2-year-olds use syntax to infer actor intentions in a rational-action paradigm

Melissa Kline (mekline@mit.edu)
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology

Jesse Snedeker (snedeker@wjh.harvard.edu)
Department of Psychology
Harvard University

Abstract

Verbs may refer to the means (*I bumped into the lamp*) or outcome (*I broke the lamp*) of an action (cf. Rappaport Hovav & Levin, 2010; Talmy, 1985). Do young children expect language to encode this distinction? Children's imitation patterns suggest that they analyze nonlinguistic events in these terms. When a head-touch is the simplest action available, toddlers include just the outcome, not the means, in their own imitation (Gergely, Bekkering, & Király, 2002). We ask whether syntax influences this inference. An experimenter with her hands occupied made a toy activate with a head-touch, using either Means-focused (*I'm daxing to my toy*) or Outcome-focused language (*I'm daxing my toy*). Toddlers then imitated the action. Means- but not Outcome-focus language encouraged children to include the distinctive head-touch, overriding the standard 'rational imitation' effect. This suggests that toddlers' knowledge of argument structure includes an understanding of a means/outcome divide in verb meaning.

Keywords:, agents, goals, event perception, development, argument structure, verb meaning, imitation

Introduction

Models of language acquisition have primarily focused on two sources of information: the language input a child receives, and the language-specific knowledge or biases she brings to the table. A third important, but less well understood, factor in language learning is the role of non-linguistic knowledge (e.g. social cognition or physical reasoning). While the semantic representations used for language may not be identical to those already available to the pre-linguistic child, non-linguistic cognition provides an important scaffold for early language. Many of the principles guiding children's guesses about noun meaning have parallels in their core knowledge about objects (cf. Bloom, 2000; Landau, Smith, & Jones, 1988; Xu, Dewar, & Perfors, 2009). For instance, young children's noun learning is very sensitive to object shape, but this shape bias for count noun learning operates most strongly in cases where shape is a strong cue to (nonlinguistic) category membership, such as for objects that are solid rather than liquid/amorphous, or with intentional shapes like tools or artifacts (Booth & Waxman, 2002; Soja, Carey, & Spelke, 1991) In contrast, there has been little research on how children's understanding of verb meaning and syntactic structure is shaped by their early knowledge about human agents and their actions and goals.

Conceptually, events can be broken down into smaller components (e.g. *agent, path, cause*), some of which can then be encoded in a sentence (e.g. *I crossed the floor, I danced, I danced across the floor*.). Curiously, many of these components of meaning are also available to prelinguistic children in some form (Gergely et al., 2002; Leslie, 1984; Luo, Kaufman, & Baillargeon, 2009; Phillips & Wellman, 2005; Woodward, 1998). Linguistic theories assume that these pre-linguistic conceptual structures form the basis of adult semantics (cf Csibra & Gergely, 2007; Pinker, 2007; Tomasello, 2003). However, this assumption has not been extensively tested: there is very little work that systematically links early event perception and language acquisition.

The flexibility and power of language emerges from the combination of two components: a lexicon of words (stable associations between concepts and forms) and syntactic rules that govern how words are combined into sentences. Critically, these components must be linked, because syntactic structures constrain the inferences we make about word, and particularly noun, meanings (cf. Gleitman, 1990). Syntactic structures can convey meaning even when the content words are unknown: the sentence *The gorp blicked the dax to the wug* expresses an event with three participants that has something to do with transferring something. Likewise, individual verbs call out for particular kinds of sentences, clustering in meaning-based groups: We can *dance/run/skip into the kitchen*, but we *enter/exit/approach the kitchen* (c.f. Levin, 1993). We can both *push* (means) or *break* (outcome) *a table*, but broadly speaking, distinct verb meanings yield distinct syntactic patterns (for instance, though both *break* and *touch* transitives are grammatical, we can say *the table broke*, but not **the table pushed*). Understanding these relationships has been a central

goal of linguistic theory since the 1960s (c.f. Fillmore, 1968; Goldberg, 2006; Hale & Keyser, 1993).

Here, we focus on a single, very abstract distinction that runs through the verb lexicon and applies to predicates spanning a wide range of meanings. Specifically, we examine the distinction between *what* happens (the outcome), and *how* it comes about (the means or manner). This distinction is promising because it plays a key role in the syntactic realization of verbs, and it appears to mirror a critical distinction in infants' prelinguistic understanding of goals (i.e. outcomes) and rational action. By 12 months old, infants know a great deal about people's actions and their goals: they make predictions about what an agent's goal is and the actions she will take to achieve it, by reasoning about how efficient a means of action is for reaching a particular goal (Gergely et al., 1995; Phillips & Wellman, 2005; Woodward, 1998).

Substantial linguistic evidence suggests that there is a broad means/outcome distinction in verb meaning: a nonstative verb may refer to the means of action (*I bumped into the lamp*) or to its outcome (*I broke the lamp*), but it has been argued that a verb cannot refer to both (Talmy, 1985; Rappaport Hovav & Levin, 2010). This division appears across languages and is not limited to a single semantic field: verbs for means include manners of motion, action, and instruments (*run, push, hammer*), and outcome verbs include both paths and change-of-state results (*enter, break.*) Often these verbs can be expanded to a predicate that describes both a means and an outcome, but with only one element lexicalized in the main verb, for instance:

(1)   a. I ran into the room
        b. I entered the room running
(2)   a. I hammered the vase into pieces
        b. I broke the vase with a hammer

Levin & Rappaport Hovav (2010) propose that the distinction between manner and result (means and outcome) verbs is instantiated at the level of argument structure (the semantic representation of verbs and their arguments), with two distinct categories of verbal roots that can be inserted in different semantic frames. Are young children in fact aware of the distinction between manner and result, and do they associate these verb classes with different sentence structures?

Most prior investigations of the link between language and conceptual representation of events have relied on linguistic analyses, rooted in judgments of sentence meaning and grammaticality (cf Dowty, 1991; Jackendoff, 1983; Levin & Rappaport Hovav, 2005; Talmy, 1985). The existing work on children's initial representations of means and outcomes for

language has focused on the motion domain (i.e. manner-of-motion vs. path-of-motion, cf. Pulverman, Song, Hirsh-Pasek, Pruden, & Golinkoff, 2013, though see Bunger & Lidz, 2014). Only recently has adult psycholinguistic research attempted to address the question of whether motion verbs like these are related to means/outcome verbs in other domains (such as Change-of-State, e.g. *push* vs. *break*). Initial findings indicate that participants trained on an outcome verb category (e.g. change-of-state results) are more likely to prefer outcome interpretations for verbs from a different domain (i.e. path verbs for motion events). Likewise people trained on manner-of-<u>motion</u> verb categories show a subsequent preference for manner-of-<u>action</u> verbs (Geojo, 2014), suggesting an overarching 'means' category. Here, we test the hypothesis that a general conceptual distinction between means and outcome is central for human language, guiding infants' initial expectations about how syntax predicts meaning. In particular, we establish how older infants use these event perspectives for language by studying how they use information provided by syntax to shift between means and outcome perspectives on goal-directed events.

To answer these questions, we make use of existing paradigms in cognitive development for understanding how children represent and make inferences about events. In a classic study of children's early understanding of goal-directed actions, Gergely, Bekkering & Király (2002) presented toddlers with a novel event – an experimenter touched a toy with her head, making it light up. In one condition, there was an explanation for this event (her hands were occupied.) In the other, her hands were available. When the experimenter could have chosen to use her hands, children produced head-touch imitations, but when her hands were occupied, the children touched the toy with their hands. This finding suggests that toddlers can analyze events in terms of theirs means and outcomes: when the experimenter performed the simplest action available to her, children inferred that the outcome, but not the specific means, should be imitated. When she freely choose a more difficult means, children inferred that this means was important to her intention. In the current study, we ask whether syntax influences children's inferences about the relevance of means and outcomes.

In the study below, we explore infants' expectations for events accompanied by either Means-focus or Outcome-focus language. The two conditions use novel verbs in frames known to bias English-speaking adults and children toward means (e.g. *I'm daxing to*

the toy) or outcome (*I daxed the toy*) (Hohenstein, 2005; Slobin, 2006). The latter sentence is somewhat ambiguous - we can both *push the table* (means) and *break the table* (outcome). Critically though, there is a strong association across languages between transitive sentences and causal (outcome) verbs, even though English allows for a wide range of verbs in this sentence context (Levin & Rappaport Hovav, 2005). This typological association suggests that a link between transitivity and outcome semantics may be especially robust.

Here we expect that children will begin with a relatively simple mapping that reflects the broad patterns present in adult language. Specifically, we test two hypotheses about how form-to-meaning mappings might emerge from early conceptual knowledge. First, infants might begin with an early expectation that syntax will mark the difference between means and outcome verbs; this would lead them to make different predictions about event interpretation following different syntactic structures. On this hypothesis, we would expect that Means-focused, but not Outcome-focused, language would encourage children to interpret *dax* as a means verb and include the distinctive head-touch, overriding the 'rational imitation' effect. On the other hand, infants might have only a general mapping between events and verbs, focusing on simpler correspondences like events versus objects, or between one-entity versus two-entity events (all events in the present study have two participants, the experimenter and the toy) (Yuan, Fisher, & Snedeker, 2012). In this case, any preferences about which event perspectives would then arise from their expectations about the important parts of events - focusing on goals rather than the means by which they are achieved (Phillips & Wellman, 2005; Woodward, 1998). If this is the case, they might initially expect *any* novel verb to map to event outcomes, and only learn that syntax can mark this distinction at a later point in development.

Experiment

Methods

Participants 30 children ages 1;7-2;11 (mean age 2;2) participated at the Boston Children's Museum. An additional 24 children were not included in the analysis, because they were uncooperative or did not finish the study (6), because of experimenter error (5), or because they did not make any imitation attempts (13). This high dropout/refusal rate may be due to the
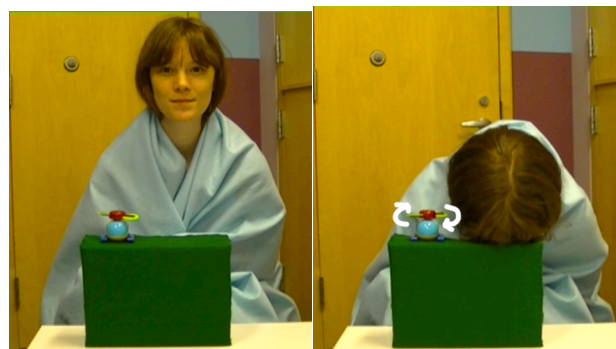

Figure 1: The novel event seen by all study participants

very short session in the unfamiliar museum environment; children had only 2-3 minutes to interact with the experimenter before being asked to imitate her. All children remained with their parent or guardian throughout the session, and received a sticker as thanks for their participation.

Stimuli In place of the 'light up' box used by Gergely et al, participants interacted with a small green box with a toy helicopter on top (see Figure 1). The helicopter could be covertly activated (making its propeller blades spin) by the experimenter using a button/handle on the back side of the box. No children discovered the true mechanism of the toy during the experiment, although some children (or their older siblings) found the button after the conclusion of the experiment. The experimenter also used a blue blanket to wrap her shoulders, arms, and hand during the demonstration phase.

Procedures This study is based on the 'hands-occupied' version of the rational-imitation study conducted by Gergely et al. As a warmup, participants first played a short 'Simon Says' game in which they were invited to wave, clap their hands, and touch their head. Then the experimenter claimed to be cold and wrapped a blanket around herself, leaving her hands occupied. She then said *Watch this, I'm gonna dax to my toy* (Means-focus) or *Watch this, I'm gonna dax my toy* (Outcome-focus) and leaned over to touch the toy with her head, activating the helicopter toy on contact (Figure 1). She then lifted her head and repeated the description (*Wow, I daxed (to) my toy!*) The whole demonstration was repeated once, including the two sentences modeling the novel verb in the appropriate frame. Finally, the experimenter offered the toy to the child, saying, *Now it's your turn! Can you dax (to) the toy*? The box was passed to the child, who was allowed to play with it and attempt the action (note
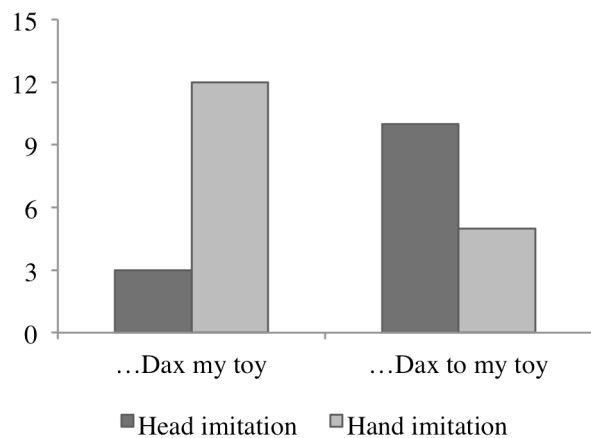
Figure 2: Histogram of actions after Outcome-focus (*Dax my toy*) and Means-focus (*Dax to my toy*) descriptions

that this means that the experimenter could no longer activate the box.) Neutral prompts like *It's your turn, can you show me?* and up to two more repetitions of the target sentence were given until the child made contact with the toy or indicated that they wanted to leave the study. All sessions were recorded on video for later analysis.

Results

In contrast to other versions of the 'light box' imitation studies, we coded the infants' initial actions only. This was for two reasons: first, because there was no interruption between the demonstration and the performance of the imitation, most children quickly performed an intentional action toward the box. Second, because the experimenter did not actually activate the helicopter (and because pilot testing revealed that some children realized that they could spin the helicopter propellers by grabbing them with their hand and twisting), actions toward the toy after the first action reflected new information the child had generated about how the toy worked.

This first intentional action toward the box was coded as to whether not it included contact between the participant's head and the top of the box; a simpler 'first body-part contact' rule was not used because some children discovered that they could pick up the (lightweight) box and bring it to their foreheads, which was coded as a 'head touch' action. Coding was performed by the first author or by research assistants who where not blind to hypothesis, and then by a second coder who was blind to condition assignment. Inter-rater agreement was very high (97% agreement, the single disagreement was resolved by discussion.)

We found that children in the Means-focus language condition were significantly more likely to produce a head-touch than children in the Outcome-focus condition (Figure 2, Fisher's exact test, $p < 0.05$)[1]. When syntax supported an outcome verb interpretation (*Daxing my toy*), children rarely included the head-touch. However, when syntax supported a means interpretation (*Daxing to my toy*), 2-year-olds inferred an appropriate verb meaning and reinterpreted the experimenter's action as relevant for their own imitation. This finding is particularly striking given that since the experimenter's hands were occupied, this unusual action could have been 'explained away' as the simplest means to an end, as it was in the original (nonlinguistic) study (Gergely et al 2002).

Discussion

This work provides significant new insight into what toddlers know about the meanings of the verbs and sentence structures they are just beginning to use. We show that even from relatively early in language acquisition, children are making rational inferences by combining what they know about the nonlinguistic world with what they know about the syntax of their native language.

Prior research has largely focused separately on toddlers' awareness of categories and internal structure of events, on the one hand, and their awareness of syntax and verbs as a syntactic category on the other. In principle, it might be the case that toddlers start out with a very coarse mapping between these two kinds of representations, expecting for instance that verbs map to events rather than to objects, but not realizing that particular verbs (with accompanying argument structures) map to specific events or sub-event features until later in development. In contrast, these results suggest that 2-year-olds are aware that either means or results may be labeled by a main verb, but that these types of verbs tend to appear in different syntactic contexts. Furthermore, they know that these same semantic categories in language are linked to their rich cognitive system for understanding goal directed actions.

The present results are also relevant for understanding the cognitive basis of rational imitation. Gergely et. al (2002) interpreted their imitation finding to mean that infants are truly thinking about a tradeoff between goals and other actions when they see and

---

[1] This comparison is also significant if the 13 children who completed the session without fussing out, but who refused to perform any actions with the toy, are included as a third 'no action' categorical response. There was no difference in imitation-attempt

imitate the 'strange manner' events used in this study. Some challenges have been offered to this interpretation, proposing instead that these differences in imitation result from lower-level perceptual differences or limitations on children's own facility with different kinds of body-bends (Beisert et al., 2012; Paulus, Hunnius, Vissers, & Bekkering, 2011). However, while such alternative explanations may play a role in the size of the effect of the main hands-free/hands-occupied difference, there are several other studies showing that changes in the social context (i.e. whether the action is demonstrated for the child or for someone else) can influence which imitations children are likely to make (Chen & Waxman, 2013; Király, Csibra, & Gergely, 2013). The present study, as well, adds to the body of evidence that children imitations are not purely the result of physical limitations or perceptual salience (all children in this study were of equivalent age and saw identical demonstrations). Instead, information carried by syntax alone modulated children's imitations in a way that is consistent both with what is known about argument structure semantics and with the rational-action theory of children's imitations.

Much work will remain to understand the role of language in young children's interpretation of the social world and other people's actions. The language children heard in this study may be serving several purposes – in addition to conveying semantic information through the structure of the sentence, it may also serve to indicate to the child that the experimenter is knowledgeable about the toy, and that they intend to teach something to the child, factors which are known to influence how children imitate in this kind of context (Bonawitz et al., 2011; Chen & Waxman, 2013). Converging work from imitation and other paradigms will be necessary to better understand how children determine the meanings of new verbs, and then integrate them with the nonlinguistic and social contexts of those sentences.

In addition, further work remains to determine more closely the exact inferences that young children are making about the meaning of the verb *dax* in different sentential contexts. Previous studies of young children's understanding of manner and path verbs (a subset of the means/outcome distinction) have established that infants begin to notice these event categories at around 10-12 months old, (cf. Pulverman et al., 2013), and that by 4 years old children's expectations about neutrally presented novel verb meanings mirror the proportions found in their different native languages (cf. Maguire et al., 2010). In

between, how do toddlers represent the meanings of verbs and sentences? This experiment finds that when presented with a prepositional phrase (*dax to my toy*), children are more likely to consider the means of action to be part of the intended meaning. However, we do not know whether they simultaneously believe that the actual outcome (the toy lighting up) is *not* part of the verb's meaning. The theory of lexical semantics proposed by Rappaport Hovav & Levin (2010) involves a lexicon of verbs with a *complementary* distribution of meanings: a verb can either refer to a means or an outcome, but not both[2]. The persistence of infants' attempts to light the box after it initially fails might provide a measure of whether they believe that a sentence meaning includes the outcome (i.e. predicting that children who hear *dax the box* would persist longer in attempting to make the outcome take place). In future work, we hope to explore the developmental trajectory of the finding in the present experiment, and to discover whether children are aware that verbs label either *just* the means or *just* the outcome.

This line of work represents an important advance for the study of both language acquisition and early cognition. Young children are able to draw on many sources of information to understand goal directed action; they can not only use language as a general communicative signal, but also use syntax as a specific guide to another person's intentions. For language learning, these results suggest that children's early knowledge of argument structure includes an understanding of a fundamental means/outcome divide in verb meaning and of how this divide is syntactically encoded.

References

Beisert, M., Zmyj, N., Liepelt, R., Jung, F., Prinz, W., & Daum, M. M. (2012). Rethinking "Rational Imitation" in 14-Month-Old Infants: A Perceptual Distraction Approach. *PLoS ONE*, *7*(3), e32563.

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

---

[2] There are a small number of proposed exceptions to this rule; Levin & Rappaport Hovav respond that these examples are cases of polysemy. In either case the means/outcome divide is a strong regularity in verb lexicons if not an exception-less rule.

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, *120*(3), 322–30. doi:10.1016/j.cognition.2010.10.001

Booth, A. E., & Waxman, S. R. (2002). Word learning is "smart": evidence that conceptual information affects preschoolers' extension of novel words. *Cognition*, *84*(1), B11–B22. doi:10.1016/S0010-0277(02)00015-X

Bunger, A., & Lidz, J. (2004). Syntactic bootstrapping and the internal structure of causative events. In *Proceedings of the 28th Annual Boston University Conference on Language Development*. Boston, MA: Cascadilla Press.

Chen, M. L., & Waxman, S. R. (2013). "Shall we blick?": Novel words highlight actors' underlying intentions for 14-month-old infants. *Developmental Psychology*, *49*(3), 426–431. doi:10.1037/a0029486

Csibra, G., & Gergely, G. (2007). "Obsessed with goals": Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, *124*(1), 60–78.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, *67*(3), 547–619.

Fillmore, C. (1968). The case for case. In E. Bach & R. T. Harms (Eds.), *Universals in Linguistic Theory* (pp. 1–90). New York: Holt Rinehart and Winston.

Geojo, A. (2014). *Breaking and Entering: Verb Semantics and Event Structure* (Doctoral dissertation). Harvard U.

Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: Rational imitation in preverbal infants. *Nature*, *415*(6873), 755–755.

Gergely, G., Nadasdy, Z., Csibra, G., & Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*, 165–193.

Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*, 3–55.

Goldberg, A. E. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford University Press.

Hale, K., & Keyser, S. (1993). On argument structure and the lexical expression of syntactic relations. In K. Hale & S. Keyser (Eds.), *The view from Building 20: Essays in honor of Sylvain Bromberger*. Oxford: Oxford University Press.

Hohenstein, J. M. (2005). Language-Related Motion Event Similarities in English- and Spanish-Speaking Children. *Journal of Cognition and Development*, *6*(3), 403–425.

Jackendoff, R. (1983). *Semantics and Cognition*. Cambridge, MA: MIT Press.

Király, I., Csibra, G., & Gergely, G. (2013). Beyond rational imitation: Learning arbitrary means actions from communicative demonstrations. *Journal of Experimental Child Psychology*, *116*(2), 471–486. doi:10.1016/j.jecp.2012.12.003

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*(3), 299–321.

Leslie, A. M. (1984). Spatiotemporal continuity and the perception of causality in infants. *Perception*, *13*, 287–305.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.

Levin, B., & Rappaport Hovav, M. (2005). *Argument realization*. Cambridge, UK: Cambridge University Press.

Luo, Y., Kaufman, L., & Baillargeon, R. (2009). Young infants' reasoning about physical events involving inert and self-propelled objects. *Cognitive Psychology*, *58*(4), 441–486. doi:10.1016/j.cogpsych.2008.11.001

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. doi:doi: DOI: 10.1016/j.cognition.2009.10.002

Paulus, M., Hunnius, S., Vissers, M., & Bekkering, H. (2011). Imitation in infancy: rational or motor resonance? *Child Development*, *82*(4), 1047–1057. doi:10.1111/j.1467-8624.2011.01610.x

Phillips, A., & Wellman, H. (2005). Infants' understanding of object-directed action. *Cognition*, *98*, 137–155.

Pinker, S. (2007). *The Stuff of Thought: Language as a Window into Human Nature*. London: Penguin Books.

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.

Rappaport Hovav, M., & Levin, B. (2010). Reflections on manner/result complementarity. In E. Doron, M. Rappaport Hovav, & I. Sichel (Eds.), *Syntax, Lexical Semantics, and Event Structure* (pp. 21–38). Oxford, UK: Oxford University Press.

Slobin, D. (2006). What makes manner of motion salient? Explorations in linguistic typology, discourse, and cognition. In M. Hickman & R. M. Stephane (Eds.), *Space in languages: Linguistic systems and cognitive categories* (pp. 59–81). Amsterdam: John Benjamins.

Soja, N. N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: object terms and substance terms. *Cognition*, *38*(2), 179–211.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. *Language Typology and Syntactic Description*, *3*, 57–149.

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1–34.

Xu, F., Dewar, K., & Perfors, A. (2009). Induction, overhypotheses, and the shape bias: In B. M. Hood & L. R. Santos (Eds.), *The Origins of Object Knowledge* (pp. 263–284). Oxford University Press.

Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: Simple structural cues to verb meaning. *Child Development*, *83*, 1382–1399.