Cognition 192 (2019) 104020

Contents lists available at ScienceDirect

Cognition

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

Original Articles

Crossing to the other side: Language influences children's perception of event components☆

Haruka Konishi[a,*], Natalie Brezack[b], Roberta Michnick Golinkoff[c], Kathy Hirsh-Pasek[d]

[a] Missouri Western State University, United States
[b] University of Chicago, United States
[c] University of Delaware, United States
[d] Temple University, United States

ARTICLE INFO

Keywords:
Perception of events
Language and thought
Relational term learning

ABSTRACT

Infants appear to progress from universal to language-specific event perception. In Japanese, two different verbs describe a person crossing a *bounded ground* (e.g., street) versus an *unbounded ground* (e.g., field) while in English, the same verb – *crossing* – describes both events. Interestingly, Japanese *and* English 14-month-old infants form categories of Japanese ground distinctions in nonlinguistic events while by 20 months, *only* Japanese-reared infants retain this ability. Five experiments were conducted to investigate the role that language plays in children's ability to form categories of Japanese ground-path distinctions. Experiments 1a and 1b first replicated and extended prior research (Göksun et al., 2011) by showing that 14-month-old English-reared children formed categories of Japanese ground-path while 23-month-old children did not in the presence of general language. Experiment 2a paired a single novel word with different Japanese ground categories and found that language *weakened* 14-month-old infants' categorization abilities. Experiment 2b showed that labeling these event types differentially allowed 23-month-olds to *recognize* the Japanese ground-path distinctions that they otherwise would not have detected. To assess whether language uniquely encouraged categorization of Japanese ground-path in Experiment 2b, two different tones were paired with ground-path categories in Experiment 3. The results of Experiments 2b and 3 suggested that language but not tones encouraged ground-path categorization. This study is among the first to show that language can be used to heighten and weaken children's categorization of "non-native" event components.

1. Introduction

Processing events is crucial for learning relational terms such as verbs and prepositions. Relational term learning requires three steps (Golinkoff & Hirsh-Pasek, 2008; Golinkoff et al., 2002; Hirsh-Pasek, Golinkoff, Hennon, & Maguire, 2004). First, infants must discriminate between the non-linguistic components of actions that words encode. Upon viewing a woman running out of a house, for example, children must detect the *path* (the trajectory, e.g., *exiting* the house) and *manner* (the way in which the action occurs, e.g., *running*) of the action. Second, children must categorize these event components across varying contexts. That is, children must come to recognize that a manner like running is stable even when performed by a new figure (e.g., girl vs. boy) or in a new context (e.g., playground vs. hallway). Third, children must attach a label to the action referent (Golinkoff et al., 2002). Discriminating between different actions, categorizing actions, and attaching labels to action referents contribute to learning relational terms. Yet there is no cross-language agreement regarding which event components are lexicalized or where to draw the line between similar manners like skipping and running. For example, English typically encodes a figure's manner of motion in the verb (e.g., *hopping*) and a figure's path in an optional satellite prepositional phrase (e.g., *around* the tree), whereas a language like Spanish often encodes the figure's path in the verb (e.g., *salir*; exit) (e.g., Slobin, 2001; Talmy, 2000).

Does the way languages encode event components drive our non-linguistic event perception? The notion that language may influence event perception reflects a long-standing language and thought debate. One view holds that the semantic and syntactic structures of language can affect how speakers perceive semantic domains and organize non-linguistic domains of space (e.g., Levinson, 1996; Gumperz & Levinson, 1996; Pederson et al., 1998), time (Boroditsky, Schmidt, & Phillips, 2003), properties (Imai & Mazuka, 2007; Li, Dunham, & Carey, 2009)

☆ This research was funded by grants to the third and fourth authors: From NSF (SBR9615391), NIH (547208), and IES (R305A100215).
* Corresponding author at: Department of Education, Missouri Western State University, Murphy Hall Office Q, St. Joseph, MO 64507, United States.
E-mail address: hkonishi@missouriwestern.edu (H. Konishi).

https://doi.org/10.1016/j.cognition.2019.104020
Received 30 April 2018; Received in revised form 25 June 2019; Accepted 29 June 2019
Available online 25 July 2019
0010-0277/

and objects (Lucy, 1996; Lucy & Gaskins, 2001; 2003). According to this view, language can have persistent effects on specific domains but contemporary researchers agree that language does not affect all aspects of cognition.

An alternative perspective suggests that cognition is universal and there is minimal to no effect of language (Li & Gleitman, 2002; Majid, 2010; Malt, Sloman, Gennari, Shi, & Wang, 1999; Regier, Kay, Gilbert, & Ivry, 2010). This view holds that language only influences perception when the speaker is thinking in that language (*Language-on-language effect*; Fisher & Gleitman, 2002; Li & Abarbanell, 2018; Gennari, Sloman, Malt, & Fitch, 2002; Trueswell & Papafragou, 2010). "Native speakers not only learn and use the individual lexical items their language offers, but also learn the kinds of meanings typically expressed by a particular grammatical category in their language, and come to expect new members of that category to have similar meanings (p. 25) (Fisher & Gleitman, 2002)." These probabilistic patterns of a language may influence how adults perceive objects and events when language is spontaneously recruited, even in non-linguistic tasks (Papafragou, Hulbert, & Trueswell, 2008; Perry & Lupyan, 2013; Slobin, 2001; Trueswell & Papafragou, 2010).

Recent research considers a more nuanced perspective on the language and thought debate: While prior work has focused on adults and older children with relatively established lexicons, it is possible that the effect of language may differ for infants and toddlers. Additionally, prior work has treated all event components synonymously; however, certain categories may vary in the degree to which they are impacted by language input (Choi & Hattrup, 2012). For example, pre-verbal 2.5-month-olds recognize that a containment relation involves one object moving into another larger object that has an opening which is a critical concept for English prepositions like *in* and *on* (Hespos & Baillargeon, 2001). In contrast, some argue that linguistic input may play a role in how early infants form categories of path of motion (Choi & Hattrup, 2012; Pruden, Roseberry, Göksun, Hirsh-Pasek, & Golinkoff, 2013). Although English-reared 7- to 9-month-olds form a category of path (e.g., *over*) when viewing events labeled with a novel word, they fail to form the category without language accompaniment until 13 months (Pruden et al., 2013). These studies suggest that perhaps, containment relations are less influenced by language relative to path of motion, as infants as young as 2.5 months form categories of containment. A more nuanced perspective on the language and thought debate might be to consider whether the effect of language may differ depending on how children perceive event components that are less common in their native language especially when children are first acquiring language (Choi, 2006; Gleitman & Papafragou, 2013; Hespos & Spelke, 2007; Li, Abarbanell, Gleitman, & Papafragou, 2009; Malt & Wolff, 2010; Regier et al., 2010; Whorf, 1956).

2. Japanese ground-path distinctions

To investigate the role of language on the perception of event components, we contrast two languages that categorize event components differently. Here, we focus on the category of *ground,* the reference point of an event's path (e.g., a field or a street), which is encoded differently in English and Japanese. English often conflates manner in the main verb (e.g., run *around* the tree) and English prepositional phrases describe both the trajectory of the figure and spatial properties of the ground object (e.g., *across* implies a stable surface).

In contrast, Japanese tends to encode path information in the main verb. Japanese has two different types of path verbs: *directional path* and *ground path verbs*. *Directional path (DP) verbs* involve the figure's trajectory of motion relative to a reference point (e.g., *iku* 'go' [from somewhere] or *kaeru* 'return' [to someplace]). These verbs define the direction of motion relative to a source or goal and are not specific to the ground on which the motion occurs (Muehleisen & Imai, 1997). Such DP verbs are not necessarily specific to Japanese; they are also commonly expressed in English (e.g., *come*). *Ground-path (GP) verbs* encode the nature of the ground with the figure's trajectory of motion (e.g., *wataru* 'go across,' *koeru* 'go over,' *nukeru* 'pass through'). For example, the Japanese GP verb *wataru* ('go across') implies that the ground (e.g., a bridge) is a surface between two edges that have vertical extent and serve as boundaries. When crossing a ground that is continuous and has no boundaries on the edges (e.g., a field), the Japanese verb *toru* should be used. These GP verbs are restricted in terms of the grounds they describe (Muehleisen & Imai, 1997). The question is how children learn to encode the categorical distinctions of Japanese ground-path verbs.

3. Universal to language-specific event perception

How do children discern the event components their language expresses? While this paper investigates the construct of ground, other research has shown that by the second year of life, children discriminate and categorize manners and paths in nonlinguistic events that are expressed in their language (Gentner & Bowerman, 2009; Gentner, 1982; Mandler, 1992; Pruden et al., 2013; Pruden, Göksun, Roseberry, Hirsh-Pasek, & Golinkoff, 2012; Pulverman, Golinkoff, Hirsh-Pasek, & Buresh, 2008; Pulverman, Song, Hirsh-Pasek, Pruden, & Golinkoff, 2013). Infants may start their language journey with universal, nonlinguistic event constructs that are then impacted by how the ambient language draws attention to specific event components. This process has been referred to as "semantic reorganization" (Göksun, Hirsh-Pasek, & Golinkoff, 2010) and "attentional narrowing," the term we adopt here.

Evidence for attentional narrowing exists in how infants process 'in' versus 'on' in Korean and English. English divides events with containment (*in*) and support *(on)* relations into two semantic categories. By contrast, Korean linguistically marks spatial relations based on whether objects are placed in a tightly fitting or loosely fitting relation using *kkita* ('tight-fit'), *nohta* ('loose-fit' in support relations), and *nehta* ('loose-fit' in containment relations). The verb *kkita* describes a tight-fit relation between objects, cross-cutting the English categories of *put in* and *put on* (Choi & Bowerman, 1991). That is, putting a ring *on* a finger and a hand *in* a glove are both described with the same verb *kkita* ('tight-fit') in Korean (Choi, 2006). English does not mark this tight-fit or loose-fit relation encoded in Korean verbs. However, studies show that 5-month-old English-reared infants discriminate between tight- and loose-fit relations, suggesting that infants are initially sensitive to distinctions that are not typically lexicalized in English (Hespos & Spelke, 2004). English-reared children show weakened sensitivity to tight-fit relations by 29 months because this distinction is not highlighted in their language, while their Korean-reared peers maintain this distinction as it is relevant to their language (Choi, 2006). English-speaking adults show little sensitivity to the tight- versus loose-fit distinction when measured with similarity judgment tasks, but can categorize appropriately when asked to group spatial relations into two categories (Hespos & Spelke, 2004). In sum, infants in their first year of life display sensitivity to a range of event components. They eventually attend more to the event components their native language expresses, and weaken their attention to event components their language does not encode (McDonough, Choi, & Mandler, 2003). This paper asks whether English-reared children experience a similar course of development for the category of *ground* by first noticing distinctions marked in Japanese that they will later ignore.

4. How do Japanese and English infants perceive grounds in events?

Prior research has examined English- and Japanese-reared infants' attention to ground categories. In a nonlinguistic study, Göksun et al. (2011) investigated whether 14- and 19-month-old English- and Japanese-reared infants could detect the difference between bounded versus unbounded instances of Japanese ground-path categories. Infants were

familiarized to a dynamic scene in which a figure (e.g., man) crossed a single ground (e.g., a road) in silence. In Japanese, this event would be described by the verb *wataru*. At test, infants viewed a split-screen video showing either two new scenes from the same category (within-category condition; e.g., a railroad track vs. a street, both *wataru* category members) or two scenes from different categories (across-category condition; e.g., a bridge from the *wataru* category vs. a field from the *toru* category). If infants detected the categorical distinction encoded in Japanese ground-path verbs, infants in the across-category condition would look longer toward the novel ground (relative to familiarization) presented at test while infants in the within-category condition would look equally to both grounds.

Critically, Göksun et al. (2011) found that both 14-month-old English and Japanese infants looked longer to the novel ground type at test *only* in the across-category condition; infants did not distinguish between the two ground types presented in within-category test trials. Such results suggest that both Japanese- and English-reared infants display similar attention to ground-path distinctions in dynamic, non-linguistic events at 14 months. However, at 19 months of age, English-reared infants looked equally to across-category test trials, no longer discriminating between Japanese ground distinctions, while Japanese-reared infants retained these ground distinctions (Göksun et al., 2011). These data suggest that within the span of only 5 months, attentional narrowing occurs in the semantic domain as infants progress from initially discriminating between ground types to only attending to distinctions expressed in their ambient language.

5. The present studies

However, aside from Göksun et al. (2011), no other studies have explored infants' ability to discriminate between or categorize ground-path events in the *absence of language*. Given the novelty of Göksun et al.'s findings, it must be replicated and extended when language is generally drawing attention to events without providing specific labels. Thus, we first ask whether 14-month-old (Experiment 1a) and 24-month-old (Experiment 1b) English-reared infants discriminate between Japanese ground-path categories in the presence of general attention-getting language (e.g, "Wow, look at her!"), thus extending this prior research to examine the effects of general language as opposed to silence on ground-path categorization. Furthermore, little is known about *how* the process of attentional narrowing occurs with event components, though language exposure is assumed to facilitate this process. The present study begins to address these gaps by assessing infants' ability to form ground-path categories and by examining the role that various types of language may play in children's language-specific event perception.

6. The role of language in weakening and heightening attention to event components

6.1. Heightening attention to object and event categories

Linguistic labels facilitate young children's ability to form categories of objects (e.g., Balaban & Waxman, 1997; Booth & Waxman, 2002; Ferry, Hespos, & Waxman, 2010; Nazzi & Gopnik, 2001; Plunkett, Hu, & Cohen, 2008) while nonlinguistic acoustic tones do not (Fulkerson & Waxman, 2007). Similarly, increasing evidence supports language's role in heightening toddlers' attention to various aspects of events. Casasola and Bhagwat (2007) demonstrated that 18-month-old English-reared children formed an abstract categorical representation of support (on) when hearing a novel spatial preposition (e.g., "She puts it *toke* [on]") during habituation but not when hearing a novel count noun (e.g., "It is a *toke*") or when viewing the events in silence. The present studies explore whether using two unique words for two different Japanese ground-path categories helps 24-month-old English-reared children (Experiment 2b) attend to non-native event components. We also investigate whether tones have the same effect as language on 24-month-olds' ground-path categorization (Experiment 3).

6.2. Weakening attention to event components

As children gain familiarity with how their native language encodes events in relational terms, it may be adaptive to *ignore* information about events that are not lexically marked. Choi (2006) found that 18-month-old English-reared children with greater vocabularies relative to their peers and who already produced the English preposition *on* showed decreased attention to the Korean tight-fit/loose-fit relation, while children with smaller vocabularies who did not produce *on* still attended to these distinctions. Göksun et al. (2010) also found that children's level of language proficiency influenced their perception of non-native event distinctions. However, more research is needed to understand whether language *weakens* categorization of event components. Here, we begin to explore this question by investigating whether language can be used to weaken 13-to-15-month-old's ability to form ground-path categories (Experiment 2a). This work has implications for understanding how and why children begin to ignore non-native event distinctions.

7. Experiment 1a: Do 13- to 15-month-old English-reared infants form Japanese ground-path categories in the presence of general language?

Although Göksun et al. (2011) found that infants between 13 and 15 months discriminate between ground-path categories, the generalizability of the findings remain unknown. The present study addresses this point by examining how different types of language and exposure to multiple ground-path categories and exemplars may influence categorization. First, Göksun et al. (2011) familiarized infants to one type of ground-path category (bounded) before test trials, whereas the children in the present study will be shown two types of ground-path categories during Category Exposure trials (bounded and unbounded; see Trial Types below), creating an opportunity for children to form two categories. Second, children will see multiple exemplars of bounded and unbounded grounds during the Category Exposure trials. Providing children with multiple exemplars from each category has shown to promote categorization of objects (Bornstein, Arterberry, & Clay, 2010). Finally, the prior study was entirely *non-linguistic*. The present study extends Göksun et al. (2011) by overlaying general language (e.g., language that draws attention to the videos without labeling the events to promote categorization). We use general language here (Experiment 1a) to later contrast it with specific language that might encourage categorization (Experiment 2a). However, it is possible that general language will reduce infants' detection of the two ground-path categories by calling attention to both types of events in the same way. To examine the effects of general language (Experiment 1a) versus specific language (Experiment 2a) on ground-path discrimination with younger and older infants, we use general language in Experiment 1a.

7.1. Method

7.1.1. Participants

Twenty-four 13- to 15-month-old full-term monolingual English-reared infants (M = 14.1, SD = 0.69; 10 males; 7 Caucasian, 14 missing demographic information; all mid- or high-SES) participated in the study. Three additional children were excluded from the final sample due to fussiness. Two additional infants were excluded, as they were considered to be outliers (see below under Results for the criterion). Infants' vocabulary scores were collected via the MacArthur-Bates Communicative Development Inventory, Short Form, Level I (MCDI; Fenson et al., 1994), a reliable and valid vocabulary measure assessing 8- to 30-month-olds' receptive language and communication development via parental report. Infants' receptive vocabulary ranged from 9 to

67 words (M = 29.7, SD = 15.9).

7.2. Stimuli

7.2.1. Visual stimuli

The events used in all trial types were a series of video clips similar to those used in Göksun et al. (2011) depicting a female actor (average height of 40 pixels) walking across different types of grounds from the left to right side of the screen. *Wataru* grounds (railroad track, road, bridge, running track, and trail) were flat surfaces with bounded edges, such as curbs. *Toru* events were grounds with no obvious boundaries, or continuous planes (soccer field, playground, softball field, hockey field, and meadow). The videos looped twice so that infants saw the actor continually in motion, engaging in the same crossing event two times. Across all trials and conditions, the pace of walking was controlled using a metronome to ensure that the actor crossed each ground in 6 s. Stimuli were videotaped outdoors in a variety of locations. Videos filled the screen during Category Exposure and Category Familiarization trials and were paired, playing side-by-side simultaneously, during Category Salience and Categorization Test trials (see "Trial Types" below).

7.2.2. Auditory stimuli

A female speaker recorded sentences in infant-directed speech using Audacity. The audio prompts were not specific to the crossing event categories (e.g., "Wow, look at her!") and functioned to draw infants' attention to the screen. This audio was played only during Category Exposure trials; Category Salience, Category Familiarization, and Categorization Test trials appeared in silence.

7.2.3. Procedure

Infants sat comfortably on a parent's lap approximately 39 in. from a 24-inch monitor. Parents were instructed to close their eyes and refrain from talking or directing their child's attention. A camera, hidden behind a small hole in a black curtain underneath the television, recorded infants' faces for later eye gaze coding. The video lasted three minutes and 12 s.

7.3. Trial types

The video stimuli were designed to test whether infants formed categories of the two ground types. First, *a priori* preference for either ground type was assessed (one *Category Salience* trial). Then, infants were exposed to both ground types (twelve *Category Exposure* trials). Next, infants were familiarized to a single ground type (three *Category Familiarization* trials). Finally, infants' ground type preference, relative to Category Familiarization, was tested (two *Categorization Test* trials). Based on the finding that infants detect across-category comparison (Göksun et al., 2011), all infants saw across-category comparisons alongside same-category videos during Categorization Test trials (e.g., railroad track – *wataru* vs. field – *toru*). A 3-second baby face video accompanied by the tune "Oh, Susanna" separated each phase of the experiment to renew infants' interest in the video and reorient infants' looking to the center of the screen.

7.3.1. Category salience trial

The two events that later appear simultaneously as the Categorization Test trial (12 secs) were shown side-by-side in silence to determine whether infants had an *a priori* preference for either Test event.

7.3.2. Category exposure trials

Category Exposure trials were designed to show children various exemplars from the *wataru* and *toru* category. Twelve full-screen Category Exposure trials displayed three different grounds from the *wataru* (bounded ground) category and three from the *toru* (unbounded ground) category. Each 6-second clip appeared twice accompanied by the sentence "Wow, look at her!" *Wataru* and *toru* grounds were presented in alternation. Presentation order was counterbalanced across infants such that half of the infants saw a *toru* trial first while the other half saw a *wataru* trial first.

7.3.3. Category familiarization trials

The purpose of the Category Familiarization trials was to reorient infants' attention to a single ground-path category. Three new videos not shown during Category Familiarization trials of either *wataru* or *toru* grounds (counterbalanced across infants) were played in *silence* on the full screen (12 secs each). Half of the infants saw *wataru* grounds and the other half saw *toru* grounds during Category Familiarization trials.

7.3.4. Categorization test trials

Test trials were intended to assess whether infants formed a category of ground path distinctions, by looking at one event longer than the other. A split-screen video showed a novel ground from the *wataru* category (e.g., bridge) and a novel ground from the *toru* category (e.g., field) simultaneously on either side of the screen for 12 s in silence. The same pair of trials was shown twice. Four between-subjects conditions ensured that the side of the screen on which the familiar ground category (relative to the Category Familiarization trials) appeared was counterbalanced.

Thus, within each condition, the Categorization Test trial was identical to the Category Salience trial and contained events infants had not encountered during Category Exposure or Category Familiarization trials. If infants form ground-path categories, they would likely show a preference for one of the two ground types at Test by looking significantly longer to one event over the other. They may show a preference for the out-of-category event (novel ground type) at Test relative to the Category Familiarization trials they saw immediately prior (Fagan, 1984), or infants may instead show a preference for the same ground type seen during Category Familiarization (familiarity preference). Either preference would indicate that infants differentiate between the two ground types. If infants do not categorize ground-path distinctions, they should look equally to both simultaneously-presented Test events.

7.4. Coding and reliability

The study employed the Intermodal Preferential Looking Paradigm (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Golinkoff, Ma, Song, & Hirsh-Pasek, 2013), which measures infants' eye gaze towards side-by-side video events (Salience and Test). Trained research assistants blind to the hypotheses coded offline recordings of infants' visual fixations to each event using Supercoder (Hollich, 2008). Coding was done frame-by-frame at a rate of 30 frames per second. During Exposure and Familiarization trials, infants' looking to the full screen was coded; for Salience and Test trials, visual attention to the left and right side of the screen was coded. The coding was the same in Experiments 1a, 1b, 2a, and 2b. Twenty percent of videos in each experiment were coded by a second person for inter-coder reliability (ranged between r = 0.97 to 0.98).

Because prior literature suggested that infants may show a novelty preference at Test (Fagan, 1984), a novelty-preference score (NPS) was calculated by dividing looking time towards the novel category Test event (relative to Familiarization trials) by the sum of looking towards the novel and familiar category Test events across both Test trials. This proportion of time looking to the novel category is a measure widely used in infant studies (e.g., Göksun et al., 2011; Pruden et al., 2013). Proportions above 0.50 indicate that infants looked longer to the event from the novel category than the event from the familiar category; a proportion below 0.50 indicated increased looking to the event from the familiar category.

7.5. Results

Data were examined for possible outliers by computing of the Category Salience data (see Trial Types below; Pruden et al., 2012, 2013). If infants' z scores were 2 standard deviations above or below the mean, their data were excluded. Two additional infants were excluded by this criterion, a practice that ensured that only infants who examined each of the events during the Category Salience trial were included in the final sample.

7.6. Category salience trial

A paired-sample *t*-test of NPS compared infants' looking time towards the novel category event ($M = 0.44$, $SD = 0.17$) and familiar category event ($M = 0.55$, $SD = 0.15$). Infants did not show a preference for either event during Category Salience $t(23) = 1.5$, $p = .12$, $d = 0.75$ (Fig. 2a). Thus, any difference that occurred at Test could not be attributed to an *a priori* preference for either Test event.

7.7. Category exposure trials

Infants' proportion of looking times (in seconds) to each of the 12 Category Exposure trials were examined to assess their attention over the course of this phase. A 12 (Exposure trial) repeated-measures analysis of variance (ANOVA) revealed a main effect, $F(11,13) = 3.98$, $p = .01$, $\eta_p^2 = 0.77$. As anticipated, there was a significant difference in looking time between the first Category Exposure trial ($M = 0.90$, $SD = 0.05$) and the last Category Exposure trial ($M = 0.80$, $SD = 0.17$), $t(23) = 3.24$, $p = .004$, $d = 0.79$, indicating that infants' looking to the Category Exposure trials declined.

7.8. Category familiarization trials

A similar ANOVA was run with proportion of looking time to the Category Familiarization trials: A 3 (Category Familiarization trial) repeated measures ANOVA yielded no main effect, $F(2,22) = 1.45$, $p = .25$, $\eta_p^2 = 0.11$. The first and last Category Familiarization trial did not significantly differ $t(23) = 1.7$, $p = .10$, $d = 0.43$, suggesting that infants maintained their attention during Category Familiarization trial events.

7.9. Categorization test trials

Infants' proportion of looking towards the novel category during each of the two Test trials were averaged together, resulting in a single NPS. To account for the possible effect of vocabulary on NPS, infants were divided into high versus low vocabulary groups by a median split (Choi, 2006). We also considered the effect of gender, as previous studies have found differences in NPS by gender (Pulverman et al., 2013). Infants' Category Salience preferences were included in Test trial analyses, to ensure that preference for Test events are above and beyond initial preferences during Category Salience. A repeated-measures ANOVA with trial type (trial: Category Salience vs. Test trials) as a within-subject variable, and gender and vocabulary level (high vs. low) as between-subject variables, yielded a main effect of trial such that there was a difference in NPS during the Category Salience and Test phase, $F(1,20) = 12.76$, $p = .002$, $\eta_p^2 = 0.42$, but no interactions with gender or vocabulary. Infants looked longer towards the event from the novel ground-path category ($M = 0.58$, $SD = 0.19$) compared to the event from the familiar ground-path category ($M = 0.41$, $SD = 0.19$) at Test. Further, a one-sample *t*-test found that proportion of looking to the novel category during Test exceeded chance $t(23) = 2.01$, $p = .05$ $d = 0.86$. These results suggest that 13–15-month-olds categorize Japanese ground distinctions in the presence of general language. We next assess whether we can also replicate and extend Göksun et al. (2011) findings by examining whether 21-to 24-month-old English-reared children categorize Japanese ground-path.

8. Experiment 1b: Do 21- to 24-month-old children detect Japanese ground-path categories accompanied by general language?

Extending Göksun et al. (2011), the same stimuli as in Experiment 1a were used with a new group of older children. The study examines whether 21- to 24-month-old children categorize ground-path events accompanied by general language; prior literature suggests that such distinctions might be lost by this age as children gain more experience with their native language (e.g., Choi, 2006; Göksun et al., 2011).

8.1. Method

8.1.1. Participants

Twenty-two 21- to 24-month-old ($M = 22.7$, $SD = 1.4$; 13 males; children's race was not collected at the time of this experiment; all mid- or high-SES) full-term monolingual English-reared children participated. Children's productive vocabulary as measured by the MCDI, Level II ranged from 17 to 100 words ($M = 53.4$, $SD = 25.6$). Two additional children were excluded from further analyses due to fussiness. Following the same outlier procedure as Experiment 1a, 2 additional children were excluded from analyses.

8.2. Procedure

The same procedure, design, and visual and auditory stimuli as Experiment 1a were used (Fig. 1).

8.3. Results

8.3.1. Category salience trial

A paired-sample *t*-test assessed whether children had *a priori* preferences for either event during Category Salience. Children showed equal looking towards the familiar category event ($M = 0.47$, $SD = 0.10$) and the novel event category ($M = 0.51$, $SD = 0.10$), $t(21) = 0.89$, $p = .38$, $d = 0.30$, (Figs. 2.1, 2b).

8.3.2. Category exposure trials

A 12 (Category Exposure trial) repeated-measures analysis of variance (ANOVA) was performed on proportion of looking time to the Category Exposure trials; there was no main effect, $F(11,11) = 1.8$, $p = .16$, $\eta_p^2 = 0.64$. In addition, there was no significant difference in looking time between the first ($M = 0.88$, $SD = 0.10$) and the last ($M = 0.85$, $SD = 0.19$) Category Exposure trial, $t(21) = 0.96$, $p = .34$, $d = 0.19$, indicating that these older children maintained their looking across the 12 Category Exposure trials.

8.3.3. Category familiarization trials

A 3 (Category Familiarization trial) repeated measures ANOVA comparing proportion of looking time to the Category Familiarization trials yielded no main effect, $F(2,20) = 1.5$, $p = .24$, $\eta_p^2 = 0.13$. Children maintained their attention across trials, as there was no significant difference in looking time between the first and third Category Familiarization trial, $t(21) = 1.4$, $p = .16$, $d = 0.39$.

8.3.4. Categorization test trials

A repeated-measures ANOVA with trial type (trial: Category Salience vs. Test trials) as a within-subject variable and gender and productive vocabulary level (high; low) as between-subject variables found no main effect of trial, $F(1,21) = 2.3$, $p = .13$, $\eta_p^2 = 0.09$, nor any interactions, indicating that children looked equally to novel and familiar events during Category Salience and Test. Children looked approximately equally towards the events from the novel ($M = 0.47$, $SD = 0.11$) and familiar ($M = 0.52$, $SD = 0.11$) ground-path categories

Trial	Visual Stimuli	Auditory Stimuli				
	Experiment	1a	1b	2a	2b	3
	Age Range (months)	13-15	21-24	13-15	21-24	21-24
Category Salience (Pre-Test) 12 seconds total 1 split-screen trial played twice	novel bounded vs. unbounded 	Silent				
Category Exposure 6 seconds each 12 single-screen trials (3 novel bounded, 3 novel unbounded) Each played twice	novel bounded 	"Wow, look at her!"	"Wow, look at her!"	"She's walking *toke* the road."	"She's walking *toke* the road."	400Hz tones
	novel unbounded 	"Wow, look at her!"	"Wow, look at her!"	"She's walking *keet* the soccer field."	"She's walking *keet* the soccer field."	800Hz tones
Category Familiarization 12 seconds total 3 single-screen trials each played twice	novel unbounded 	Silent				
Categorization (Test) 12 seconds total 1 split-screen trial played twice	bounded vs. unbounded (same as Category Salience) 	Silent				

Fig. 1. Design and sample stimuli of experiment 1a, 2a, 2b, and 3.

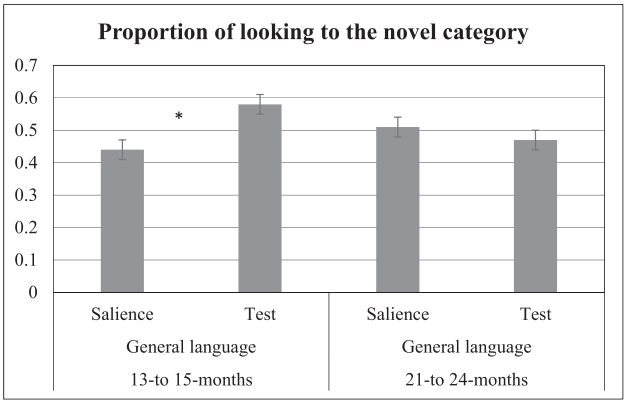

Fig. 2a. 13-to 15-month-old and 21-to 24-month-old children's proportion of looking time to the novel ground during salience and test when general language accompanies events.

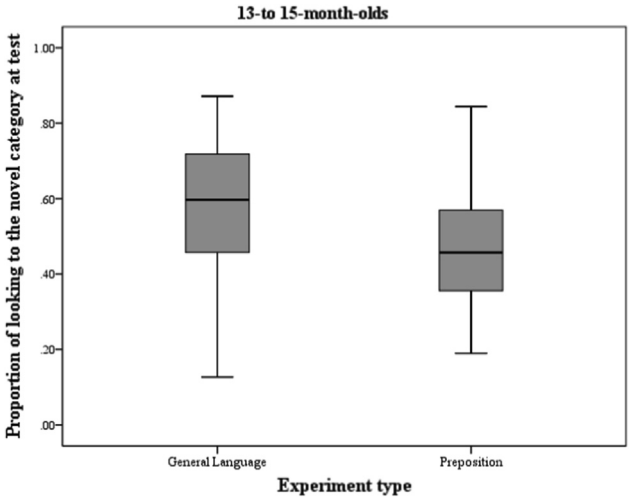

Fig. 2b. 13-to 15-month-olds' proportion of looking to the novel category at test when general language and prepositions accompany events.

at Test. Further, a one-sample *t*-test found that proportion of looking to the novel category during Test did not exceed chance $t(21) = 0.40$, $p = .68$, $d = 0.45$ (Figs. 2a, 2b).

8.3.5. Cross experiment analysis

To investigate the extent to which 13- to 15-month-olds' ground-path categorization skills (Experiment 1a) differ from those of 21- to 24-month-olds (Experiment 1b), a cross-experiment analysis was conducted. A repeated-measures ANOVA with trial type (trial: Category Salience vs. Test trials) as a within-subject variable and age group (13- to 15-month-olds vs. 21- to 24-month-olds) as a between-subjects variable revealed no main effect of trial $F(1,44) = 3.62$, $p = .06$, $\eta_p^2 = 0.07$, but a significant interaction of age group $F(1,44) = 14.29$, $p = .001$, $\eta_p^2 = 0.23$. Post-hoc analysis revealed that 13- to 15-month-olds looked longer to the novel ground type at Test ($M = 0.58$, $SD = 0.19$) than the 21- to 24-month-olds ($M = 0.47$, $SD = 0.11$, $t(44) = 3.78$, $p = .000$).

8.4. Discussion

Experiment 1a examined whether English-reared infants formed categories of Japanese ground-path in the presence of general language. Because Göksun et al. (2011) demonstrated that English-reared infants discriminated between Japanese ground-path categories at 13–15 months, we predicted that infants in this study would perform similarly. Yet, it was also possible that the general language accompanying the events might weaken infants' sensitivity to the ground-path differences by using identical language to call attention to both events. Results showed that infants at 14 months, infants formed categories of

Japanese ground-path. Experiment 1b assessed whether Göksun et al. (2011) findings with children at 18 to 20 months could be generalized here with 21-to 24-month-old children; neither group of older children discriminated between Japanese ground-path categories. The present studies replicate and extend those of Göksun et al. (2011) – even with the inclusion of general language, additional ground-path exemplars, and a different experimental design. This work provides compelling evidence for the developmental trajectory of attentional narrowing where infants progress from universal-to language-specific event perception. However, whether language input drives how infants perceive events in a language-specific way is unknown.

To investigate this question, the next pair of experiments evaluates whether language can be used to weaken and heighten children's categorization of Japanese ground-path. That is, we use a single label to test whether language can *discourage* 13–15-month-old English-reared infants' categorization in Experiment 2a. We then use two different labels to evaluate whether language can *heighten* 21–24-month-olds' categorization of Japanese ground-path in Experiment 2b.

9. Experiment 2a: Can English-reared 13–15-month-olds' detection of Japanese ground-path categories be weakened by pairing a single label with events from different categories?

This study addresses whether hearing a single prepositional phrase paired with exemplars from *wataru* (bounded grounds) and *toru* (unbounded grounds) categories influences infants to collapse the two categories. On this logic, infants who form a single ground-path category should not show a significant preference for either novel or familiar event at Test. Previous studies have found that language heightens children's ability to process event components (such as tight-fit/loose-fit; Casasola, Bhagwat, & Burke, 2009); thus the present study is one of the first to explore whether language can be used to *weaken* categorization of events (Plunkett et al., 2008). Because the literature was effectively silent on the amount of exposure children might need to collapse across the two categories, we sought to determine if 12 Exposure trials would be sufficient to decrease infants' categorization of Japanese ground-path.

9.1. Method

9.1.1. Participants

Twenty-four 13- to 15-month-old full-term monolingual English-reared infants (M = 14.03, SD = 0.95; 10 males; 9 Caucasian, 1 African American, 1 mixed ethnicity, 3 missing demographic information; all mid- or high-SES) were recruited. Infants' receptive vocabulary as measured by the MCDI, Level I ranged from 18 to 85 words (M = 27.81, SD = 16.48). Two additional infants were excluded from further analyses due to fussiness (N = 1). Following the same outlier procedure as previous Experiments, 1 additional child was excluded from analyses.

9.1.2. Procedure

The procedure, design, and visual stimuli were identical to that of Experiment 1a with the exception of the auditory stimuli presented during the 12 Category Exposure trials. The purpose of the Category Exposure trials was to pair a single prepositional phrase with exemplars from the *wataru* (bounded grounds) and *toru* (unbounded grounds) categories with the intent of collapsing these two categories. Additionally, the ground in the video (e.g., field, sidewalk, etc.) was labeled. Infants heard the sentence "Look, she's walking *toke* the ___ (ground)!" (rather than the general language, "Wow, look at her!" used in Experiment 1a). Novel prepositions were used in sentences structured in the format: "Look, she's walking *toke* the ___ (ground)!" for three reasons. First, novel verbs have been shown to be difficult for English-reared children to learn (e.g., Hirsh-Pasek & Golinkoff, 2006). Second, path verbs (such as *across* and *exit*) are less common in English than in other languages (Talmy, 1985), and therefore acquired later (Selimis & Katis, 2010), so using a prepositional phrase rather than a path verb might aid infant comprehension. Third, since the familiar manner verb "walking" was used repeatedly in the Category Exposure sentences, children might focus more readily on the ground rather than the action. Finally, the specific ground being crossed was mentioned in the sentence to rule out the possibility that the novel word might be interpreted as an adverb. For example, the sentence "She's walking *toke*…" might cause children to interpret *toke* as an adverb such as *fast*. The novel word *toke* was selected as Casasola et al. (2009) used this nonsense word as a preposition.

9.2. Results

9.2.1. Category salience trial

A paired-sample t-test revealed that infants did not show significantly longer looking to either of the events; there was no preference for either the familiar event (M = 0.50, SD = 0.16) or the novel event (M = 0.49, SD = 0.16) during Category Salience $t(23) = 0.15$, $p = .87$, $d = 0$.

9.2.2. Category exposure trials

A 12 (Category Exposure trial) repeated-measures analysis of variance (ANOVA) revealed no main effect, $F(11,13) = 1.24$, $p = .34$, $\eta_p^2 = 0.51$. As seen in Experiment 1a with the same age range, there was a significant difference in looking time between the first (M = 0.88, SD = 0.08) and last Category Exposure trials (M = 0.79, SD = 0.22), $t(23) = 2.03$, $p = .05$, $d = 0.54$, suggesting that infants decreased their looking as the Category Exposure trials progressed.

9.2.3. Category familiarization trials

A 3 (Category Familiarization trial) repeated measures ANOVA yielded no main effect, $F(2,22) = 0.31$, $p = .73$, $\eta_p^2 = 0.03$. Infants maintained their looking during Category Familiarization; there was no significant difference between the first (M = 0.70, SD = 0.29) and last trials (M = 0.67, SD = 0.24), $t(23) = 0.35$, $p = .72$, $d = 0.07$.

9.2.4. Categorization test trials

A repeated-measures ANOVA with trial type (Category Salience vs. Test trials) as a within-subject variable and gender and vocabulary level (high; low) as between-subject variables revealed no main effect of trial, $F(1,23) = 0.35$, $p = .55$, $\eta_p^2 = 0.01$ and no interactions. In addition, children's proportion of looking towards the novel ground category (M = 0.48, SD = 0.15) at Test was not significantly different than chance $t(23) = .51p = .45$. Infants looked approximately equally to both ground-path categories during Test trials, suggesting that pairing a single novel label discouraged categorization of Japanese ground-path distinctions.

9.2.5. Cross experiment analysis

To examine the effect of the type of language stimuli (general language vs. novel preposition) on 13- to 15-month-olds' ability to form ground-path categories, we conducted a cross-experiment analysis (Experiment 1a vs. Experiment 2a). A repeated-measures ANOVA with trial type (trial: Category Salience vs. Test trials) as a within-subject variable and experiment type (general language vs. single novel label) as a between-subjects variable revealed a main effect of trial, $F(1,46) = 5.58$, $p = .02$, $\eta_p^2 = 0.11$, and a significant interaction of experiment type $F(1,46) = 9.65$, $p = .003$, $\eta_p^2 = 0.17$. Post-hoc analysis showed that 13- to 15-month-olds who heard general language (M = 0.58, SD = 0.19) looked longer to the novel ground type at Test than the 13- to 15-month-olds who heard a single novel label for both ground types (M = 0.48, SD = 0.15, $t(46) = 3.10$, $p = .003$). This suggests that 13–15-month-olds who heard general language were more adept at forming categories than those who heard a single novel word paired with ground-path categories. We next explore whether language can also heighten attention to non-native event components that

children have already learned to ignore.

10. Experiment 2b: Can 21- to 24-month-old English-reared children's detection of Japanese ground-path categories be heightened through the use of two prepositional phrases paired differentially with each category?

Experiment 1b indicated that 21- to 24-month-old children do not categorize Japanese ground-path events. To examine whether differentiating between ground types could be enhanced with language, we paired one novel spatial preposition (*keet*) with the grounds from the *wataru* (bounded) category and another novel spatial preposition (*toke*) with the grounds from the *toru* (unbounded) category. If novel prepositions encourage children to discriminate between *wataru* and *toru* grounds, they should look longer towards one type of ground-path category at Test. However, if the brief language exposure they receive during Exposure trials is insufficient to strengthen children's detection of ground-path categories, they should show no preference for either ground type at Test.

10.1. Participants

Eighteen 21- to 24-month-old English-reared children (M = 22.3, SD = 1.3; 12 males; 2 Caucasian, 1; 3 missing information on race; all mid- or high-SES) participated. Children's productive vocabulary as measured by the MCDI, Level II, ranged from 8 to 86 words (M = 44.9, SD = 21). Six additional children were excluded from further analyses due to fussiness (N = 4) and caregiver interference (N = 2). Additional data from 5 children were excluded from further analyses as their Category Salience looking times were outliers as defined in prior Experiments.

10.2. Procedure

The procedure, design, and visual stimuli were the same as Experiment 1b (Fig. 1). However, during Exposure, when children saw *wataru* (bounded) exemplars, they heard the sentence "Look, she's walking *keet* the ___(e.g., road, bridge)!" and when they saw *toru* (unbounded) exemplars, they heard, "Look, she's walking *toke* the ___(e.g., field, playground)!" *Keet* was selected as the novel word because Casasola and Wilbourn (2004) used this nonsense word as a preposition. Exposure trials were designed to promote ground-path categorization by pairing different prepositions with *wataru* (bounded) versus *toru* (unbounded) grounds. After Exposure, children who saw three types of *wataru* grounds during Familiarization should look longer at the novel *toru* ground at Test. In contrast, those who saw three types of *toru* grounds during Familiarization should look longer at the novel *wataru* ground at Test. A preference for neither event at Test would suggest that the differential language did not prompt children to form Japanese ground-path categories (Fig. 1).

10.3. Results

10.3.1. Category salience trial

A paired-sample t-test was conducted to assess whether children had *a priori* preferences for either ground type. This test revealed significantly longer looking times towards familiar events (M = 0.56, SD = 0.13), $t(17) = 2.1$, $p = .05$, $d = 1$, suggesting that as a group, children had an *a priori* preference for the familiar category (relative to Familiarization) event. Infants' Category Salience looking time is included in Test trial analyses, as done in all prior Test analyses, to account for *a priori* Category Salience preferences.

10.3.2. Category exposure trials

A 12 (Category Exposure trial) repeated-measures analysis of variance (ANOVA) revealed no main effect of looking time across Category Exposure trials, $F(11,7) = 2.6$, $p = .1$, $\eta_p^2 = 0.80$. However, there was a significant difference in looking between the first Category Exposure trial (M = 0.91, SD = 0.09) and the last Category Exposure trial (M = 0.74, SD = 0.32), $t(17) = 2.3$, $p = .03$, $d = 0.72$, suggesting that children showed a significant decline in looking across the 12 Category Exposure trials.

10.3.3. Category familiarization trials

A 3 (Category Familiarization trial) repeated measures ANOVA yielded no main effect of proportion of looking, $F(2,16) = 0.38$, $p = .68$, $\eta_p^2 = 0.03$. There was no significant difference in looking between the first (M = 0.78, SD = 0.29) and third (M = 0.77, SD = 0.23) Category Familiarization trial $t(18) = 0.11$, $p = .99$, $d = 0.001$, suggesting that children maintained their attention during Category Familiarization trial events.

10.4. Categorization test trials

A repeated-measures ANOVA with trial type (trial: Category Salience vs. Test trials) as a within-subject variable and gender and vocabulary level (high; low) as between-subject variables revealed a main effect of trial and no interactions such that the proportion of time children looked to the novel ground type increased significantly from Category Salience to Test, $F(1,18) = 8.1$, $p = .01$, $\eta_p^2 = 0.36$. A paired-samples t-test comparing children's novelty preference in the Category Salience phase to Test phase confirmed that children looked significantly longer at the event from the novel category at Test (M = 0.59, SD = 0.12) compared to the Familiar category (M = 0.43, SD = 0.13). In addition, a one-sample t-test found that looking time to the event from the novel category at test was significantly above chance $t(17) = 3.3$, $p = .004$, $d = 1.2$. Taken together, these findings suggest that infants displayed a significant preference for the novel ground-path category event at Test, a preference that was significantly different from their initial Category Salience preference (Figs. 3a, 3b).

10.5. Discussion

Experiment 2a investigated whether a single nonsense preposition, paired with two different ground-path categories, would discourage 13- to 15-month-old English-reared infants' detection of Japanese ground-path categories. The findings of Experiments 1a and 2a together demonstrate that (a) 14-month-old English-reared infants form categories of Japanese ground-path, when accompanied by general attention-eliciting language; and (b) a single novel label paired with both ground types appears to weaken 14-month-old infants' categorization abilities. This is one of the first studies to show that language (i.e., a novel label) can reduce infants' existing categorization of non-native event

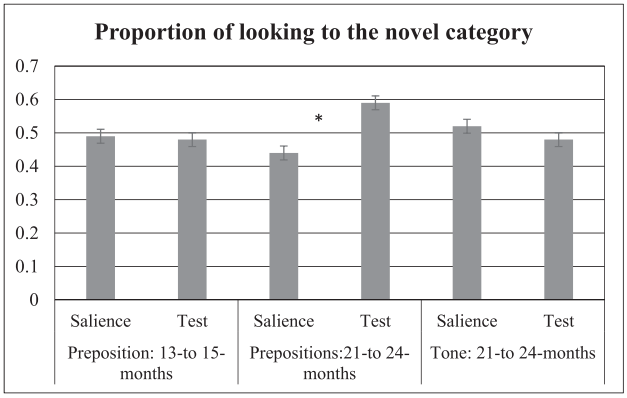

Fig. 3a. 13-to 15-month-old and 21-to 24-month-old children's proportion of looking time to the novel ground during salience and test when novel prepositions and tones accompany events.

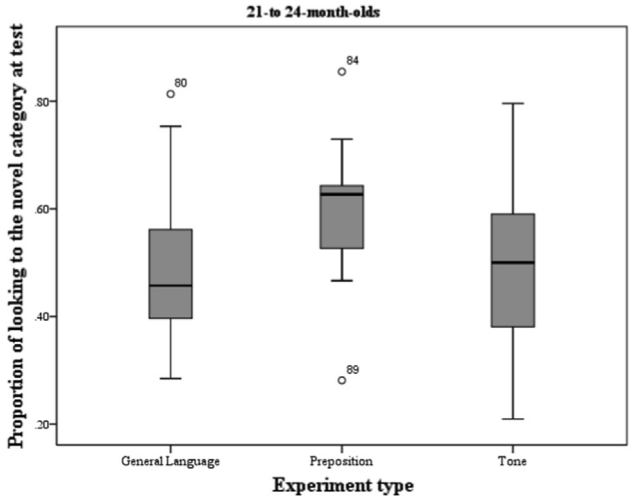

Fig. 3b. 21-to 24-month-old children's proportion of looking time to the novel ground test when general language, prepositions and tones accompany events.

components.

While these experiments have shown that language can *weaken* attention to event categories, an additional question remains regarding whether language can also *enhance* attention to event components that children have learned to ignore. Experiment 1b showed that by 21–24 months, children do not categorize Japanese ground-path when paired with general attentional language; yet when two novel words were paired with *wataru* (bounded) and *toru* (unbounded) grounds, children formed ground-path categories (Experiment 2b). Current findings resemble previous research showing that a novel label facilitates 18-month-old English-reared children's categorization of support relations (i.e., placing one object on another) that they did not demonstrate when the stimuli were seen in silence (Casasola & Bhagwat, 2007). These studies generally demonstrate that language may be key in facilitating or weakening children's attention to event categories. However, it is still not clear if language is unique in its ability to enhance ground-path categorization. Although the facilitative effects of auditorily-presented linguistic labels on object categorization has been established (Balaban & Waxman, 1997; Booth & Waxman, 2002; Ferry et al., 2010), the underlying mechanism is unclear. One interpretation is that language *qua* language promotes categorization in those studies. Another interpretation is that the benefits of language for categorization might rest on the fact that language provides low-level attentional cues.

Further, it is unclear whether the effect of language on *object* categorization functions similarly in *event* categorization. A handful of studies have investigated how language affects categorization of spatial relations (Casasola et al., 2009; Choi, 2006), but it is yet unknown whether tones might facilitate categorization of non-native event components. General, non-differentiating language does not promote categorization of ground-path distinctions in 21–24-month-olds (Experiment 1b) while unique, differentiating language does so (Experiment 2b). However, it is possible that children attended to ground-path differences in Experiment 2b due to a *general auditory* effect. Perhaps a reliable and consistently different auditory stimulus paired with each ground type (such as tones) would be as effective as novel labels in promoting categorization. Indeed, when prior work used labels that were digitally edited such that they could not be identified as count nouns, they also enhanced infants' object categorization (Balaban & Waxman, 1997). Thus, the use of a tone that accompanied the events might promote categorization of ground-path event distinctions. We next examine whether tones matched for duration and frequency to the auditory stimuli in Experiment 2b would enhance 21-24-month-olds' attention to ground-path categories, as differential language had previously. If it is language *qua* language at work, children should no longer categorize ground-path when paired with tones. If true, this would suggest that a mere overlap between auditory and visual stimuli was not sufficient to drive the effect in Experiment 2b (e.g., Robinson & Sloutsky, 2007). Instead, language might be unique in facilitating categorization.

11. Experiment 3: Can 21- to 24-month-old English-reared children's categorization of Japanese ground-path be heightened by pairing each ground type with a different tone?

Here, we examined whether *language* was responsible for encouraging categorization of ground-path distinctions, or whether distinctive nonlinguistic auditory "labels" that matched the paralinguistic features of the language in Experiment 2 might facilitate categorization. If children categorize ground-path in the presence of nonlinguistic tones, they should look longer toward the novel ground type at Test, showing that tones are sufficient for spurring categorization. However, if tones are insufficient for promoting categorization, children should show no preference for either ground type at Test. We hypothesized that if language uniquely facilitates ground-path categorization above and beyond attentional features of the linguistic labels, tones would not increase toddlers' detection of Japanese ground-path categories.

11.1. Participants

Twenty monolingual English-reared 21- to 24-month-old English-reared children (M = 22.6, SD = 1.0, 11 males; 15 Caucasian, 1 African American, 1 mixed ethnicity, 3 missing demographic information; all mid- or high-SES) participated. Children's productive vocabulary as measured by the MCDI, Level II ranged from 10 to 100 words (M = 49.1, SD = 27.8). Seven additional children were excluded from further analyses due to fussiness (N = 2), looking at the screen for less than 50% of the video (N = 1), parental interference (N = 1), and sibling interference (N = 1).

11.2. Procedure

The procedure, design, and visual stimuli were the same as the previous experiments (Fig. 1). However, instead of speech, the exemplars from the *wataru* (bounded) category were accompanied by 400 Hz tones while exemplars from the *toru* (unbounded) category were accompanied by 800 Hz tones (counterbalanced tone-category pairings across children) during Category Exposure trials. The purpose of the Category Exposure trials was to examine whether two distinct tones would also encourage categorization of ground-path events. Tones were matched in frequency, duration, and volume to words in Experiment 2b (Table 1). That is, when children saw *wataru* exemplars, in place of the words from Experiment 2b, "Look, she's walking *keet* the ___(e.g., road, bridge)!" they heard 400 Hz tones at the onset of each word lasting the length of the word. When children saw *toru* exemplars, they heard 800 Hz tones, matched to the timing and duration of each word in a sentence such as, "Look, she's walking *toke* the __(e.g., field, playground)!"

11.3. Results

11.3.1. Category salience trial

A paired-sample t-test of NPS comparing infants' looking time towards the novel event (M = 0.52, SD = 0.15) and the familiar event category (M = 0.48, SD = 0.15) trials revealed that infants did not show a preference for either event during Category Salience, t (19) = 0.63, p = .54, d = 0.14.

11.3.2. Category exposure trials

A 12 (Category Exposure trial) repeated-measures analysis of

Table 1
Example of Experiment 2b sentence matched with Experiment 3 tones. Words and pauses (———) between words in are listed in the first row. Timing in seconds of word and pause onsets and offsets are listed in the second row. Single tones lasting the length of each word, from word onset to offset (either 400 Hz or 800 Hz depending on condition), were used in Experiment 3. The third row shows tones corresponding to the timing listed in the second row in the 400 Hz condition.

Experiment 2b Words and Pauses	Onset-Offset	Experiment 3 tones
Wow	0–0.626	400 Hz
———	0.626–0.892	———
she's	0.892–1.249	400 Hz
———	1.249–1.297	———
walking	1.29–1.699	400 Hz
———	1.699–1.822	———
toke	1.822–2.157	400 Hz
———	2.157–2.235	———
the	2.235–2.468	400 Hz
———	2.468–2.517	———
hockey	2.517–2.904	400 Hz
———	2.904–2.956	———
rink	2.956–3.358	400 Hz

variance (ANOVA) revealed a marginal main effect, $F(11,9) = 3.76$, $p = .07$, $\eta_p^2 = 0.165$. There was a significant difference between the first Category Exposure trial ($M = 0.84$, $SD = 0.07$) and the last Category Exposure trial ($M = 0.70$, $SD = 0.23$), $t(19) = 2.52$, $p = .02$, $d = 0.56$, suggesting that attention to the events declined from the first to the last trial.

11.3.3. Category familiarization trials

A 3 (Category Familiarization trial) repeated measures ANOVA yielded no main effect, $F(2,18) = 0.33$, $p = .72$, $\eta_p^2 = 0.03$. The first and third Category Familiarization trials did not significantly differ, $t(20) = 0.36$, $p = .72$, first trial: $M = 9.3$ $SD = 2.3$, third trial: $M = 9.2$, $SD = 2.7$, suggesting that children maintained their attention during Category Familiarization trial events.

11.3.4. Categorization test trials

A repeated-measures ANOVA with trial type (trial: Category Salience vs. Test trials) as a within-subject variable and gender and vocabulary level as between-subject variables produced no main effect of trial, $F(1,19) = 0.09$, $p = .77$, $\eta_p^2 = 0.006$, and no interactions. In addition, children's proportion of looking towards the novel ground category ($M = 0.48$, $SD = 0.15$) at Test was not significantly different than chance, $t(19) = .17p = .86$, $d = 0.04$ (Figs. 3).

11.3.5. Cross experiment analysis

To determine the effect of the type of stimuli (general language vs. prepositions vs. tones) on 21- to 24-month-olds' Japanese ground-path categorization, we conducted a cross-experiment analysis (Experiment 1b vs Experiment 2b vs. Experiment 3). A repeated-measures ANOVA with trial type (trial: Category Salience vs. average of Test trials) as a within-subject variable and experiment type (experiment: general language vs. prepositions vs. tones) as a between-subject variable yielded no main effect of trial $F(1, 58) = 0.78$, $p = .37$, $\eta_p^2 = 0.01$, but a significant interaction with experiment type $F(1, 58) = 6.7p = .002$, $\eta_p^2 = 0.16$. Children who heard two novel prepositions looked longest at the novel ground-path category, suggesting that children's ability to form categories of ground-path distinctions was greater for those who were exposed to bounded and unbounded grounds paired with two distinct novel prepositions ($M = 0.59$, $SD = 0.12$) than those who heard general language ($M = 0.47$, $SD = 0.11$, $t(38) = -3.24$, $p = .002$) or two distinct tones ($M = 0.48$, $SD = 0.15$, $t(36) = 2.64$, $p = .012$).

11.4. Discussion

Experiment 3 investigated whether pairing tones of different frequencies that replaced each word spoken in the linguistic stimuli in Experiment 2b would prompt English-reared children to form Japanese ground-path categories. If language uniquely influences ground-path categorization, we would *not* expect tones to do so. Supporting our prediction, children exposed to tones did not form categories of Japanese ground-path as children of this age had previously when the categories were paired with differentiating language. Children in this experiment were between 21 and 24 months of age and presumably, judging by the findings from Göksun et al. (2011) and Experiment 1b, do not form categories of Japanese ground-path events in silence or with general attentional language. Findings from Experiment 2b and the present study suggest that English-reared children's categorization abilities can be strengthened using language – even after three minutes of exposure – but not tones. This suggests that above and beyond the attentional function associated with non-linguistic auditory stimuli, *language* facilitates categorization of event components.

12. General discussion

"How do children, starting from an initially equivalent base, become native speakers of their language?" (Bowerman & Levinson, 2001, p. 10). This is the broad question underpinning this set of studies. Since this question was initially asked, research has revealed that children detect components of events that are expressed by the world's languages well before they speak. However, by age 3, children have settled on the components of events their native language encodes (Maguire et al., 2010). Becoming a native speaker seems to involve "thinking for speaking" (Slobin, 2001), or perceiving events in ways consistent with the native language.

The present studies addressed a series of questions related to understanding the Bowerman and Levinson query. The first set (Experiments 1a and 1b) speaks to the issue that infants start out "… from an initially equivalent base." Experiment 1a showed that while Göksun et al. (2011) utilized non-linguistic stimuli, 13- to 15-month-old English-learning infants in the present study categorized Japanese ground-path events even when they were accompanied by general non-differentiating language and additional ground-path categories and exemplars. Experiment 1b then showed that the "initially equivalent base" changes as children are exposed to their ambient language; 21- to 24-month-old children did not form categories of Japanese ground-path in the presence of general language, in line with Göksun et al. (2011). Furthermore, cross-experiment analysis showed that 13-to 15-month olds' ability to categorize Japanese ground-path was significantly greater than that of 21-to 24-month-olds. This set of findings provides compelling evidence for a developmental process of attentional narrowing in the event domain, as we succeeded in replicating prior research using general language and additional exemplars and ground-path categories.

The second set of studies (Experiments 2a and 2b) contribute to our understanding of the process by which infants form categories of events encoded in their native language (Bowerman & Levinson, 2001). Clearly, this process begins early, manifested in infants' perception of events around the time they speak their first words. Both Experiments 2a and 2b showed that language influences categorization of event components – language can discourage category distinctions infants otherwise form (Experiment 2a) and strengthen categories they otherwise ignore (Experiment 2b). Additionally, language is unique in its influence on event categorization; Experiment 3 found that tones did not facilitate categorization as language did in Experiment 2b.

12.1. Universal conceptions to language-specific distinctions

Language has been hypothesized to guide infants' attention to the

event components relevant to their native language (Choi, 2006; Göksun et al., 2011; Hespos & Spelke, 2004). Previous studies used language to heighten children's attention to spatial relations (e.g., *tight-fit*; Casasola et al., 2009) but no prior work has yet attempted to weaken children's ability to categorize event components. Experiment 2a is the first to test this question directly by investigating whether language can be used to discourage 14-month-olds' categorization of Japanese ground-path events. Results showed that 14-month-old English-reared infants fail to form ground-path categories when a single novel preposition was used to describe two different types of Japanese "crossing" events. What occurred in the laboratory may well be analogous to what happens in the world: When infants hear different types of events described with a single word, they may unite these events under a single lexical expression – such as "across." Indeed, Göksun et al. (2010) found that English-reared children with larger vocabularies do not categorize the Japanese ground-path events while their less talkative peers did. Perhaps specific types of words or the amount of language in children's lexicons are associated with reduced ability to form categories of non-native events. Interestingly, infants' detection of non-native events was decreased in the laboratory in a mere three minutes of exposure.

Language not only discourages children's event categorization of event components, it also heightens this ability. As one of the few studies to examine how language promotes categorization (Casasola et al., 2009), Experiment 2b found that 23-month-old English-reared children categorize Japanese ground-path in the presence of two distinct prepositions. These findings suggest that experience hearing unique language paired with different ground-path categories may be crucial in highlighting the relational commonality and differences between event categories.

12.2. Does language uniquely drive language-specific event perception?

A remaining question is whether children's detection of ground-path categories is a language-specific effect or could be facilitated by any non-linguistic auditory stimuli. A broad range of studies in the object categorization literature has identified this finding: Words (including content-filtered words) promoted successful categorization, but non-linguistic sounds failed to have the same effect even when they mimicked the characteristics of speech (e.g., melodies, mechanical sounds, and, in some cases, mouth sounds) (Balaban & Waxman, 1997; Fulkerson & Haaf, 2003; Namy, 2001; Woodward & Hoyne, 1999).

Consistent with the object categorization research, 21–24-month-olds categorized Japanese ground-path only when they heard two novel prepositions paired with each ground type during Category Exposure trials (Experiment 2b) but not when they heard general language (Experiment 1b) or tones (Experiment 3). These results support the language *qua* language approach instead of the low-level auditory explanation suggesting that language is not merely serving as an attention-director. Instead, the two novel prepositions (*toke* and *keet*) may be encouraging children to attend differentially to the properties of the two Japanese ground-path categories (*wataru* and *toru*). Indeed, English prepositions such as *through* encode aspects of the ground; perhaps the use of prepositions signaled children to attend to the characteristics of the ground and encouraged category formation.

Yet, recent work by Ferguson and Waxman (2016) found that certain social cues may cause children to consider tones communicatively, similar to language. Actors in this study first used tones in the context of a social conversation. Following the exchange, infants used the tones to categorize objects. Such social cues appear to be powerful in influencing children's object categorization. Therefore, future research should consider whether distinct tones used in a communicative context may affect ground-path categorization.

12.3. Language and thought

The results of the present study have implications for the language and thought debate; supporting the weaker version of the Whorfian perspective: Japanese ground-path categories may be a universally-available event component for infants but language may direct children's attention to these event components (Bowerman & Choi, 2001; Choi & Bowerman, 1991; Gentner & Boroditsky, 2001; Talmy, 1985). Current findings as well as work by Göksun et al. (2011) provide a new perspective on the language and thought debate by suggesting that linguistic categories can influence event perception prior to the time when children can produce linguistic expressions for these categories. These studies indicate that although English-reared and Japanese-reared 13-to-15-month-olds form Japanese ground-path categories when faced with *non-linguistic events*, only Japanese children maintain this ability at older ages (by 18-to-20-months in Göksun et al., 2011; by 21–24 months in the present study in the presence of minimal general language). These findings suggest that as English-reared children gain language knowledge, they pay less attention to event components such as Japanese ground-path distinctions that are not obligatorily encoded in their native language. A similar pattern is observed with the Korean *tight-fit/loose-fit* distinction: although English-reared children initially detect the *tight-fit/loose-fit* distinction, with increased language knowledge (Choi, 2006), these children no longer make this distinction by 29 months of age (but see Gürcanli, Landau, & Wilson, 2010, for evidence that English speakers can linguistically encode the tight/loose distinction if they choose to). This collection of studies suggests that depending on the type of event component, and whether that event component is routinely encoded in one's native language, language can indeed influence event perception prior to the time that toddlers produce these distinctions in their own language production.

Given that the research on language and thought has primarily focused on adults and older children, it is not surprising that the effects of language on perception may have been underestimated. The malleability of young children's perception of ground-path categories demonstrates how, in a controlled setting, children use the presence of two words or of one word to guide their event category formation. Just as infants can differentiate between non-native phonemes more adeptly than adults, perhaps infants are also primed to note which event categories are marked by language. The present work suggests that the relationship between language and thought may differ depending on when it is assessed and the type of event component investigated. The effect may be seen more starkly when children are first acquiring language and before they can produce much language.

12.4. Malleability of event perception

We have shown that language can enhance young children's categorization of Japanese ground-path (Experiment 2b), suggesting that semantic domains may differ in how readily non-native distinctions can be heightened after attentional narrowing has taken place. The theoretical assumption is that similar to the phonological domain, infants of all linguistic backgrounds may perceive a set of event components that will either be weakened or strengthened as a function of their exposure to the ambient language (Göksun et al., 2010). Should such distinctions not be encoded in their native language, children will then weaken those distinctions. But can adults, presented with these same stimuli, readily learn those categories? To explore this question, we conducted two pilot studies. In both, adults with no exposure to Japanese were shown the same videos used in Experiment 2b in which two novel prepositions were paired with crossing events. Thus, for example, *wataru* (unbounded) exemplar videos were paired with, "Look, she's walking *keet* the ___(e.g., road, bridge)!" while *toru* (unbounded) exemplars were paired with, "Look, she's walking *toke* the ___(e.g., field, playground)!"). After watching the Category Salience, Category Exposure, and Category Familiarization videos, adults were asked to judge which of the two simultaneously presented Test videos were more similar to the Familiarization videos seen previously. In this pilot study, 12 out of 27 (44%) adults correctly chose the familiar, as opposed to the

novel, ground type at Test, which was not significantly different from chance ($t(26) = .28p = .59$)). Because adults experienced difficulty in forming Japanese ground-path categories, a second pilot study with identical stimuli provided 18 adults additional prompting to ensure all adults understood the task. Even with additional prompting about the structure of the task, only 11 out of 18 adults (61%) correctly chose the familiar video at Test, which did not differ from chance ($t(17) = 0.94$, $p = .36$). This provides some preliminary evidence that adults may struggle more than children in categorizing non-native ground distinctions.

While adults in these initial pilot studies did not show evidence of categorizing Japanese ground types, as the 24-month-olds did in Experiment 2b, it is still possible that adults could form categories if given more explicit instructions about the task. For example, if adults were told to that they are going to learn nonsense words in the videos as labels for different events from another language, they may be pushed to form categories. Additionally, introducing the novel words in a verb frame ("She's *toking* the field") may encourage category formation, as adults may have interpreted the novel prepositions in the present study as adverbials modifying the manner of walking. Certainly, adults can perceptually differentiate between ground types; when participants were told what the novel words referred to after the end of the study, all understood the difference between bounded and unbounded grounds. Future studies could elucidate under what conditions adults form categories of ground-path distinctions.

Our pilot data revealed that under certain conditions, English-speaking adults have difficulty categorizing Japanese ground-path distinctions, which may suggest that these categories require more time and exemplars to be noticed by adults. However, under certain conditions, adults have been seen to categorize non-native event components. Hespos and Spelke (2007) showed that when English-speaking adults were familiarized to either a tight- or loose-fit relation in silence and later asked to decide if a new exemplar was similar to the familiarized relation, they grouped exemplars by the tight-fit/loose-fit distinction. Perhaps this relates to the fact that English has many ways of encoding tight and loose fit (e.g., close, snug, and baggy, unconnected) while the same is not true of bounded versus unbounded ground categories, as English does not encode these concepts obligatorily in its relational terms (Landau & Jackendoff, 1993). Similarly, Shafto, Havasi, and Snedeker (2014) reported that English-speaking adults initially interpreted novel verbs to correspond to the manner of motion but can shift to perceiving the path element of these events when given concentrated input (Hespos & Spelke, 2004; Shafto et al., 2014). This may not be surprising given that English also contains path verbs (e.g., exit). However, since English does not make distinctions between surfaces with boundaries versus those without, heightening this distinction may be more difficult. Taken together, these findings may suggest that event components encoded in one's first language (e.g., tight-fit/loose-fit and path verbs) could be easier for adults to learn in a second language than event components (e.g., boundedness) that are not commonly instantiated in the first language. Still, more work is necessary to fully understand how adults may be prompted to categorize non-native event components.

12.5. Implications

Infants' event categorization has implications for language learning. In the phonological domain, individual differences in infants' abilities to discriminate two vowels at 6 months significantly predicts language outcomes at 13, 16, and 24 months of age (Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005; Tsao, Liu, & Kuhl, 2004). An open question is whether a similar pattern may be true for the semantic domain: Do individual differences in infants' abilities to categorize native event components have consequences for their later vocabulary acquisition? To acquire motion verbs like running, infants must recognize that "running" can refer equally well to Carl Lewis circling the track or Grandma running to the car. Thus, categorization of motion events may be a prerequisite to acquiring relational terms. In fact, Konishi, Stahl, Golinkoff, and Hirsh-Pasek (2016) found that individual differences in forming nonlinguistic categories of path and manner at 13- to 15-months uniquely predicted children's knowledge of verbs but not general vocabulary at 27- to 33-months. This preliminary finding suggests that the ability to categorize event components may be a critical step in verb and preposition learning. Future work may consider investigating whether the ability to ignore non-native distinctions such as ground-path early on may relate to children's English vocabulary level, as prior research has identified links between vocabulary level and sensitivity to *tight-fit/loose-fit* distinctions (Choi, 2006).

12.6. Conclusion

To acquire verbs and prepositions, children progress from perceiving events in a language-general to a language-specific way. Although researchers have hypothesized that language influences children's perception of events, few studies have examined this question explicitly. The present study provides compelling evidence for the role that language plays in weakening and heightening categorization of non-linguistic event components. Language exposure may affect event perception, influencing children to pay more attention to distinctions common in their native language and less attention to those that are less prevalent in the ambient language. Additionally, these findings speak to the malleability of the narrowing process in the semantic domain as toddlers heightened their categorization of Japanese ground-path with relatively little exposure, while preliminary results indicate that adults may experience more difficulty. If language-specific biases emerge through a process of narrowing in the semantic domain influenced by language input, infants' early ability to attend to event components may offer us insight into how children become native speakers of their language.

References

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology, 64*(1), 3–26.

Booth, A. E., & Waxman, S. R. (2002). Object names and object functions serve as cues to categories for infants. *Developmental Psychology, 38*(6), 948–957.

Boroditsky, L., Schmidt, L., & Phillips, W. (2003). Sex, syntax, and semantics. In Gentner, & Goldin-Meadow (Eds.). *Language in mind: Advances in the study of language and cognition* (pp. 61–79). Cambridge, MA: MIT Press.

Bornstein, M. H., Arterberry, M. E., & Clay, M. (2010). Infant object categorization transcends diverse object relations. *Infant Behavior and Development, 33*(1), 7. https://doi.org/10.1016/j.infbeh.2009.10.003.

Bowerman, M., & Choi, S. (2001). Shaping meanings for language: Universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman, & S. C. Levinson (Eds.). *Language acquisition and conceptual development* (pp. 475–511). Cambridge: Cambridge University Press.

Bowerman, M., & Levinson, S. C. (2001). *Language acquisition and conceptual development.* Cambridge: Cambridge University Press.

Casasola, M., & Bhagwat, J. (2007). Do novel words facilitate 18-month-old infants' spatial categorization? *Child Development, 78*, 1818–1829.

Casasola, M., Bhagwat, J., & Burke, A. S. (2009). Experience with a label can give a boost. *Developmental Psychology, 45*(3), 711–723.

Casasola, M. M., & Wilbourn, M. P. (2004). 14-month-old infants form novel word spatial relation associations. *Infancy, 6*(3), 385–396. https://doi.org/10.1207/s15327078in0603_4.

Choi, S. (2006). Influence of language-specific input on spatial cognition: Categories of containment. *First Language, 26*(2), 207–232.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition, 4*(1), 83–121.

Choi, S., & Hattrup, K. (2012). Relative contribution of cognition, perception, and language on spatial categorization. *Cognitive Science, 36*, 102–129.

Fagan, J. F. (1984). The relationship of novelty preferences during infancy to later intelligence and later recognition memory. *Intelligence, 8*(4), 339–346.

Fenson, L., Dale, P. S., Reznick, S., Bates, E., Thal, D. J., Pethick, S. J., ... Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5), 1–185.

Ferry, A. L., Hespos, S. J., & Waxman, S. R. (2010). Categorization in 3-and 4-month-old infants: An advantage of words over tones. *Child Development, 81*(2), 472–479.

Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not tones) facilitate object categorization: Evidence from 6-and 12-month-olds. *Cognition, 105*(1), 218–228.

Fulkerson, A. L., & Haaf, R. A. (2003). The influence of labels, non-labeling sounds, and sources of auditory input on 9-and 15-month-olds' object categorization. *Infancy, 4*(3), 349–369.

Ferguson, B., & Waxman, S. R. (2016). What the [beep]? Six-month-olds link novel communicative signals to meaning. *Cognition, 146*, 185–189.

Fisher, C., & Gleitman, L. R. (2002). Language acquisition. In H. F. Pashler (Series Ed.) and C. R. Gallistel (Volume Ed.), Stevens' handbook of experimental psychology learning and motivation (pp. 445–496). New York: Wiley.

Gennari, S. P., Sloman, S. A., Malt, B. B., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition, 83*(1), 49–79. https://doi.org/10.1016/S0010-0277(01)00166-4.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. Center for the Study of Reading Technical Report; no. 257.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman, & S. Levinson (Eds.). *Language acquisition and conceptual development* (pp. 215–256). Cambridge, UK: Cambridge University Press.

Gentner, D., & Bowerman, M. (2009). Why some spatial semantic categories are harder to learn than others: The typological prevalence hypothesis. In J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura, & Ş. Özçalışkan (Eds.). *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* (pp. 465–480). New York, NY: Psychology Press.

Gleitman, L., & Papafragou, A. (2013). New perspectives on language and thought. In K. Holyoak, & R. Morrison (Eds.). *Cambridge handbook of thinking and reasoning*. New York: Oxford University Press.

Göksun, T., Hirsh-Pasek, K., Golinkoff, R. M., Imai, M., Konishi, H., & Okada, H. (2011). Who is crossing where? Infants' discrimination of figures and grounds in events. *Cognition, 121*(2), 176–195.

Göksun, T., Hirsh-Pasek, K., & Golinkoff, R. M. (2010). Trading spaces: Carving up events for learning language. *Perspectives on Psychological Science, 5*(1), 33–42.

Golinkoff, R. M., & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in Cognitive Sciences, 12*(10), 397–403. https://doi.org/10.1016/j.tics.2008.07.003.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language, 14*(1), 23–45.

Golinkoff, R. M., Chung, H. L., Hirsh-Pasek, K., Liu, J., Bertenthal, B. I., Brand, R., & Hennon, E. (2002). Young children can extend motion verbs to point-light displays. *Developmental Psychology, 38*(4), 604–614.

Golinkoff, R. M., Ma, W., Song, L., & Hirsh-Pasek, K. (2013). Twenty-five years using the intermodal preferential looking paradigm to study language acquisition: What have we learned? *Perspectives on Psychological Science, 8*(3), 316–339.

Gumperz, J. J., & Levinson, S. C. (1996). *Rethinking linguistic relativity*. Cambridge, UK: Cambridge University Press.

Gürcanli, Ö., Landau, B., & Wilson, C. (2010, January). Spatial factors that contribute to linguistic choice in English. Paper presented at the 84th Annual Meeting of Linguistic Society of America, Baltimore, MD.

Hespos, S. J., & Baillargeon, R. (2001). Reasoning about containment events in very young infants. *Cognition, 78*(3), 207–245.

Hespos, S. J., & Spelke, E. S. (2004). Conceptual precursors to spatial language. *Nature, 430*, 453–456.

Hespos, S. J., & Spelke, E. S. (2007). Precursors to spatial language: The case of containment. In M. Aurnague, M. Hickmann, & L. Vieu (Eds.). *The categorization of spatial entities in language and cognition* (pp. 233–245). Amsterdam, Netherlands: John Benjamins Publishing Company.

Hirsh-Pasek, K., Golinkoff, R. M., Hennon, E. A., & Maguire, M. J. (2004). Hybrid theories at the frontier of developmental psychology: The emergentist coalition model of learning as a case in point. In D. G. Hall, & S. Waxman (Eds.). *Weaving a lexicon* (pp. 173–204). Cambridge, MA: MIT Press.

Hirsh-Pasek, K., & Golinkoff, R. M. (Eds.). (2006). *Action meets words: How children learn verbs*. New York: Oxford University Press.

Hollich, G. (2008). *Supercoder: A program for coding preferential looking (Version 1.7.1). [computer software]*. West Lafayette: Purdue University.

Imai, M., & Mazuka, R. (2007). Language-relative construal of individuation constrained by universal ontology: Revisiting language universals and linguistic relativity. *Cognitive Science, 31*(3), 385–413. https://doi.org/10.1080/15326900701326436.

Konishi, H., Stahl, A., Golinkoff, R. M., & Hirsh-Pasek (2016). Individual differences in non-linguistic event categorization predict later motion verb comprehension. *Journal of Experimental Child Psychology, 151*, 18–32.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the "critical period". *Language Learning and Development, 1*(3–4), 237–264.

Landau, B., & Jackendoff, R. (1993). Whence and whither in spatial language and spatial cognition? *Behavioral and Brain Sciences, 16*(2), 255–265.

Levinson, S. C. (1996). Frames of reference and Molyneux's question: Cross-linguistic evidence. In P. Bloom, M. Peterson, L. Nadel, & M. F. Garrett (Eds.). *Language and space* (pp. 109–169). Cambridge, MA: MIT Press.

Li, P., & Abarbanell, L. (2018). Competing perspectives on frames of reference in language and thought. *Cognition, 170*, 9–24. https://doi.org/10.1016/j.cognition.2017.09.005.

Li, P., Dunham, Y., & Carey, S. (2009). Of substance: The nature of language effects on entity construal. *Cognitive Psychology, 58*(4), 487–524. https://doi.org/10.1016/j.cogpsych.2008.12.001.

Li, P., & Gleitman, L. (2002). Turning the tables: Language and spatial reasoning. *Cognition, 83*(3), 265–294.

Lucy, J. A. (1996). The scope of linguistic relativity: An analysis and review of empirical research. In J. J. Gomperz, & S. C. Levinson (Eds.). *Rethinking linguistic relativity* (pp. 37–69). Cambridge: Cambridge University.

Lucy, J., & Gaskins, S. (2001). Grammatical categories and the development of classification preferences: A comparative approach. In M. Bowerman, & S. Levinson (Eds.). *Language acquisition and conceptual development* (pp. 257–283). Cambridge, UK: Cambridge University Press.

Lucy, J. A., & Gaskins, S. (2003). Interaction of language type and referent type in the development of nonverbal classification preferences. In D. Gentner, & S. Goldin-Meadow (Eds.). *Language in mind: Advances in the study of language and thought* (pp. 465–492). Cambridge, MA: MIT Press.

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., Imai, M., Haryu, E., Vanegas, S., ... Sanchez-Davis, B. (2010). A developmental shift from similar to language-specific strategies in verb acquisition: A comparison of English, Spanish, and Japanese. *Cognition, 114*(3), 299–319.

Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language, 40*(2), 230–262.

Malt, B. C., & Wolff, P. (2010). *Words and the mind: How words capture human experience*. New York: Oxford University Press.

Mandler, J. (1992). How to build a baby: II. Conceptual primitives. *Psychological Review, 99*(4), 587–604.

McDonough, L., Choi, S., & Mandler, J. M. (2003). Understanding spatial relations: Flexible infants, lexical adults. *Cognitive Psychology, 46*(3), 229–259.

Majid, A. (2010). Words for parts of the body. In B. C. Malt, & P. Wolff (Eds.). *Words and the mind: How words capture human experience* (pp. 58–71). Oxford, NY: Oxford University Press.

Muehleisen, V., & Imai, M. (1997). Transitivity and the incorporation of ground information in Japanese path verbs. In M. H. Verspoor, K. Lee, & E. Sweetser (Eds.). *Lexical and syntactic constructions and the construction of meaning* (pp. 329–346). Amsterdam, Netherlands: John Benjamins Publishing Company.

Namy, L. L. (2001). What's in a name when it isn't a word? 17-month-olds' mapping of non-verbal symbols to object categories. *Infancy, 2*, 73–86.

Nazzi, T., & Gopnik, A. (2001). Linguistic and cognitive abilities in infancy: When does language become a tool for categorization? *Cognition, 80*(3), B11–B20.

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition, 108*, 155–184. https://doi.org/10.1016/j.cognition.2008.02.007.

Pederson, E., Danziger, E., Wilkins, D., Levinson, S., Kita, S., & Senft, G. (1998). Semantic typology and spatial conceptualization. *Language, 74*(3), 557–589.

Perry, L. K., & Lupyan, G. (2013). What the online manipulation of linguistic activity can tell us about language and thought. *Frontiers in Behavioral Neuroscience, 7*, 122. https://doi.org/10.3389/fnbeh.2013.00122.

Plunkett, K., Hu, J., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition, 106*(2), 665–681.

Pruden, S. M., Göksun, T., Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2012). Find your manners: How do infants detect the invariant manner of motion in dynamic events? *Child Development, 83*(3), 977–991.

Pruden, S. M., Roseberry, S., Göksun, T., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Infant categorization of path relations during dynamic events. *Child Development, 84*(1), 331–345.

Pulverman, R., Golinkoff, R. M., Hirsh-Pasek, K., & Buresh, J. S. (2008). Infants discriminate manners and paths in non-linguistic dynamic events. *Cognition, 108*(3), 825–830.

Pulverman, R., Song, L., Hirsh-Pasek, K., Pruden, S. M., & Golinkoff, R. M. (2013). Preverbal infants' attention to manner and path: Foundations for learning relational terms. *Child Development, 84*(1), 241–252.

Regier, T., Kay, P., Gilbert, A. L., & Ivry, R. B. (2010). Language and Thought: Which Side Are You on, Anyway? In B. C. Malt, & P. Wolff (Eds.). *Words and the mind: How words capture human experience* (pp. 165–182). Oxford, NY: Oxford University Press.

Robinson, C. W., & Sloutsky, V. M. (2007). Linguistic labels and categorization in infancy: Do labels facilitate or hinder? *Infancy, 11*(3), 233–253.

Selimis, S., & Katis, D. (2010). Motion descriptions in English and Greek: A cross-typological developmental study of conversations and narratives. Retrieved from: *Linguistik Online, 42*(2)http://www.linguistik-online.de/42_10/selimisKatis.html.

Shafto, C. L., Havasi, C., & Snedeker, J. (2014). On the plasticity of semantic generalizations: Children modify their verb lexicalization biases in response to changing input. *Developmental Psychology, 50*(3), 794–808.

Slobin, D. I. (2001). Form-function relations: How do children find out what they are? In M. Bowerman, & S. C. Levinson (Eds.). *Language acquisition and conceptual development* (pp. 406–449). Cambridge, UK: Cambridge University Press.

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

Talmy, L. (2000). *Toward a cognitive semantics, Vol. 2*. Cambridge, MA: MIT Press.

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

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development, 75*(4), 1067–1084.

Whorf, B. L. (1956). *Language, thought, and reality*. Cambridge, MA: MIT Press.

Woodward, A., & Hoyne, K. (1999). Infants' learning about words and sounds in relation to objects. *Child Development, 70*(1), 65–77.