of two lists of the stimulus events. In each list, the 24 encoding videos were shown in a fixed random order, and the 24 memory videos were shown in the same fixed random order. All filler events shown during the memory task were identical to those that participants had seen before, and all target events were different—half had Manner changes and half Path changes. Each participant saw only one kind of change (Manner or Path) for a given target event, and changes were distributed across the two lists so that an event that appeared with a Manner change in one list appeared with a Path change in the other. Presentation of events during the memory task was experimenter-paced, allowing participants to move on to a new event as soon as a judgment had been provided for the previous one. A recentering stimulus was shown between all videos during both phases of the experiment. For adults, this stimulus was a crosshair displayed just above the center of the screen. Instead of a crosshair, preschoolers saw a video in which colorful objects (e.g., stars and smiley faces) flew around the screen. This animated version of the recentering stimulus allowed

[2] This is a replication of the procedure reported as the Nonlinguistic task in Trueswell and Papafragou (2010), using 20 new adult participants for this paper.

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.

Stimuli were presented to all participants on the same Tobii 1750 remote eyetracking system. Responses were made verbally, and sessions in which participants completed the Linguistic task were audio-recorded. Experimenters recorded memory judgments by hand.

2.5. Data coding and analysis

2.5.1. Transcription and coding of event descriptions

Participant descriptions of target events in the Linguistic condition were transcribed and coded by hand. A second person transcribed 25% of the target event descriptions (a total of 48 descriptions). Out of 320 words provided by participants in these event descriptions, there were 12 disagreements between the transcriptions, none of which would have affected the way event descriptions were coded for analysis.

Words or phrases that referred to instruments (e.g., "roller skates") or the agent's manner of motion (e.g., "skater," "skating," "riding") were coded as Manner mentions, and those that referred to either the path endpoint (e.g., "goal," "soccer net"), the agent's trajectory of motion (e.g., "into"), or the relationship between the agent and the path endpoint (e.g., "entering") were coded as Path mentions.[3] For example, an utterance like (3a) would be coded as including both Manner and Path, whereas (3b) includes only Manner information and (3c) includes only Path information:

(3) a. The boy skated into the net.
b. The boy skated.
c. The boy went into the net.

Transcriptions were coded by one of the authors (A.B.), and then codes were reviewed by a second author (A.P. or J.T.). Disagreements about coding were resolved as they arose. Mean length of utterance (MLU) for each subject was calculated based on word counts. Event descriptions were not provided by preschoolers for 4 of the 120 Linguistic trials; an additional five event descriptions from preschoolers were excluded because they were ungrammatical.

2.5.2. Analysis of eye movement data

Eye movement data from target trials were analyzed to assess the effects of age and task on encoding of event components. Only eyegaze data from the encoding phase were analyzed. Data samples (50 per second) were time-locked to the onset of each video, and analyses were performed on raw eyegaze coordinates from each sample. Trackloss was determined separately for each eye by the Tobii ClearView software. On each sample, ClearView assigns a Validity code (0–4) to the data reported for each eye that indicates the system's certainty that it has recorded the correct eyegaze coordinates. Our data set includes samples with validities of 0 (the system is certain that it has recorded the correct coordinates and that data has been assigned to the correct eye) or 1 (recorded coordinates are correct, but the system has only recorded one eye). Missing data (samples from either eye with Validity >1, including blinks) were counted as trackloss for a given eye. If trackloss occurred for only one eye, gaze coordinates were taken from the eye with no trackloss; if both eyes passed the validity screening, gaze coordinates from the two eyes were averaged together. Trials with global track-loss of >33% were excluded from the analysis (no adult trials; Preschoolers in the Nonlinguistic condition, 4 of 120 trials; in the Linguistic condition, 8 of 120 trials). Subjects with four or more excluded target trials were excluded from analysis entirely and replaced in the design ($n = 1$ preschooler, Linguistic condition).

Two dynamic spatial scoring regions were defined for each target video: a Manner region, defined as an invisible rectangle around the instrument used by the agent as the means of motion (e.g., the roller skates), and a Path endpoint region, defined as an invisible rectangle around the stationary Path endpoint of the motion event (e.g., the soccer net). The Manner region never included the head or torso of the agent. Because the agent did not move independently, these omitted areas did not provide any additional information about the manner of motion, and this division allowed looks to the agent (not reported) and the instrument to be analyzed separately. The decision to include only the path endpoint in the Path region was made on the basis of previous work in our labs showing that viewers of motion events like these tend to make early looks that project the agent's trajectory—the region of space through which the agent moved toward the path endpoint—from the starting point of motion to a visible path endpoint, but not to fixate intermediate points in the motion trajectory (Papafragou et al., 2008). The decision to omit the trajectory from the Path region was made for two reasons: First, the trajectory itself was never visible in the events, i.e., there was never a marked path, road, or jet stream that the agent followed to reach the path endpoint, and viewers rarely fixate empty regions of space. Second, linguistic encoding of paths is achieved via reference to some fixed point along the path (a source, a route midpoint, or a goal), and in our stimuli the Path endpoints provided these reference points. Thus, when participants consulted the Path endpoint of our dynamic events, they collected information not only about the reference object itself, but also about the linguistically relevant path, i.e., the relation of the agent to that object. On average, Manner regions subtended 6.14° (width) × 2.67° (height) visual angle for adults and 6.90° × 2.67° for preschoolers, Path regions subtended 8.21° × 8.98° visual angle for adults and 9.22° × 8.98° for preschoolers. Exact sizes of these regions for each target video are given in Appendix A, Table A1.

Because the instrument moved across the screen toward the path endpoint during the videos, 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

[3] Nouns ("roller skates," "soccer net") were included in event component counts as well as prepositions and verbs because they are related transparently to the regions coded as Manners and Paths in the eyetracking analysis.

video. A coder then repositioned the Manner region by hand for each frame of the video, recording the coordinates of this region in each successive frame to a file. Manner region size remained constant across frames, as did the size and position of the Path region. For the analysis, the coordinate file was merged with the eyetracking data file. An eyegaze sample was defined as being within a region of interest if its coordinates fell within the region as defined for that moment in the video. Manner and Path regions sometimes overlapped near the end of events, as instruments moved "into" or "onto" path endpoints. Cases of overlap were resolved by assigning gaze to the event component that was visible in the top layer of the visual display—always the instrument—rather than to the occluded one. 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 s. Any looks within a region were included in the analysis, regardless of duration.

2.5.3. Analysis of memory data

Accuracy during the memory phase was assessed by comparing Same/Different judgments collected by hand with actual changes to target stimuli in each list.

2.5.4. Statistical analyses

The reliability of trends observed in the data was tested using multilevel mixed logit modeling with crossed random intercepts for Subjects and Items (after Baayen, Davidson, & Bates, 2008). This analysis offers a more sophisticated treatment of random factors (Subject, Item) than the tradition of analyzing subject and item means separately. Logit modeling, a kind of linear regression, is also a better treatment of the data because coding values are binomial (absent vs. present) at the trial level. Mixed logit models include random factors like subject and item in a single model, allowing categorical data to be treated as categorical rather than as proportions of subject or item means as in a traditional analysis of variance (ANOVA). Moreover, because the data that we report are categorical factors reported as proportions, they violate the ANOVA assumption that variances will be homogenous across the entire data set (Jaeger, 2008). See Jaeger (2008) and Manning (2003) for a more complete description of this kind of model and direct comparisons between the output of mixed logit modeling and ANOVAs. For the sake of continuity with previous research in this field (and ease of interpretation), we have also included the results of complementary ANOVA analyses on subject and item means.

3. Results

3.1. Spoken event descriptions

Our analyses of the descriptions that were provided by participants in the Linguistic condition revealed that children were more likely than adults to omit motion event components from their descriptions of target events. Children produced event descriptions that were shorter than those produced by adults (preschool MLU in words = 5.8; adult = 7.7). Table 1 provides a coarse report of the semantic content of those utterances: we found that adults mentioned *both* Manners and Paths in their event descriptions 70% of the time, whereas preschoolers mentioned both event components only 26% of the time, replicating previous production results (Özyürek et al., 2008; Papafragou & Selimis, 2010). Looking at mention of each event component separately, we see that both preschoolers and adults were more likely to mention Manners than Paths in their event descriptions. Adults mentioned Manner of motion in every utterance and Paths in about two-thirds of these utterances. Preschoolers were more likely than adults to drop event components, mentioning Manners slightly less often than adults did and Paths only about half as often as adults.[4] For the roller skating event depicted in Fig. 1, for example, most adults provided a description that included both the skates (Manner) and the net (Path), as in (4a). Preschoolers, on the other hand, were more likely to provide a description like (4b), including only the skates.

(4) a. Adult
A boy rollerbladed into a soccer net.
b. Preschooler
The boy was skating.

3.1.1. Mixed logit modeling of event component mention

We tested the reliability of these observations using multilevel mixed logit modeling with crossed random intercepts for Subjects and Items as described in Section 2.5.4 above. Binary values at the trial-level for mention of Manner and Path in event descriptions (mentioned, not mentioned for each Event Component) were modeled using Age (Adult, Preschool) as a (between-subjects) first-level fixed factor. The model revealed main effects of Age and Event Component (both $p < 0.0001$), but no interaction between the two (Table 2).[5]

[4] Our finding that both children and adults were more likely to omit Paths than Manners from their event descriptions may appear surprising, given the well-attested goal bias in production (e.g., Allen et al., 2007; Lakusta & Landau, 2005; Papafragou, 2010) and in nonlinguistic event representation in children (e.g., Lakusta, Wagner, O'Hearn, & Landau, 2007; Papafragou, 2010). As we show later, memory for Paths was better than memory for Manners, despite the fact that Manners were more likely to be mentioned. We suspect that what looks like a bias against Paths is actually driven by an increase in the salience of our Manners of motion compared to previous studies that have reported higher production of Path elements. The following characteristics of our stimuli may have made Manners more interesting to participants. First, we used Manners of motion that relied on instruments (e.g., riding a skateboard, driving a car) rather than intrinsic movements of the agent's body (e.g., running, rolling). The very presence of these instruments, as well as the fact that a different instrument was used in each target event, probably played a role in increasing the salience of Manners of motion. In addition, in all of our events, the Manner of motion was causally related to the change in location (i.e., the boy uses the skateboard to reach the net) rather than being incidental to it (e.g., a tomato spins while falling off a cliff), a manipulation that has been argued to affect linguistic and gestural encoding of motion event components (e.g., Croft, 1991; Goldberg, 1997; Kita et al., 2007).

[5] One might be tempted by an informal inspection of Table 1 to anticipate a sizeable interaction between Age and Event Component. However, this is an illusion of looking at these data as the bounded variable proportion (see Jaeger, 2008). Our statistics were done either in logit space or using elogit transformations to avoid this error. Mean elogit transformed values for proportion mention of each event component can be found in Table A2 in Appendix A.

Table 1
Proportion of mention of event components in linguistic descriptions.

	Both Manner and Path	Manner only	Path only	Neither
Adults	0.70 (±0.07)	0.30 (±0.07)	0.00 (±0.00)	0.00 (±0.00)
Preschoolers	0.26 (±0.08)*	0.65 (±0.08)*	0.06 (±0.03)	0.03 (±0.02)

Note: Values indicate proportions (± standard errors of subject means) of event descriptions provided by participants in each age group. Significance tests for an effect of participant age on each coded production category were performed using multilevel mixed logit modeling with crossed random intercepts for Subjects and Items.
* Significantly different from adult production at $p < 0.01$.

Table 2
Fixed effects from best-fitting multilevel linear model of Manner and Path mention in event descriptions.

Effect	Estimate	S.E.	z-Value
Intercept	3.35	0.63	5.34*
Age (Adult vs. Preschool)	2.41	0.58	4.12*
Event component (Manner vs. Path)	−4.47	0.49	−9.10*

P-values come from a model that did not include the interaction term; a model that added the interaction term did not reliably improve the fit, based on a chi-square test of the change in −2 restricted log likelihood between the two models (Steiger, Shapiro, & Brown, 1985). Formula in *R*: DepVar ~ Age + Event Component + (1|Subject) + (1|Event).
* Significant at $p < 0.0001$.

3.1.2. ANOVA for event component mention

To complement the mixed logit modeling described above, repeated-measures ANOVAs were also conducted on elogit-transformed subject and item means for the proportion mention of Manner and Path elements in event descriptions, with participant Age (Adult, Preschool) as an independent variable. The results of these analyses echoed that of the multilevel modeling, with significant main effects of both Age (subject: $F(1,18) = 17.723$, $p < 0.0005$; item: $F(1,22) = 22.635$, $p < 0.0001$) and Event Component (subject: $F(1,18) = 50.634$, $p < 0.0001$; item: $F(1,22) = 77.939$, $p < 0.0001$) but no interaction between the two (subject: $F(1,18) = 1.031$, $p = 0.32$; item: $F(1,22) = 2.595$, $p = 0.13$).

3.2. Eye movements

Given the finding that preschoolers tended to omit event components from their event descriptions more often than adults, we examined the eye-movement data to ask whether children were also paying less attention than adults to these components while viewing the events. Eyegaze patterns from the Nonlinguistic task reveal whether Manners and Paths were available to the conceptual system, and eyegaze patterns from the Linguistic task reveal whether these components were also available to language. Informal inspection of the eye movements depicted in Fig. 3 reveals important similarities in allocation of attention to event components across age groups. In both tasks, adults and preschoolers allocated approximately the same amount of attention to Manners and on a similar time course (Fig. 3A). In the Linguistic task, but not the Nonlinguistic task, there was a gradual increase in attention to Manners over the viewing period for both age groups, with preschoolers increasing their attention to Manners earlier than adults. This eyegaze pattern is consistent with a strategy in which viewers allocated more attention to event components they planned to talk about: recall that both adults and preschoolers were highly likely to mention manners of motion in their event descriptions. Inspection of Fig. 3B reveals that adults and preschoolers in both tasks also allocated attention to Paths similarly throughout the viewing period. This similarity in allocation of attention to Manners and Paths across age groups and tasks suggests that, at some level of processing, these event components were just as available to preschoolers as they were to adults. That is, even if these components did not make it to the surface in preschoolers' utterances as often as they did for adults, they were available for both event apprehension and sentence planning.

3.2.1. Mixed elogit modeling of eye movements

We tested the reliability of these trends using multilevel mixed elogit modeling, following Barr (2008), using crossed random intercepts for Subjects and Items. Elogit-transformed proportions of Path looking minus elogit-transformed proportions of Manner looking were modeled within the five 1-s windows depicted in Fig. 3 using Task (Linguistic, Nonlinguistic) and Age (Adult, Preschool) as first-level predictors (Table 3). Task was a reliable predictor of looking patterns ($p < 0.05$) starting in the fourth window (3–4 s after video onset), reflecting the increase in looks to Manner in the Linguistic but not the Nonlinguistic task. In addition, there was a significant Age × Task interaction in the third analysis window (2–3 s from video onset), reflecting the earlier increase in Manner looks by preschoolers compared to adults.

3.2.2. ANOVA for eye movements

To complement the modeling described above, we conducted ANOVAs on subject and item means (averaged across 1-s blocks of the viewing time) of the same dependent variable used in the mixed models, with Task (Linguistic, Nonlinguistic) and Age (Adult, Preschool) as independent variables. The results of the subject analysis echoed that of the multilevel modeling, with a significant main effect of Task in the fourth (3–4 s after video onset; $F(1,36) = 6.418$, $p = 0.02$) and fifth (4–5 s after video onset; $F(1,36) = 7.303$, $p = 0.01$) windows of analysis, and an Age × Task interaction in the third window (2–3 s after video onset; $F(1,36) = 6.560$, $p = 0.01$). The item analysis showed a main effect of Task in the fifth analysis window

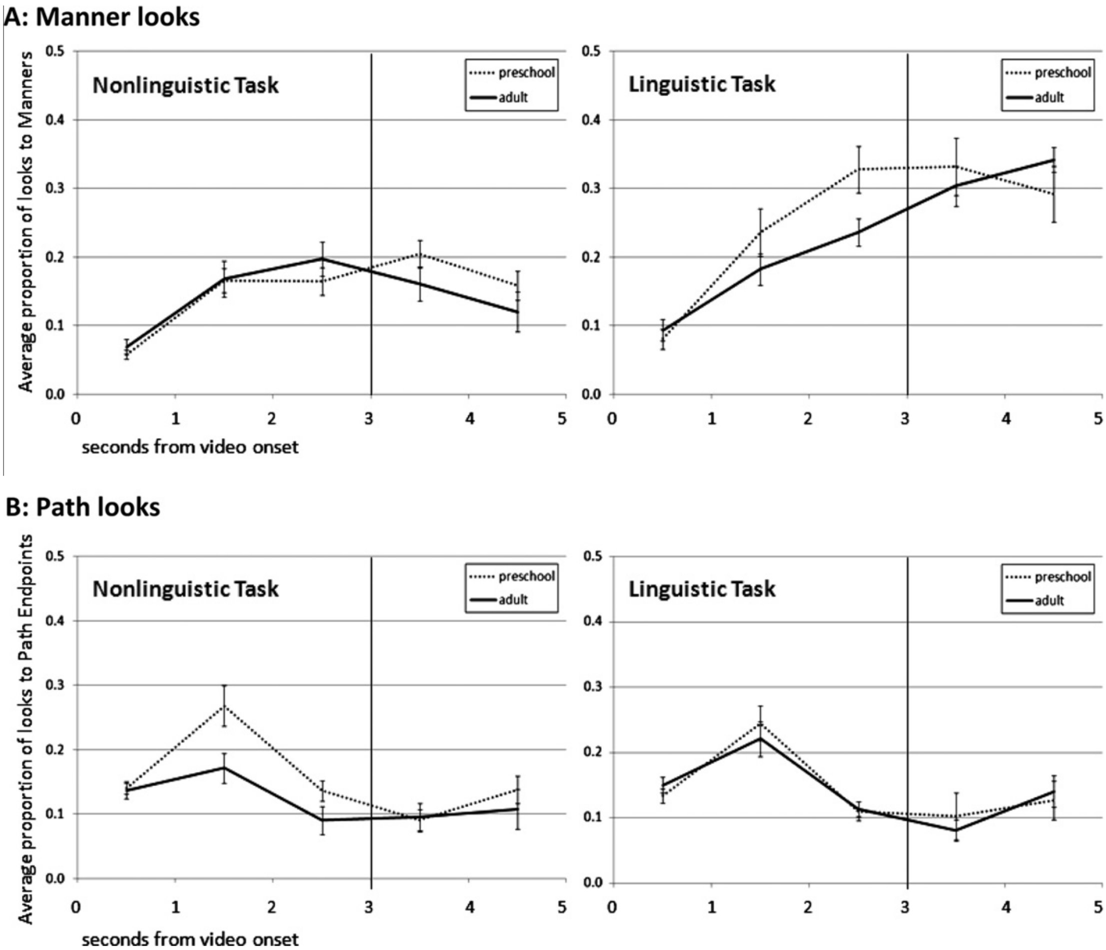

Fig. 3. Proportion of looks to Manners (A) and Paths (B) of motion events by task. Proportions are calculated based on all looks to events, and so are not expected to add up to 1 across these scoring regions. Gaze preferences are averaged across 1-s units of the viewing period. The full viewing period for adults is presented, and just the first 5 s of viewing for preschoolers. The vertical line that appears at second 3 indicates the beep that occurs during the viewing phase in each task. Error bars represent standard error of subject means.

Table 3
Fixed effects from best-fitting multilevel linear models of looking preference (Path minus Manner) in each time window of the viewing period.

Time window (sec from video onset)	Effect	Estimate	S.E.	*t*-Value
2–3 s	Intercept	−0.65	0.66	−0.98
	Age (Adult vs. Preschool)	0.63	0.40	1.59
	Task (Linguistic vs. Nonlinguistic)	−0.22	0.40	−0.56
	Age × Task	−1.27	0.57	−2.24*
3–4 s	Intercept	−0.78	0.43	−1.83
	Task (Linguistic vs. Nonlinguistic)	−1.24	0.35	−3.57*
4–5 s	Intercept	−0.17	0.43	−0.40
	Task (Linguistic vs. Nonlinguistic)	−1.41	0.36	−3.93*

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*: 2–3 s: DepVar ∼ Age ∗ Condition + (1|Subject) + (1|Event); 3–4 s: DepVar ∼ Condition + (1|Subject) + (1|Event); 4–5 s: DepVar ∼ Condition + (1|Subject) + (1|Event). For the 0–1, and 1–2 s windows, no models tested provided a better fit than an empty model with no fixed effects.
* Significant at $p < 0.05$.

(4–5 s after video onset; $F(1,44) = 4.488$, $p = 0.04$), but no effect of or interaction with Age.

3.3. Memory for event components

Data collected during the memory phase provide an independent measure of the event components preschoolers included in their motion event representations. Table 4 shows percent accuracy for noticing changes to Manner and Path components during the Memory test. Our analyses of accuracy on the memory task revealed that, overall, both age groups were more accurate at picking out changes to Paths than Manners, despite the fact that this was the event component most likely to be omitted from event descriptions by both age groups. Moreover, in the Nonlinguistic task, children were just as accurate as adults at identifying changes to Manners and Paths, suggesting that there is no general age-related difference in the encoding of these event components as part of motion event representations. Interestingly, the Linguistic task improved memory performance for adults whereas it decreased memory performance for children, relative to the Nonlinguistic task.

3.3.1. Mixed logit modeling of memory accuracy

Multilevel logistic modeling of accuracy on memory for changes to Manners and Paths of motion events was carried out as described above, using binary accuracy values (correct, incorrect) for each subject-item pair. Age (Adult, Preschool), Task (Nonlinguistic, Linguistic), and Change Type (Manner, Path) were entered into the model as first-level fixed factors. A model that did not include any interaction terms showed a main effect of Change Type ($p < 0.01$), but the best-fitting model included an Age × Task interaction (Table 5).

The main effect of Change Type reflects the fact that, overall, participants were more accurate at identifying changes to Paths than changes to Manners for our target events. The results of the norming study reported in Section 2.3 suggest, moreover, that any advantage that Path-change videos produce in the memory task does not arise from an imbalance of differences in the surface details of Path- vs. Manner-change videos.

The Age × Task interaction reflects two facts about performance on our memory task: First, adults, but not preschoolers, performed better on the memory test in the Linguistic condition than in the Nonlinguistic condition. In addition, the performance of preschoolers in the Linguistic task was no different from what would be expected if they made same/different decisions for test events at random (i.e., their performance was not different from chance). The increase in accuracy for adults in the Linguistic task suggests that they used linguistic representations of the target events to support memory. It is possible that children in the Linguistic condition were less likely to take advantage of memories of their spoken descriptions during the memory test. This hypothesis is consistent with studies reporting that children may not be able to use language strategically to support memory until they are 6 or 7 years of age (cf. Hitch, Halliday, Schaafstal, & Heffernan, 1991; Kahn & Snedeker, 2010; Palmer, 2000). Papafragou et al. (2002) reported a similar finding for children in this age range engaged in a linguistic task: in their study, 4- to 5-year-old English- and Greek-speaking children were at or below chance at identifying changes to manners of motion in black-and-white drawings of motion events, whereas 10- to 12-year-olds and adults performed significantly better.

We do not believe that preschoolers' failure at the memory test in the Linguistic task can be explained by the fact that the memory test was a surprise for participants in the Linguistic task. To rule out this possibility, we ran an additional 10 4-year-olds in the Linguistic task, instructing them both to describe what happened in each cartoon and to watch each cartoon "very carefully" in preparation for questions. Even with advance warning of the memory test, children did not perform differently from chance (50% accuracy) for changes to either event component: Manner changes, 55% (±9%) accuracy ($t = 0.56$, $p = 0.59$ vs. chance); Path changes, 57% (±9%) accuracy ($t = 0.71$, $p = 0.49$ vs. chance).

In the Linguistic task, one can ask whether explicit mention of an event component at the time of encoding

Table 5
Fixed effects from best-fitting multilevel linear model of accuracy on changes to Manner and Path components during the memory phase.

Effect	Estimate	S.E.	z-Value
Intercept	1.99	0.54	3.69*
Age (Adult vs. Preschool)	−0.48	0.71	−0.68
Task (Linguistic vs. Nonlinguistic)	−0.93	0.72	−1.29
Change Type (Manner vs. Path)	−0.77	0.26	−2.98*
Age × Task	2.79	1.06	2.63*

Effects or interactions that do not appear in this model did not reliably improve the fit. Formula in R: DepVar ~ Age * Task + Change Type + (1|Subject) + (1|Event).
* Significant at $p < 0.01$.

Table 4
Proportion accuracy in identifying changes to event components.

	Adults		Preschoolers	
	Nonlinguistic task	Linguistic task	Nonlinguistic task	Linguistic task
Manners	0.67 (±0.07)	0.87 (±0.07)*,%	0.72 (±0.08)*	0.60 (±0.12)
Paths	0.83 (±0.07)*,#	0.93 (±0.04)*	0.81 (±0.06)*	0.65 (±0.12)

Values indicate the mean proportion (± standard error of subject means) of test events for which participants provided correct answers during the memory task, split by the kind of change (Manner, Path) presented. Significance tests for an effect of task and the type of event component changed on memory accuracy were performed using multilevel mixed logit modeling with crossed random intercepts for Subjects and Items.
* Significantly different from chance (0.50) at $p < 0.05$.
\# Significantly different from Manner accuracy in the same task at $p < 0.05$.
% Significantly different from adult performance in the Nonlinguistic task at $p < 0.05$.

Table 6
Fixed effects from best-fitting multilevel linear model of accuracy on Path changes by Path mention.

Effect	Estimate	S.E.	*z*-Value
Intercept	1.21	0.80	1.52
Age (Adult vs. preschool)	2.81	01.31	2.15*

Effects or interactions that do not appear in this model did not reliably improve the fit. Formula in *R*: DepVar ∼ Age + (1|Subject) + (1|Event).
* Significant at $p < 0.05$.

improved memory for that component during the later memory task. It turns out, however, that accuracy on Path changes in the memory test was independent, for both age groups, of whether Path had been mentioned in the corresponding event description. A multilevel logistic model on memory for changes to Path components was carried out as described above, with Age (Adult, Preschool) and Path mention (yes, no) as first-level fixed factors. The best-fitting model included a main effect of Age, reflecting the fact that adults perform more accurately than children in the Linguistic task, but no effect of Path mention or interaction between the two factors (Table 6). A similar analysis for Manner accuracy could not be conducted because of lack of variation in Manner mention in the Linguistic task: adults mentioned Manner 100% of the time, and children mentioned it 96% of the time (see Table 1).

Accuracy during the memory phase for both age groups was also independent of eyegaze patterns during the encoding phase. Separate multilevel logistic models of accuracy on Path and Manner changes were carried out in each of the 1-s windows described in Section 3.2, with Age (adult, preschool), Condition (Linguistic, Nonlinguistic) and Eyegaze (looks to Path regions for Path accuracy, looks to Manner regions for Manner accuracy) as first-level fixed factors. No effects of eyegaze on memory accuracy were found for either event component in any of the time windows.

3.3.2. ANOVA for memory accuracy

Repeated-measures ANOVAs were also conducted on elogit-transformed subject and item means for proportion accuracy on Manner and Path changes during the Memory phase (Change Type), with participant Age (Adult, Preschool) and Task (Nonlinguistic, Linguistic) as independent variables. The results of the subject analysis echoed that of multilevel modeling, with a significant main effect of Change Type ($F(1,38) = 9.674$, $p < 0.005$) and a significant interaction between Age and Task ($F(1,38) = 5.366$, $p < 0.05$). The item analysis showed a significant main effect of Age ($F(1,44) = 9.700$, $p < 0.005$), reflecting the fact that variance in accuracy by preschooler was less across items than across subjects, and a significant interaction between Age and Task ($F(1,44) = 16.900$, $p < 0.0005$). Mean elogit transformed values for accuracy in identifying changes to event components can be found in Table A3 in Appendix A.

4. General discussion

The experiment reported here introduced a novel method for assessing the processes that underlie language production in young children: observing children's eye movements as they examine and prepare descriptions of dynamic events. To accomplish this, we investigated the origins of linguistic omissions in children's speech, focusing on motion event descriptions. Both prior studies (Allen et al., 2007; Özyürek et al., 2008; Papafragou et al., 2006; Papafragou & Selimis, 2010) and our own investigations in the present paper have indicated that English-speaking children are more likely than adults to omit mention of the Path or Manner of motion when describing motion events (e.g., children might say "The boy was skating" rather than "The boy was skating into the net"). Our goal was to probe the mechanisms at work during utterance formulation to uncover the source of these truncated event descriptions—more specifically, to investigate whether linguistic omissions are indicative of (1) developmental differences in on-line attention allocation during event apprehension/conceptualization or (2) limitations related to the language production system or its interfaces in children. Our experiment provided a direct test of the first of these two hypotheses.

Our findings demonstrate that the initial stages of event apprehension are strikingly similar in preschoolers and adults: 4-year-old children and adults compile similar representations of dynamic motion events and, like adults, children of this age show tight relations between attention allocation and utterance preparation. Specifically, eye movements and memory for event components collected from participants in the Nonlinguistic task revealed that the conceptual event representations constructed by adults and 4-year-old children are highly similar. Eye movement patterns in the Linguistic task revealed, moreover, that the process of preparing an utterance directs attentional resources in children and adults in highly similar ways. While viewing our motion events in the Linguistic task, both age groups showed an effect of language planning, increasing their attention to the component that was more central to their description (the Manner, which both age groups tended to encode in event descriptions). Finally, in the Linguistic task, preschoolers' and adults' memory for Paths was independent of whether they had mentioned the corresponding event component in their earlier description of the motion event. Our results thus provide evidence against the hypothesis that gaps in children's linguistic output reflect differences in attention allocation during motion event conceptualization.

Our data offer support instead for the hypothesis that the source of adult-child differences in the content of event descriptions is linked to the developing language production system itself. Within this broad hypothesis, several types of more specific accounts are possible, only some of which are consistent with our data. For instance, our data argue against the possibility that children's omissions stem purely from a lack of awareness of the typological patterns of English in the domain of motion. Although children produced event descriptions that were shorter than those produced by adults, both age groups were more likely to mention Manners than Paths. Furthermore, the groups did not differ in the linguistic positions in which they encoded event components. When they included both Manner and Path in their utterances, both adults and children tended to follow the English pattern of motion event encoding, with

Manner information expressed in the main verb and Path in a satellite of the verb, as in the sentences in (5).

(5) a. Adult
An alien is driving his car into a cave.
b. Preschooler
He was driving inside of a cave.

This coheres with previous studies showing that language-specific trends in the linguistic encoding of motion event components emerge by the age of three (Allen et al., 2007; Naigles et al., 1998; Özyürek et al., 2008; Papafragou et al., 2006; Papafragou & Selimis, 2010; Slobin, 1996; Zheng & Goldin-Meadow, 2002).

Our data leave open several other possible explanations of omission patterns in children's motion descriptions. A first possibility is that children's omissions arise from differences between adults and children that are purely linguistic in nature, perhaps due to differences in the processing mechanisms involved in utterance preparation (e.g., Adams & Gathercole, 2000; Levelt, 1989; McDaniel, McKee, & Garrett, 2010; Speidel & Herreshoff, 1989) or in pragmatic awareness (e.g., Feldman, Goldin-Meadow, & Gleitman, 1978; Papafragou et al., 2006). For example, Papafragou and colleagues (2006) demonstrated that speakers use pragmatic concerns like inferential structure and relevance when choosing which elements in a conceptual representation to include in an utterance. On this account, the observed differences in output found in this study may signal that children and adults were drawing on different pragmatic strategies when deciding what was important to communicate about our motion events: children may have underestimated how much motion information would be optimal to encode, and so produced a higher proportion of underinformative (Manner-only) utterances compared to adults.

A different, perhaps complementary possibility is that children and adults differ at the attention-production interface: that is, despite similar patterns of event inspection across age groups, children may require a higher threshold of attention to a particular event component to encode that component in an utterance. This possibility is consistent with the fact that, in our data, children showed an earlier and more pronounced influence of attention allocation on production than adults did: children in the Linguistic task increased their looks to Manners before adults did. There are two underlying mechanisms that might give rise to this kind of difference. One possibility is that children are more susceptible than adults to talking about whatever captures their attention. If this is the case, then the omissions that we see in this study do not reflect the process of linguistic preparation, but rather are revealing about what children thought was most interesting about individual events. Adults either thought that other things about the events were interesting or were able to suppress their excitement about particular event components in the interest of providing fully informative event descriptions. On the other hand, these results might indeed reflect differences in the processes that underlie linguistic preparation. It may be the case that children have to direct more attention than adults to particular event components to encode those components in utterances, to the potential detriment of event components that they directed less attention to. The relative difficulty of encoding a particular event component may reflect a word retrieval effect, or it may have something to do with the way that the event component is realized in the linguistic representation. The extent to which the linguistic omission findings we presented can be explained by these or additional hypotheses remains an issue for future research.

Acknowledgements

This research was supported in part by Grant #R01-HD055498 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development to A. Papafragou and J.C. Trueswell. Correspondence concerning this article should be addressed to Ann Bunger, Department of Psychology, University of Delaware, 108 Wolf Hall, Newark, DE 19716. Email: bunger@udel.edu.

Appendix A

See Tables A1–A3.

Table A1
Motion events depicted in target videos and sizes of Manner and Path regions used in the eyegaze analysis. Region sizes are given for preschool videos in degrees of visual angle at a viewing distance of 60 cm. Adults viewed the videos under slightly different display settings, resulting in an 11% reduction in the width (but not the height) of the regions for adults as compared to preschoolers.

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

Table A2
Mention of event components in linguistic descriptions.

	Proportion		elogit-Transformed Proportion	
	Manner	Path	Manner	Path
Adults	1.00	0.70 (±0.07)	3.22	1.00 (±0.38)
Preschoolers	0.91 (±0.04)	0.32 (±0.09)*	2.32 (±0.34)	−0.64 (±0.43)*

Values indicate mean proportions and elogit-transformed mean proportions (± standard errors of subject means) of utterances in which Manner and Path components were mentioned. Significance testing for an effect of participant age on event component mention was performed using multilevel mixed logit modeling with crossed random intercepts for Subjects and Items.
* Significantly different from adult production at $p < 0.01$.

Table A3
elogit-Transformed accuracy in identifying changes to event components (compare to Table 5).

	Adults		Preschoolers	
	Nonlinguistic task	Linguistic task	Nonlinguistic task	Linguistic task
Manners	0.75 (±0.32%)	1.80 (±0.35%)	1.01 (±0.40%)	0.33 (±0.53%)#
Paths	1.58 (±0.33%)*	2.11 (±0.24%)	1.42 (±0.33%)	0.75 (±0.55%)#

Values indicate elogit-transformed mean proportions (± standard error) of test events for which participants provided correct answers during the memory task, split by the kind of change (Manner, Path) presented. Significance tests for an effect of Change-Type and Task on memory accuracy were performed using multilevel mixed logit modeling with crossed random intercepts for Subjects and Items.
* Significantly different from Manner accuracy at $p < 0.05$.
# Significantly different from adult performance in the same task at $p < 0.05$.

References

Adams, A.-M., & Gathercole, S. E. (2000). Limitations in working memory: Implications for language development. *International Journal of Language & Communication Disorders, 35*, 95–116.

Allen, S., Özyürek, A., Kita, S., Brown, A., Furman, R., Ishizuka, T., et al. (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.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Barr, D. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*, 457–474.

Bock, K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language, 48*, 653–685.

Croft, W. A. (1991). *Syntactic categories and grammatical relations*. Chicago: University of Chicago Press.

Feldman, H., Goldin-Meadow, S., & Gleitman, L. (1978). Beyond Herodotus: The creation of a language by linguistically deprived deaf children. In A. Lock (Ed.), *Action, symbol, and gesture: The emergence of language* (pp. 351–414). New York: Academic Press.

Fenson, L., Dale, P. S., Reznick, J. S., & Bates, E. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59* (5, Serial No. 242).

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj, II (Ed.). *Language development. Language, thought, and culture* (Vol. 2, pp. 301–334). Hillsdale, NJ: Lawrence Erlbaum.

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

Goldberg, A. E. (1997). The relationships between verbs and constructions. In M. Verspoor, K. D. Lee, & E. Sweetser (Eds.), *Lexical and syntactical constructions and the construction of meaning* (pp. 383–398). Amsterdam: John Benjamins.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11*, 274–279.

Hitch, G. J., Halliday, M. S., Schaafstal, A. M., & Heffernan, T. M. (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.

Huttenlocher, J., Smiley, P., & Charney, R. (1983). Emergence of action categories in the child: Evidence from verb meanings. *Psychological Review, 90*, 72–93.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446.

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

Johnson, M. H. (1995). The inhibition of automatic saccades in early infancy. *Developmental Psychobiology, 28*, 281–291.

Johnston, J. R., & Slobin, D. I. (1978). The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language, 6*, 529–545.

Kahn, M., & Snedeker, J. (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.

Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2007). Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes, 22*, 1212–1236.

Lakusta, L., & Landau, B. (2005). Starting at the end: The importance of goals in spatial language. *Cognition, 96*, 1–33.

Lakusta, L., Wagner, L., O'Hearn, K., & Landau, B. (2007). Conceptual foundations of spatial language: Evidence for a goal bias in infants. *Language Learning and Development, 3*, 179–197.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.

Mandler, J. M. (2004). *The foundations of mind: Origins of conceptual thought*. New York: Oxford University Press.

Manning, C. D. (2003). Probabilistic syntax. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probabilistic linguistics* (pp. 289–341). Cambridge, MA: MIT Press.

McDaniel, D., McKee, C., & Garrett, M. F. (2010). Children's sentence planning: Syntactic correlates of fluency variations. *Journal of Child Language, 37*, 59–94.

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

Özyürek, A., Allen, S., Furman, R., Kita, S., Brown, A., & Ishizuka, T. (2008). Development of cross-linguistic variation in speech and gesture: Motion events in English and Turkish. *Developmental Psychology, 44*, 1040–1054.

Palmer, S. (2000). Working memory: A developmental study of phonological recoding. *Memory, 8*, 179–193.

Papafragou, A. (2010). Source-goal asymmetries in motion representation: Implications for language production and comprehension. *Cognitive Science, 34*, 1064–1092.

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

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

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

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

Piaget, J. (1955). *The child's construction of reality*. London: Routledge and Kegan Paul.

Pruden, S. M., Hirsh-Pasek, K., Maguire, M. J., & Meyer, M. A. (2004). Foundations of verb learning: Infants form categories of path and manner in motion events. In *Proceedings of the 28th annual Boston University Conference on Language Development* (pp. 461–472). Somerville, MA: Cascadilla Press.

Pruden, S. M., Hirsh-Pasek, K., & Golinkoff, R. M. (2008). Current events: How infants parse the world events for language. In T. Shipley & J. Zacks (Eds.), *Understanding events: From perception to action* (pp. 160–192). New York: Oxford University Press.

Pulverman, R., & Golinkoff, R. M. (2004). Seven-month-olds' attention to potential verb referents in nonlinguistic events. In *Proceedings of the 28th annual Boston University Conference on Language Development* (pp. 473–480). Somerville, MA: Cascadilla Press.

Pulverman, R., Sootsman, J. L., Golinkoff, R. M., & Hirsh-Pasek, K. (2003). The role of lexical knowledge in nonlinguistic event processing: English-speaking infants' attention to manner and path. In *Proceedings of the 27th annual Boston University Conference on Language Development* (pp. 662–673). Somerville, MA: Cascadilla Press.

Pulverman, R., Golinkoff, R. M., Hirsh-Pasek, K., & Buresh, J. (2008). Infants discriminate manners and paths in non-linguistic dynamic events. *Cognition, 108*, 825–830.

Pulverman, R., Hirsh-Pasek, K., Golinkoff, R. M., Pruden, S., & Salkind, S. (2006). Conceptual foundations for verb learning: Celebrating the event. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 134–159). New York: Oxford University Press.

Scerif, G., Karmiloff-Smith, A., Campos, R., Elsabbagh, M., Driver, J., & Cornish, K. (2005). To look or not to look? Typical and atypical development of oculomotor control. *Journal of Cognitive Neuroscience, 17*, 591–604.

Slobin, D. I. (1996). From 'thought and language' to 'thinking for speaking'. In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). New York: Cambridge University Press.

Speidel, G. E., & Herreshoff, M. J. (1989). Imitation and the construction of long utterances. In G. E. Speidel & K. E. Nelson (Eds.), *The many faces of imitation in language learning* (pp. 181–197). New York: Springer.

Steiger, J. H., Shapiro, A., & Brown, M. W. (1985). On the multivariate asymptotic distribution of sequential chi-square statistics. *Psychometrika, 50*, 253–264.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.). *Grammatical categories and the lexicon. Language typology and syntactic description* (Vol. III, pp. 57–149). Cambridge: Cambridge University Press.

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

Zheng, M., & Goldin-Meadow, S. (2002). Thought before language: How deaf and hearing children express motion events across cultures. *Cognition, 85*, 145–175.