The Developmental Origins of Syntactic Bootstrapping

Cynthia Fisher, a Kyong-sun Jin, b Rose M. Scott c

a Department of Psychology, University of Illinois at Urbana-Champaign
b Department of Psychology, Sungshin Women’s University
c Psychological Sciences, University of California Merced

Received 18 June 2018; received in revised form 4 June 2019; accepted 4 June 2019

Abstract

Children use syntax to learn verbs, in a process known as syntactic bootstrapping. The structure-mapping account proposes that syntactic bootstrapping begins with a universal bias to map each noun phrase in a sentence onto a participant role in a structured conceptual representation of an event. Equipped with this bias, children interpret the number of noun phrases accompanying a new verb as evidence about the semantic predicate–argument structure of the sentence, and therefore about the meaning of the verb. In this paper, we first review evidence for the structure–mapping account, and then discuss challenges to the account arising from the existence of languages that allow verbs’ arguments to be omitted, such as Korean. These challenges prompt us to (a) refine our notion of the distributional learning mechanisms that create representations of sentence structure, and (b) propose that an expectation of discourse continuity allows children to gather linguistic evidence for each verb’s arguments across sentences in a coherent discourse. Taken together, the proposed learning mechanisms and biases sketch a route whereby simple aspects of sentence structure guide verb learning from the start of multi-word sentence comprehension, and do so even if some of the new verb’s arguments are omitted due to discourse redundancy.

Keywords: Syntactic bootstrapping; Word learning; Discourse

Correspondence should be sent to Kyong-sun Jin, Department of Psychology, Sungshin Women’s University, Seoul 02844, Republic of Korea. E-mail: kyongsun.jin@gmail.com and Cynthia Fisher, Department of Psychology, University of Illinois, Champaign, IL 61820. E-mail: clfishe@illinois.edu
1. Introduction

Infant language-learners receive input consisting of word sequences paired with world situations. Based on these data, infants learn to understand some words before they are 1 year old (Bergelson & Swingley, 2012; Tincoff & Jusczyk, 2012), and they begin to understand multi-word sentences early in their second year (Hirsh-Pasek & Golinkoff, 1996; Seidl, Hollich, & Jusczyk, 2003). Ultimately, preschoolers build a lexicon and grammar that support broad generalization to new sentences. Accounts of these feats of learning necessarily begin with the child’s understanding of the world: The novice, not yet knowing the words or the syntax, must try to link input utterances with aspects of accompanying scenes. Top-down knowledge derived from the scene then “supervises” word and syntax learning, investing words and their combinations with meaning. Theories of language acquisition of all theoretical stamps thus assume that independent knowledge of word and sentence meaning drives both word and syntax acquisition (e.g., Chang, Dell, & Bock, 2006; Pinker, 1984; Tomasello, 2003).

However, aspects of verb meanings in particular challenge the assumption that children can recover word and sentence meanings based only on understanding scenes, and thus in turn challenge our theories of language acquisition (e.g., Gleitman et al., 2005). The reason for this is simple: Verbs and other predicate terms do not label world events directly. Rather, verbs as instantiated in sentences denote abstract conceptualizations or construals of an event, highlighting the roles of differing subsets of participants in the event or adopting different conceptualizations of the same participants’ roles (Clark, 1990; Fillmore, 1977; Gentner & Boroditsky, 2001; Levin & Rappaport Hovav, 2005; Pinker, 1989; Rispoli, 1989). Top-down feedback from the scene thus provides ambiguous evidence for verb and sentence meaning (e.g., Gillette, Gleitman, Gleitman, & Lederer, 1999; Snedeker & Gleitman, 2004).

For example, suppose an infant and parent are playing with a shape-sorter toy as in Fig. 1. The shared goals of this interaction, embedded in a turn-taking game, establish that it is the child’s turn to fit a shape into its place. If the parent chooses to speak in this constraining context, the child might infer the parent’s likely communicative goal of exhorting the child to fit a block into the toy (e.g., Vouloumanos, Onishi, & Pogue, 2012). Nonetheless, the parent must select among multiple options to achieve this goal. She might say “The block goes here,” commenting on the block’s motion, or “You can put the block here,” commenting on the action needed to bring about that motion. This illustrates the problem: The relational meanings of sentences are not determined by events, but by the speaker’s choice of construals of events. In this case, if the child carries out the desired action of fitting a block into the toy, the resulting world event will include both the child’s causal role and the block’s motion; but the presence of a causal agent in the world event does not demand that a speaker choose a sentence that encodes the agent’s role.

The syntactic-bootstrapping theory proposes that verb learning succeeds despite this problem because children use knowledge of syntax, as well as of the observed scene, to interpret sentences and thus to figure out the meanings of verbs (e.g., Gleitman, 1990;
Syntactic bootstrapping depends on the tight links between syntax and meaning that pervade natural languages (e.g., Carlson & Tanenhaus, 1988; Goldberg, 2006; Grimshaw, 1981; Jackendoff, 1990; Levin & Rappaport Hovav, 2005; Pinker, 1989; Williams, 2015). Roughly speaking, the syntactic structure of a sentence provides information about the number and type of syntactic arguments surrounding the verb; a semantic analysis of the syntactic structure yields a semantic event structure that specifies (among other things) how many and which participant roles are encoded as arguments in the sentence. This semantic analysis requires that learners have access to at least some links between syntax and meaning, resulting from innate assumptions, language-specific learning, or both (e.g., Gleitman et al., 2005).

To a listener who already knows the verb, its meaning as instantiated in a particular sentence will reflect both the semantic event structure (largely contributed by the syntactic structure) and the verb’s root meaning, which includes the idiosyncratic semantic content that differentiates crying from laughing or baking from boiling. The key insight of syntactic bootstrapping is that some of this interpretive work can be done by a child who does not yet know the verb (e.g., Landau & Gleitman, 1985; Trueswell & Gleitman, 2007). On hearing “You can put the block here,” the child might use the syntactic structure of the sentence to derive its semantic event structure; this semantic structure could then guide the search for an appropriate event construal in the current scene—one that involves the right set of participants. This syntactically guided search constrains hypotheses about the verb’s semantic

---

Fig. 1. Schematic diagram of the origins of syntactic bootstrapping via structure-mapping. Note: The structured conceptual representations shown here represent the assumption that infants’ conceptual representations have predicate-argument structure, and that related construals of the same event can differ in their number of arguments (as in put vs. go). No commitment to a particular set of conceptual primitives is intended.
content. In this way, syntactic bootstrapping proposes that the ordinary processes of syntax-guided sentence interpretation support verb learning.

Accordingly, preschoolers assign different interpretations to new verbs presented in different syntactic contexts (e.g., Arunachalam & Waxman, 2010; Lidz, Gleitman, & Gleitman, 2003; Naigles & Kako, 1993; Papafragou, Cassidy, & Gleitman, 2007; Yuan & Fisher, 2009). For example, 2-year-olds who heard an invented verb in transitive sentences (“The duck is kradding the bunny!”) looked longer at an event in which a duck caused a bunny to bend than at an event in which the duck and bunny each waved an arm, relative to children who heard the new verb in intransitive sentences (“The duck and the bunny are kradding!”; Naigles, 1990). Via syntactic bootstrapping, the new verb’s sentence context provides clues about the intended semantic event structure (e.g., Grimshaw, 2005); the semantic content of the verb (bending vs. arm-waving) must be derived from the scene.

In addition, as we shall discuss below, most verbs occur in multiple sentence structures and take on correspondingly varying semantic event structures. A core claim of the syntactic bootstrapping theory is that children can gain increasingly refined guidance for determining verb (root) meaning by learning about the set of syntactic structures each verb accepts (e.g., Gleitman et al., 2005; Landau & Gleitman, 1985).

2. The origins of syntactic bootstrapping: The structure-mapping account

This article explores the developmental origins of syntactic bootstrapping, asking by what mechanisms children first begin to find syntax meaningful. We and our colleagues have proposed a structure-mapping account of early syntactic bootstrapping on which little or no language-specific syntactic knowledge is needed to give the child an initial sentence-structural guide. This account depends on three main claims (Fisher, 1996, 2000; Fisher, Gertner, Scott, & Yuan, 2010):

1. **Structure-mapping**: Syntactic bootstrapping begins with an unlearned bias toward one-to-one mapping from each noun phrase in a sentence to a participant role in the event construal encoded by the sentence (e.g., Fisher, 1996; Gleitman, 1990; Lidz et al., 2003). Given this bias, the set of noun phrases in the sentence becomes intrinsically meaningful to children.

   This definition appeals to an event construal rather than the world event itself; recall that one could comment on a one-participant construal of a caused-motion event, as noted in discussion of Fig. 1. For this reason, structure-mapping proposes a directional one-to-one mapping, from the set of noun phrases in the sentence to the set of participant-roles highlighted by the sentence; the reverse mapping, from the set of participants in the world scene to a set of noun phrases in a predicted sentence, would require the learner to know (by non-syntactic means) what event construal the speaker intended.

2. **Independent encoding of verb syntax**: Toddlers gather distributional facts about verbs’ syntactic behavior through listening experience, independent of access to a
useful referential context. Children thus can begin to learn which sentence structures each verb appears in before knowing what each verb means.

3. **Early abstraction**: Learners begin with a bias to represent sentences and their meanings in abstract terms that promote useful generalizations, such as noun, verb, agent, and patient, rather than only concrete, word- and situation-specific representations (e.g., Gleitman et al., 2005; Pinker, 1984; cf. Braine, 1963; Tomasello, 2003).

To illustrate, suppose our learner encounters the verb *put* in “You can put the block here,” in the context of Fig. 1. The scene offers multiple candidate meanings, including the *put* and *go* construals mentioned earlier. Via structure-mapping, a partial representation of sentence structure might guide choices among these meanings. If children represent the input sentence as including two noun phrases, they can map it onto a meaning that involves two participant roles, such as the intended *put* construal. Such experiences would permit the child to learn about *put*, including its appearance in a particular syntactic structure (e.g., with two noun phrases), the corresponding semantic event structure that it can take on (with two participant roles), and an initial estimate of its root meaning (derived from the scene). Given independent encoding of verbs’ syntactic behavior, parts of this verb knowledge could be established even when no event-derived semantic content is available. An input sentence as in Fig. 1, without the informative scene, yields data about *put*—a syntactic structure that it accepts, and therefore a semantic event structure compatible with this verb (though not its semantic content). Finally, given abstract representations of sentence form and meaning, the sentence–scene pair in Fig. 1 provides data not only about sentences with *put*, but also about English transitive sentences more generally. If children represent the noun phrases in order, for example, this experience might provide one data point suggesting that the first of two noun phrases in a sentence specifies an agent’s role. Couched in abstract terms, this knowledge should transfer to similar sentence structures containing other verbs.

What if the parent had instead said, “This block goes here”? With appropriate changes to the sentence representation in Fig. 1, the same procedure would unfold. If children represent this sentence as including one noun phrase, then they could map it onto an event-construal involving one participant role, such as the *go* construal. The consequences for verb and syntax learning then follow as described above for *put*, again with appropriate changes: The *go* sentence yields data that could contribute both to learning about *go*, and to knowledge about English intransitive sentences more generally.

As this description shows, structure-mapping depends on strong assumptions about the nature of early conceptual representations of events. First, to permit a systematic alignment of sentences and scenes via structure-mapping, infants’ non-linguistic conceptual representations must themselves have predicate-argument structure, honoring the fundamental distinction between the participants in an event and the roles they play. Second, the value of this alignment for verb and syntax learning depends on some degree of similarity between infants’ conceptual representations of participants’ roles and the thematic roles that are expressed in sentences. Similar assumptions are shared with linguistic theories of argument structure (e.g., Grimshaw, 1990; Jackendoff, 1990; Levin & Rappaport
Hovav, 2005; Williams, 2015) and are made explicit in Fig. 1, which depicts event con-
struals with a formal distinction between predicates and arguments, and with short-hand
labels for abstract roles such as agent and patient. These structured conceptual representa-
tions, in turn, permit the child to detect the correspondence between noun phrases in sen-
tences and participant roles in (construals of) events.

Evidence for these assumptions comes from at least two sources (see also Fisher &
Gleitman, 2002; Fisher & Song, 2006; Gordon, 2003; Lakusta et al., 2007; Wellwood
et al., 2015). First, deaf children who are not exposed to conventional languages invent
gestural communication systems (known as home sign) that bear striking similarities to
conventional languages (e.g., Goldin-Meadow, 2003). For example, home signers invent
distinct signs for actions and the participants in them, and often use gesture order to mark
abstract agent and patient roles. These patterns suggest that children naturally analyze
their experiences into conceptual predicates and arguments, and detect the abstract simi-
larities between the agents and patients of diverse events (e.g., breaking, giving, eating).
Second, infant cognition research suggests that infants create structured mental represen-
tations of events; in these representations, objects are individuated and linked with dis-
tinct roles such as the hitter and hittee in a collision event, the container and content in a
containment event, and so on (e.g., Stavans, Lin, Wu, & Baillargeon, 2019; Yin & Csi-
bra, 2015). These patterns suggest that young infants’ representations of events have
predicate–argument structure: They include knowledge-rich role categories that depend
on the relationships between objects rather than on the identity of the objects themselves.

Via structure-mapping, early syntactic bootstrapping is grounded in noun learning and
in simple innate expectations about predicate-argument structure. We assume that chil-
dren learn the meanings of some nouns without syntactic help (e.g., Gillette et al., 1999);
they then infer, by virtue of their referential meanings, that these words are nouns, and
thus candidate arguments of verbs. This step permits further learning via syntactic boot-
strapping, including estimating the syntactic privileges and semantic event structures com-
patible with each verb by the “counting the nouns” procedure sketched here.

Finally, we note that this account is at odds with current usage-based accounts of lan-
guage acquisition, in that it proposes early abstract representations and innate expecta-
tions about linguistic predicate-argument structure. On a usage-based account, children
represent their early experiences with language in concrete terms, tied to particular
remembered word-strings and situations (e.g., Abbot-Smith & Tomasello, 2006;
Ambridge & Lieven, 2015; Tomasello, 2003). They then gradually construct knowledge
of increasingly abstract linguistic patterns by detecting syntactic-semantic analogies
across many input sentences. This process is assumed to recruit only domain-general
mechanisms, with no built-in constraints on relationships between linguistic form and
meaning. On this account, early sentence comprehension and production are guided by
lexicalized schemata that permit limited generalization to new utterances (e.g., I’m X-ing
it, where X stands for a range of options tied to the words the child has experienced in
this slot; Ambridge & Lieven, 2015). The constructivist account contrasts with early-ab-
straction accounts of language acquisition (including our structure-mapping account),
which assume that children learn syntactic and semantic facts about each word, but are
also biased to represent the form and meaning of sentences in abstract terms (e.g., Fisher et al., 2010; Gertner, Fisher, & Eisengart, 2006; Pinker, 1989; Valian, 1986).

In what follows we first briefly review some evidence for the three core claims of the structure-mapping account. We then focus on challenges to our account that result from the ambiguity of input sentences. These challenges will prompt us to refine our view of the independent encoding of verb syntax, and to enrich the structure-mapping account with a fourth claim, that children approach language learning with an expectation of discourse continuity.

3. Testing the predictions of the structure-mapping account

3.1. Structure-mapping: Learning predicate meanings by “counting the nouns”

We have proposed that syntactic bootstrapping begins with a built-in bias toward one-to-one mapping from noun phrases in sentences to participant roles in the child’s construals of events. This proposal makes strong predictions. Perhaps most obviously, the number of noun phrases in a sentence should guide very early verb learning. Via structure mapping, the set of noun phrases becomes inherently meaningful, without further learning about the native-language syntax. If so, children should assign suitably different interpretations to new verbs appearing in simple transitive or intransitive sentences as soon as they can (a) identify some noun phrases and (b) represent them as parts of a larger sentence structure.

Infants satisfy these prerequisites early in the second year of life. By 14 months, infants treat new nouns (e.g., “This one is a blicket”), but not just any new word (“This one is blickish”), as having object-referential meanings (Booth & Waxman, 2009; Waxman & Booth, 2001). Moreover, 14- and 15-month-olds understand multi-word sentences under some circumstances (Gagliardi, Mease, & Lidz, 2016; Hirsh-Pasek & Golinkoff, 1996; Seidl et al., 2003). For instance, Seidl et al. (2003) showed infants an event in which a book hit a bunch of keys. Following this event, infants saw the book and keys side by side; 15-month-olds (but not 13-month-olds) looked longer at the book if they heard “What hit the keys?”, but at the keys if they heard “Where are the keys?”. This success required infants to identify multiple familiar words per sentence and integrate their meanings to interpret the sentence. Our account predicts that children should use syntax to guide verb learning as soon as these prerequisites are in place.

Yuan, Fisher, and Snedeker (2012) began to test this prediction. In experiments using a looking-preference comprehension task, they showed 21- and 19-month-olds two video events, presented side by side. One was a two-participant caused-motion event (e.g., one actor turned another in a swivel chair), and the other was a one-participant event (e.g., one actor bounced on a large ball). This event pair was accompanied by an invented verb in transitive sentences containing two noun–phrase arguments (e.g., “She’s gorping her”), or in intransitive sentences containing one noun-phrase argument (“She’s gorping”); a third group of children heard neutral audio (“Watch this!”). Children who heard transitive
sentences looked longer at the two-participant event than did those who heard intransitive sentences or neutral audio. This effect of syntax held for both 21- and 19-month-olds and held even when the one-participant event included a “bystander” as shown in Fig. 2, who stood idly as if awaiting her turn. In the bystander case, both test events showed two people, but they differed in whether both filled roles within a single coherent event. This result suggests that the children did not simply seek a video-clip showing the correct number of people for the sentence they heard (one for each noun phrase); instead, they mapped their representations of sentence structure onto structured conceptual representations of events, thereby assigning a relational meaning to a novel verb presented in transitive but not in intransitive sentences.

These and other data now provide strong evidence that syntactic bootstrapping guides verb learning before age 2 (e.g., Arunachalam, Escovar, Hansen, & Waxman, 2013; Messenger, Yuan, & Fisher, 2015). However, one could argue that these experiments have not yet tested the prediction stated above, that children can use the set of noun phrases from the start of multi-word sentence comprehension. To provide a stronger test of this prediction, Jin and Fisher (2014) adapted Yuan et al.’s (2012) task for younger infants, 15-month-olds. Infants saw simple animated events including a two-participant causal event (one animated box bumping another along the floor), and a one-participant event (an animated ball jumping). Infants who heard transitive sentences (“He’s kradding him”) looked longer at the two-participant event (as opposed to the one-participant event) than did those who heard intransitive sentences (“He’s kradding”) or neutral audio. This effect of syntax again held when the one-participant event included a bystander, suggesting that these younger infants also sought to map a two-noun sentence onto a coherent conceptual relation involving two participant roles.

These data provide striking evidence of the early usefulness of syntactic bootstrapping. Evidence that 15-month-olds assign different interpretations to novel transitive and intransitive verbs supports a key prediction of the structure-mapping account: Simple aspects of sentence structure guide comprehension starting early in the second year of life, at or
near the onset of multi-word sentence comprehension. Via the procedure sketched in Fig. 1, these early successes depend on the availability of structured conceptual representations. We argued above that infants’ event representations themselves possess predicate-argument structure, honoring the fundamental distinction between the entities involved in an event and the roles they play; these representations permit the alignment of sentences and event construals via structure-mapping.

Do learners need a built-in one-to-one mapping bias, or could they learn this expectation from language experience? The structure-mapping account proposes an unlearned bias, but the constructivist account sketched above would counter that any effects of syntax on interpretation must be learned from experience (e.g., Ambridge & Lieven, 2015). We would argue that evidence for syntactic bootstrapping at 15 months of age is difficult to reconcile with such a constructivist account, simply because infants near the start of multi-word sentence comprehension have had so little time to develop language-specific argument-structure constructions from experience.

### 3.2. Distributional learning about verbs

The structure-mapping inference sketched in Fig. 1 assumes an input sentence about the “here and now,” offering simultaneous access to syntactic cues to the utterance’s semantic structure and referential cues to the verb’s semantic content. But this lucky coincidence is unnecessary: Toddlers gather syntactic-distributional facts about otherwise unknown verbs from listening experience, independent of access to a useful referential context (Arunachalam, 2013; Arunachalam et al., 2013; Arunachalam & Waxman, 2010; Messenger et al., 2015; Scott, Chu, & Schulz, 2017; Scott & Fisher, 2009; Suzuki & Kobayashi, 2017; Yuan & Fisher, 2009). As a result, children can begin gathering evidence about each verb’s syntactic behavior before they know what it means.

For example, Yuan and Fisher (2009) showed 2-year-old children dialogues in which two women conversed, using an invented verb repeatedly in transitive (e.g., “Bill was blicking the duck!”) or intransitive sentences (“Bill was blicking!”). The dialogue videos showed only the two talkers, providing no referential information about what it means to blick. In later test trials, the children heard the same novel verb in isolation (“Find blicking!”) while viewing side-by-side videos showing a two-participant causal action, and a one-participant action event. Children’s interpretations of the novel verb were guided by the preceding dialogues: Those who had heard the verb in transitive dialogues looked reliably longer at the two-participant event than did those who heard it in intransitive dialogues. This finding held when children were tested one or two days after hearing the dialogues, but disappeared if children heard neutral audio or a different novel verb at test (e.g., kradding instead of blicking; Messenger et al., 2015; Scott & Fisher, 2009). This result has since been extended to children under two years old, to dialogues including more complex sentences, and to input sentences presented without a coherent discourse context (Arunachalam, 2013; Arunachalam et al., 2013; Arunachalam & Waxman, 2010; Messenger et al., 2015). Thus, before their second birthday, children create lasting
memories of the combinatory behavior of new verbs through listening experience, and do so under conditions of great referential ambiguity.

However, learning verbs’ syntactic-distributional privileges is not a simple matter of linking each verb with a single likely syntactic structure and corresponding semantic structure. As noted earlier, most verbs occur in multiple syntactic structures; these alternations are not only syntactic, but involve variations in the linking of semantic roles to syntax. To illustrate, the sentence pairs shown in (1–2) include both transitive and intransitive sentences, but differ in their semantic role assignments across the two structures. The transitive sentences in (1a) and (2a) both describe actions on objects, assigning the agent to subject position and the object acted upon (the patient or undergoer of action) to object position. However, the intransitive *break* sentence (1b) links the patient role to subject position, whereas the intransitive *sweep* sentence (2b) assigns an actor’s role to subject position. These patterns, known as argument-structure alternations, are of interest here because the verbs that participate in them tend to share similar meanings (e.g., Fillmore, 1977; Levin, 2015; Pinker, 1989). The causal alternation shown in (1) accommodates verbs denoting a change of state (*break, spill, open, close*), whereas the unspecified-object alternation shown in (2) accommodates a diverse set of verbs that denote activities without implying any particular result (*sweep, draw, drink, wash*).

(1) a. Ella broke the cup. b. The cup broke.
(2) a. Ella swept the floor. b. Ella swept.

To identify such argument-structure alternations, learners must not only count the noun–phrase arguments, but also categorize them by their semantic roles—in this case, noting that intransitive *break* has a patient subject, and intransitive *sweep* has an actor subject (Ameka, 2008; Gropen, Epstein & Shumacher, 1997).

How could children categorize the arguments accompanying a verb, as well as counting their number? Referential scenes provide one clear source of role information in both children’s and adult’s sentence comprehension (e.g., Chapman & Miller, 1975; Clark, 1973; Huttenlocher, 1974; Knoeferle, Crocker, Scheepers, & Pickering, 2005). For example, in the shape-sorter context of Fig. 1, the child might construe herself as a causal agent, the blocks as patients of caused motion, and the shape-sorter as a location for putting things. These roles are determined by the situation and the child’s understanding of it.

Distributional learning can provide an additional source of role-relevant data. Verbs impose semantic constraints on their arguments, *selection restrictions*. Different nouns will tend to name the patient subjects of intransitive *break* (breakable things), and the actor subjects of intransitive *sweep* (people). Therefore, linking nouns and their semantic categories to partial syntactic representations could help children work out which of the alternations in (1–2) each verb occurs in. Computational analyses have shown that lexical selection, together with syntax, can help to identify verbs that participate in distinct argument-structure alternations in linguistic corpora (e.g., Joanis, Stevenson, & James, 2008; Merlo & Stevenson, 2001; Twomey et al., 2014).
Scott and Fisher (2009) showed that toddlers could learn syntactic-semantic combinatorial facts about verbs simply by listening to sentences. They presented 28-month-olds with novel-verb dialogues with distributional properties characteristic of either the causal alternation (1) or the unspecified-object alternation (2). The dialogues manipulated lexical cues found to identify verbs that occur in each alternation in corpus analyses, including the animacy of nouns in intransitive subject position. At test, children saw a caused-motion event (a girl bent a boy forward and back) and a contact-activity event (the girl brushed the boy’s back with a feather duster) side by side, and heard the novel verb in transitive sentences (“The girl is dacking the boy!”). Children who had heard causal-alternation dialogues looked longer at the caused-motion event than did those who had heard unspecified-object dialogues. Children thus learned more than transitivity via listening; they linked particular nouns, semantic categories such as animacy, or both, to sentence positions surrounding a new verb. This learning about the verb’s syntactic-semantic combinatorial properties affected later verb interpretation.

Further results showed that 2-year-olds linked familiar semantic categories to novel verbs’ argument slots through listening experience (Yuan, Fisher, Kandhadai & Fernald, 2011). In an initial listening phase, 2-year-olds heard two novel verbs presented in transitive sentences. One verb consistently preceded direct objects naming animals (e.g., “I stiped the pig”); the other preceded direct objects naming household objects (e.g., “She nerked the bottle”). Each verb appeared with six nouns from the appropriate category; no pictures of animals or objects appeared during training. At test, children saw a pictured animal and object side by side, and heard test sentences such as “Which one would you stipe?”. Test pictures showed animals and artifacts not mentioned in the listening phase; therefore, children had to detect the category linked with each verb in the training sentences, and generalize to new items to find the target at test. Children looked reliably longer at the target picture, suggesting that they linked unknown verbs with categories of animals or objects based on the distributions of co-occurring nouns.

These findings show that toddlers have some of the tools they need to conquer the complexity of verbs’ syntactic behavior. When children attend to sentences, they gather data about the syntactic-semantic combinatorial privileges of unknown verbs. Much further research will be needed to determine what children can encode about the contexts of new verbs, and how they use that information to guide later sentence processing. For example, can children link categories beyond animacy to the argument slots of new verbs? If so, children might estimate diverse verbs’ selection restrictions by tracking noun-verb co-occurrences, noting that we *splash* liquids, *drink* potable liquids, and so on (e.g., Resnik, 1996). Children’s ability to do so will depend on their word and world knowledge, and on what features of the referents come to mind when children hear nouns in sentences. Category-level semantic priming effects emerge in a variety of tasks between 18 and 24 months (e.g., Friedrich & Friederici, 2005; Mani, Johnson, McQueen, & Huettig, 2013; Willits, Wojcik, Seidenberg, & Saffran, 2013), and 2-year-olds use known verbs to learn new noun meanings (Ferguson, Graf, & Waxman, 2014; Goodman, McDonough, & Brown, 1998). Such findings suggest that lexical processing, even in toddlers, reflects intricate knowledge about meaningful relationships between words.
3.3. Early abstraction

The structure-mapping account requires abstract representations of linguistic form and meaning. These abstract representations play two roles in our account: They give children access to the proposed innate bias to align noun phrases with participant roles in events, and they permit the rapid detection and extension of language-specific patterns. For example, we assume children must learn the significance of word order in English, or the use of case-markers to identify subject and object arguments in Korean or Japanese; but as soon as they begin to learn these facts, they should extend that knowledge to new verbs.

These claims are controversial. As noted earlier, a core hypothesis of constructivist accounts is that all linguistic generalizations emerge from the details of language experience, without the guidance of built-in constraints on relationships between form and meaning (e.g., Ambridge & Lieven, 2015; Tomasello, 2003). To illustrate, on a constructivist account, children might at first represent the form and meaning of an utterance such as “She’s feeding the baby” in concrete terms tied to this word combination, and to the details of a baby-feeding scene. Abstract notions that would permit experience with this utterance to guide the processing of other sentences emerge by comparison across many utterances; these might ultimately include notions such as subject and object, noun and verb, or agent and patient, but such abstractions would not be available to guide sentence interpretation and learning in infancy.

Support for early abstraction accounts comes from the early successes in syntactic bootstrapping described above, and from evidence that toddlers under 2 years old use word order to interpret new verbs. For example, Gertner et al. (2006) showed 21-month-old children a pair of caused-motion events, one in which a girl acted on a boy (causing him to bend), and another in which the boy acted on the girl (turning her in a swivel chair). These events were accompanied by a novel verb in transitive sentences, either “The boy is gorping the girl” or “The girl is gorping the boy.” Children looked longer at the event in which the subject of the sentence they heard played an agent’s role (see also Dittmar, Abbot-Smith, Lieven, & Tomasello, 2008; Fernandes, Marcus, Di Nubila, & Vouloumanos, 2006; Franck, Millotte, Posada, & Rizzi, 2013; Gavarró, Leela, Rizzi, & Franck, 2015; Gertner & Fisher, 2012). These findings imply considerable abstraction in early representations of sentence form and meaning. Presented with an unknown verb in a transitive sentence, toddlers applied their knowledge of the native language to interpret it as referring to the actions of the subject referent on the object referent.

Building on these results, Scott, Gertner, and Fisher (2018) asked whether toddlers could extend their knowledge of English word order to events other than the caused-motion events tested in the previous experiments. Experimental work on early comprehension of word order began with causal events because these represent the canonical meanings of transitive sentences (e.g., Hopper & Thompson, 1980). However, transitive sentences can describe a broad range of two-participant events, including perception events (see, hear) and actions with no specified effect (sweep, touch, tickle, hug); diverse linguistic analyses have proposed that the subjects and objects of transitive sentences map...
onto a broad range of asymmetries in semantic role prominence (e.g., Dowty, 1991; Grimshaw, 1990; Levin & Rappaport Hovav, 2005). For example, Dowty (1991) proposed a prototype concept of the thematic roles agent and patient. On this view, the subject argument of a transitive verb has more of the semantic entailments of a proto-agent, such as volitional involvement, causation, and motion; the object argument has more proto–patient entailments, such as undergoing a change of state. Dowty’s proposal, among others, raises the possibility that children might have access to a flexible range of default interpretations for transitive verbs. That is, even when the referential scene makes a causal interpretation unlikely, children might link transitive word order with an asymmetry in movement or volitional involvement.

To test this possibility, Scott et al. (2018) showed 23-month-old children a pair of animated events involving two characters, a horse and a dog. In one event, the horse put a hat on the dog; in the other, the dog passed a comb over the horse’s mane. This event pair was accompanied by a transitive sentence containing an invented verb: “The horse is meeking the dog,” or “The dog is meeking the horse.” These stimulus events were designed to suggest two-participant meanings lacking a key feature of prototypical transitive sentences: They depicted no clear change of position or state in the recipient of either action. Even so, word order guided children’s looking preferences. Children looked longer at the event in which the actor was named by the subject of the sentence they heard. A second experiment showed that 23-month-olds extended their knowledge of English word order to new verbs in transitive sentences accompanying figure-ground motion events. In one event, schematized in Fig. 3, a flower repeatedly circled a ball; in the other, the ball jumped over the flower. These events suggest meanings with even fewer of the features of prototypical transitive sentences: They involved no contact between the subject and object referent. Children nonetheless interpreted word order systematically, linking the subject of the sentence with the more mobile event participant. The finding that toddlers systematically extended their knowledge of transitive word order

![Sample test-event stimuli from Scott, Gertner, and Fisher (2018).](image)

---

**Fig. 3.** Sample test-event stimuli from Scott, Gertner, and Fisher (2018). In the event shown on the left, the flower moved around the ball in a circular path, while the ball tilted gently in place; in the event shown on the right, the ball jumped back and forth over the flower, while the flower tilted gently in place. The arrows indicate path of motion.
to such events suggests that children may map word order onto a broad range of asymmetries in semantic roles. This pair of studies is part of a broader effort to determine what kinds of events young children readily construe as coherent two-participant events, and what kinds of asymmetries in those participants’ roles children readily map onto word order in transitive sentences (e.g., Arunachalam & Dennis, 2019; Chestnut & Markman, 2016; Fisher & Song, 2006; Gleitman, Gleitman, Miller, & Ostrin, 1996; Kline, Snedeker, & Schultz, 2017; Landau & Gleitman, 2015; Naigles & Kako, 1993).

4. Counting the missing nouns: The ambiguity of sentences

The work summarized above supports several key predictions of the structure-mapping account. These data suggest that simply finding some noun phrases gives children a rough but useful “first pull on the bootstraps,” yielding partial representations of sentence structure that guide interpretation and provide an abstract format for new learning. We have also seen hints of how children might move beyond these first steps, by encoding both syntactic and semantic combinatorial facts about verbs from listening experience.

However, difficulties for our account arise from the pervasive syntactic ambiguity of sentences. To a novice learner, each sentence must be a noisy indicator of its true syntactic structure. For example, most strikingly, many languages allow arguments to be omitted (e.g., Allen, 2008; Bowerman & Brown, 2008). In languages such as Korean or Mandarin, noun-phrase arguments can be omitted from sentences whenever their referents are recoverable in the discourse context. We can think of this as roughly the same contexts in which English speakers would use unstressed pronouns (Huang, 1984). If asked “Where’s my new camera?”, an English speaker might reply “I broke it” (with overt pronouns marking the arguments accompanying the verb in this sentence); in contrast, in such a context a Korean speaker could produce the translation-equivalent of “I broke” or “broke.” Argument-dropping is pervasive in casual speech, including in speech to children. For example, in analyses of child-directed Japanese, more than 85% of transitive sentences were missing one or both arguments (Matsuo et al., 2012; Rispoli, 1989); similarly high rates of omission characterize child-directed speech in Korean, Hindi, and Mandarin (e.g., Clancy, 2009; Lee & Naigles, 2005; Narasimhan, Budwig, & Murty, 2005).

This raises a challenge for syntactic bootstrapping, perhaps particularly for the structure-mapping account. In the shape-sorter game of Fig. 1, a Korean-speaking (or Mandarin- or Hindi-speaking) parent would typically omit some or all of the intended arguments. How could the learner “count the nouns” if they are mostly missing?

In thinking about how children might solve this problem, we consider two likely sources of data: probabilistic combinatorial learning about verbs and discourse structure. The first of these two sources requires us to enrich our notion of distributional learning about verbs.
4.1. Probabilistic combinatorial learning about verbs

Individual sentences are ambiguous, but they jointly yield probabilistic data about verbs’ syntactic distributions (e.g., Hoff & Naigles, 2002; Lee & Naigles, 2005; Twomey et al., 2014; Ural et al., 2009). Given infants’ and toddlers’ prowess in statistical learning (e.g., Lany & Saffran, 2010; Newport, 2016), we might expect toddlers to profit from probabilistic data about whether each verb is used transitively or intransitively. For example, given enough sentences in Turkish, a language permitting pervasive argument-dropping, toddlers might learn that some verbs occur with more noun phrases per sentence than others, on average (Ural et al., 2009).

Evidence reviewed above suggests that young children have some of the tools they need to accomplish this learning. However, most experimental work on linguistic-distributational learning, including the dialogue-based syntactic bootstrapping studies reviewed earlier, tests children in brief experiments with all-or-none, or nearly all-or-none, probabilities. For example, Yuan and Fisher (2009) presented English-learning toddlers with dialogues in which a new verb appeared only in transitive sentences with two noun phrases, or only in intransitive sentences with one noun phrase.

In contrast, to permit learning about verbs’ syntactic distributions from noisy data (and thus to permit syntactic bootstrapping), toddlers must encode syntactic-semantic combinatorial facts about verbs, retaining multiple options for each verb, as noted above. This learning must be probabilistic, not deterministic, and children must use this knowledge to parse sentences despite ambiguity. These requirements for verb learning are strikingly like findings from a related literature on verb bias: When adults or 5-year-olds identify a verb, they retrieve probabilistic syntactic-semantic combinatorial knowledge about that verb, and use it to resolve syntactic ambiguity (e.g., Garnsey, Pearlmutter, Myers, & Lotockey, 1997; Kidd & Bavin, 2005; Snedeker & Trueswell, 2004). For example, in the sentence “Tickle the pig with the fan,” the underlined prepositional phrase could be attached to the preceding verb, specifying an instrument for tickling, or to the noun phrase, specifying which pig to tickle (Snedeker & Trueswell, 2004). Resolution of this ambiguity depends on whether the verb in the sentence occurs with instrument prepositional phrases often (e.g., tickle) or rarely (e.g., choose). Effects of verb bias emerge quickly, guiding online parsing.

Based on the striking similarities between the role of verb bias in online parsing, and the requirements for structure-guided verb learning, we and others argue that verb bias and syntactic bootstrapping reflect the same underlying phenomenon (e.g., Trueswell & Gleitman, 2007; Wonnacott, Newport & Tanenhaus, 2008). Probabilistic distributional learning creates syntactic-semantic combinatorial knowledge about verbs that plays two roles. First, it permits syntactic bootstrapping. Children gather syntactic-semantic combinatorial facts about each verb (permitting them both to count and categorize the arguments assigned to that verb in sentences) and use that knowledge to guide learning of verb meaning. Second, from early in development, verb combinatorial knowledge is used online to identify the intended structure and lexical content of sentences, reducing ambiguity (this is the familiar effect of verb bias).
Investigations of this proposal involve studying probabilistic learning about the syntactic biases of verbs or other word classes (e.g., Culbertson & Newport, 2015; Hudson Kam & Newport, 2005; Lin & Fisher, 2017; Perek & Goldberg, 2015; Qi, Yuan, & Fisher, 2011; Thothathiri & Rattinger, 2016; Twomey, Chang, & Ambridge, 2016; Wonnacott, 2011; Wonnacott et al., 2008). One key result from this literature is that learners simultaneously consider syntactic co-occurrence probabilities at multiple levels of abstraction and use them to choose “rationally” whether to link new learning to a specific word, or to extend it to a broad class of words that appear in the same contexts (see also Gerken, 2010; Reeder, Newport & Aslin, 2013).

Based on this emerging literature, we strongly assume that children have some ability to handle noisy data in learning about verbs’ syntactic distributions. But we also assume that learners face serious data sparseness problems in estimating the syntactic-semantic combinatorial behavior of each verb. For example, in languages with frequent argument-dropping, children might rarely or never get a chance to observe the true number of arguments typically assigned to relatively infrequent verbs (Bowerman & Brown, 2008; Narasimhan et al., 2005). This brings us to our second likely source of constraint.

4.2. Discourse structure

How arguments are realized in sentences depends on their place in a larger discourse. Noun-phrase arguments are pronominalized or omitted, not arbitrarily, but when their referents can be recovered from the situational or linguistic context (e.g., Allen, 2008; Clancy, 2003; Du Bois, 1987; Narasimhan et al., 2005; Prince, 1992). For example, arguments are more often omitted when their referents are given rather than new, or if they refer to entities that are present rather than absent. Nearby sentences in connected discourse tend to share arguments; as a result, typical sentences in casual speech introduce at most one new lexical argument (as opposed to pronominal or null arguments; Allen, 2008; Du Bois, 1987). To a learner who analyzes sentences in isolation, these patterns are noise to be overcome by probabilistic combinatorial learning; but to one who analyzes utterances in context, they might be useful.

Given these patterns, we and others propose that a bias to expect discourse continuity in comprehension increases linguistic support for verb learning, by letting learners collect evidence for argument structure across sentences in a discourse (Clancy, 1996; Narasimhan et al., 2005). For example, recent mentions might let toddlers identify sentences as transitive despite missing arguments as in (3).

(3) Question: What’s Mom, doing?  
Answer: [Øi] Talking to Grandma.

Past evidence yields reason to suspect that children could use established topics to recover missing arguments in this way. First, argument realization in toddlers’ own speech shows sensitivity to discourse structure (e.g., Allen, 2008; Clancy, 2003;
Like adults, children omit arguments more often when their referents are given rather than new, or present rather than absent. Second, in comprehension, toddlers and preschoolers link ambiguous phrases such as pronouns with referents that have been recently and prominently mentioned (e.g., Hartshorne, Nappa, & Snedeker, 2015; Horowitz & Frank, 2015; Lidz, Waxman, & Friedman, 2003; Saylor & Ganea, 2007; Song & Fisher, 2005, 2007).

Jin et al. (in preparation) asked whether Korean-learning 2.5-year-olds could learn whether a new verb appears in transitive or intransitive sentences despite missing arguments, with appropriate discourse support. Korean makes a good test case for this purpose: It has a canonical SOV (subject-object-verb) word order, and it permits pervasive argument dropping. Korean has nominal case-markers, affixes on nouns that identify their grammatical roles; however, these case markers are optional, and they are often absent in casual speech (e.g., Kim, 2008; Lee, 2006). As a result, children frequently encounter sentences that might be either transitive or intransitive, consisting of a single noun followed by a verb.

In two experiments, Jin et al. (in prep.) used the dialogue training method introduced by Yuan and Fisher (2009), but created dialogues that provided little or no direct syntactic evidence for transitivity. For example, in the dialogues for one experiment, two women used the invented verb *thomita* in transitive or intransitive sentences embedded in natural discourses. As shown in Fig. 4, all novel-verb transitive sentences had dropped subjects, thus only one overt noun phrase; no accusative case markers were included. Crucially, these sentence subjects were omitted in appropriate discourse contexts, those in which the referent had been prominently mentioned, including in an immediately preceding question. In the intransitive dialogues, all novel-verb sentences also contained one overt noun phrase. Therefore, the number of noun phrases appearing with the new verb was identical in the transitive and intransitive dialogues. Moreover, due to the SOV word order of Korean, the transitive and intransitive answers to the context question (bold text in Fig. 4) had the same word-order—one noun followed by a new verb. One key difference between the transitive and intransitive dialogues was whether this single noun was the same noun that appeared in the question (in intransitive dialogues), or a new noun (in transitive dialogues).

In the test trials (Fig. 4), children heard the novel verb in isolation (an English gloss is “Find thomming!”) while viewing side-by-side videos depicting a two-participant causal action (one actor swings the leg of a seated actor), and a one-participant action enacted by two people (both actors wheel their arms). Children’s interpretations of the new verb were guided by the dialogues: Those who had heard the verb in transitive dialogues looked reliably longer at the two-participant event than did those who had heard it in intransitive dialogues. This effect disappeared if children heard a different novel verb at test, one not heard in the dialogues (as in “Find mwupping!”). Thus, Korean toddlers who heard our transitive dialogues learned that an invented verb could be transitive without ever encountering it in a sentence with two noun-phrase arguments or with an accusative case marker, when the discourse context strongly implied a missing argument. In additional studies in English, Jin (2015) provided converging evidence for the
contribution of discourse structure by comparing coherent dialogues similar to English translations of the dialogues in Fig. 4 to “scrambled” dialogues in which the same critical novel-verb sentences were presented without appropriate discourse support. The effect of transitive versus intransitive dialogues on children’s looking preferences at test was found in the coherent-dialogue condition, but not the scrambled-dialogue condition, confirming that an informative discourse context helps children to recover missing arguments.

By what mechanisms might a preceding discourse context help children to infer a missing argument for an unknown verb? One possibility is that this inference could result primarily from probabilistic combinatorial learning about language that spans utterance boundaries. That is, children might learn to expect referents that are linguistically marked as prominent to be re-mentioned in subsequent utterances (e.g., Arnold, Brown-Schmidt, & Trueswell, 2007; Han, 2006). In this case, children could have learned that repeated use (as opposed to omission) of terms such as “Grandma” suggested a different grandmother was being acted on.

More generally, children might posit a missing argument based on the semantic or pragmatic constraints of their model of the discourse. Researchers have proposed that each sentence is interpreted relative to a propositionally encoded representation of the
discourse; this representation includes a set of discourse referents and what has been said about them (e.g., Bock & Brewer, 1985; Grosz & Sidner, 1986; Johnson-Laird, 1983; Kintsch, 1988). Each new utterance alters the model, thus affecting interpretation of the next. A propositionally encoded situation model provides a common workspace within which to integrate linguistically provided information with world knowledge and situational constraints; as a result, children can reason about meaningful links between utterances using the same inferential mechanisms that permit them to understand events they see (e.g., Kehler & Rohde, 2013; Lynch et al., 2008; Sullivan & Barner, 2016). In the case of our Korean dialogues, the context question preceding each novel verb sentence (e.g., “What’s Grandma doing?”) should highlight the referent’s possible actions in the child’s model of the linguistically described situation; this in turn should bias the child to assume that the same referent (and her actions) might form part of the answer.

Thus, based on linguistic expectations, inference-making within a situation model, or both, a bias toward discourse continuity might allow children to recover the missing arguments of unknown verbs. This process will not be simple, as opportunities for error abound. Question-answer pairs like those in our dialogues create strong discourse constraints, but many kinds of adjacent sentences in conversation will license omitted arguments. For example, if we say “Grandma brought a present,” natural continuations could include Grandma (e.g., “(She) just arrived”), the present (“(It) is very big”), or both (“(She) made (it)”). In English, any of these references to previously mentioned entities would be marked by overt pronouns; but in Korean, likely usage would involve null pronouns. How might the novice learner infer which argument (the person or the present) was missing in a particular continuation, or infer one missing argument for the *arrive* or *is big* sentences, but two missing arguments for the *make* sentence?

A comparison to the interpretation of overt pronouns both illustrates the problem and suggests a possible solution. A pronoun provides an overt cue to seek an antecedent in the discourse model, and it offers linguistic features to guide that search (e.g., gender, number; Arnold et al., 2007; Gelman & Raman, 2003; Saylor, Ganea, & Vázquez, 2011). Omitted arguments offer no such cue, but they still leave the verb itself, including its morphology, and any overt arguments or modifiers, to launch and guide the search. In languages permitting argument omission, syntactic-semantic combinatorial knowledge about familiar verbs prompts adults to seek antecedents for null elements in the discourse model (e.g., Mandarin: Yang, Gordon, Hendrick, & Wu, 1999; Korean: Kwon & Sturt, 2013). We speculate that the same could be true in the learning system. As children gather partial knowledge of the combinatorial properties of verbs (both counting and categorizing the arguments appearing with the verb), and about verb morphology (e.g., Rispoli, 1989), this learning might invite integration of suitable discourse referents into sentences with missing arguments. This speculation represents a generalization of the verb bias proposal laid out earlier. In this way, learning about the syntactic-semantic distributions of verbs might guide not only identification of the structure and content of each sentence, but also its linking with the discourse context.
5. Conclusion and further challenges

We began with a learning problem, arguing that children need linguistic evidence about the number of arguments accompanying a verb to offset the ambiguities of events as evidence for verb meanings. The structure-mapping account of early syntactic bootstrapping proposes a mechanism whereby children could receive such linguistic evidence from the start of multi-word sentence comprehension, before learning much about the syntax of the native language. We argued that children might begin by interpreting each noun phrase as an argument, and thereby get some aid for sentence interpretation from even a partial sentence representation (Fisher, 1996; Gleitman, 1990; Lidz et al., 2003).

Here we focused on challenges posed by languages that permit frequent argument dropping. In languages such as Korean, noun–phrase arguments are freely omitted when their referents are recoverable in context. As a result, learners cannot count on the noun phrases in each sentence to display all the arguments of the verb. This challenge prompted us to enrich the structure-mapping account in two ways.

First, we refined our notion of distributional learning about verbs to include probabilistic learning about verbs’ syntactic-semantic combinatorial privileges. To conquer the complexities of verbs’ syntactic distributions, we argued that children must not only count the noun phrases in single sentences, but also categorize them semantically to identify their likely argument roles and track multiple syntactic-semantic combinatorial options for each verb. This procedure links syntactic bootstrapping in early childhood with the use of verb bias in online processing: The same probabilistic distributional learning procedure creates verb biases that guide online sentence parsing and permits syntactic bootstrapping (Trueswell & Gleitman, 2007; Wonnacott et al., 2008).

Second, we added a fourth mechanism to the structure-mapping account, proposing that an expectation of discourse continuity allows children to gather linguistic evidence for the arguments accompanying a verb across sentences in a coherent discourse (see also Allen, 2008; Narasimhan et al., 2005). Building on evidence that toddlers use discourse context to interpret ambiguous overt noun phrases such as pronouns or unknown nouns (e.g., Horowitz & Frank, 2015; Lidz et al., 2003; Saylor & Ganea, 2007; Song & Fisher, 2005, 2007; Sullivan & Barner, 2016), we reviewed new evidence that toddlers can recover a missing argument for a brand-new verb with appropriate discourse support (Jin et al., in prep.). These findings raise many questions for future research about the developmental source of children’s expectation of discourse continuity, and how children integrate their discourse-guided referential expectations with distributional evidence about the syntactic-semantic combinatorial properties of verbs.

Enriched with these two modifications, the structure-mapping account sketches a path whereby young learners can go beyond “counting the nouns” to estimate each verb’s syntactic distributions. This estimate incorporates multiple syntactic-semantic combinatorial options, and it can use the sets of noun-phrase arguments co-occurring with each verb to help determine its meaning even if some of those arguments are lost to discourse redundancy.
But missing arguments are only one example of the noisy relationship between noun phrases in sentences and the arguments that accompany verbs. Just as individual sentences can have too few noun phrases to display their intended arguments, they can also have too many. To illustrate, "on the couch" is an argument accompanying put in (4a), but not like in (4b), where it is part of a complex noun phrase argument. “With Susie” is an argument accompanying chat in (5a) but not walk in (5b), where it is an adjunct phrase indicating accompaniment. Distinguishing arguments from adjuncts is tricky, but central intuitions are that arguments are semantically obligatory (it takes two to chat, but not to walk), and their roles depend on the verb (Dowty, 2003; Koenig, Mauner, & Bienvenue, 2003; Schütze & Gibson, 1999; Tutunjian & Boland, 2008). Roughly speaking, arguments but not adjuncts yield information about the semantic structures compatible with each verb.

(4) a. She put [the puppy] [on the couch].
   b. She liked [the puppy on the couch].

(5) a. She chatted [with Susie].
   b. She walked [with Susie].

Via structure-mapping, “extra” noun phrases accompanying novel verbs should cause errors. Several investigations confirm this prediction (e.g., Dautriche et al., 2014; Pozzan & Trueswell, 2015). For instance, Gertner and Fisher (2012) found that 21-month-olds tended to interpret intransitive sentences with two noun phrases conjoined in subject position as if they were transitive, mistakenly linking “The boy and the girl are pilking!” with an event in which a boy acted on a girl.

To escape such errors, children must learn the language-specific features that identify complex noun phrases and begin to identify which constituents might be arguments rather than adjuncts. We speculate that the same two data-sources examined here, probabilistic syntactic-semantic combinatorial learning about verbs and an expectation of discourse continuity, could support these learning tasks, in much the same way that we argued they support the recovery of missing arguments.

For example, consider the prepositional phrase “on the couch” in (4). As noted earlier, both adults and preschoolers use the biases of familiar verbs to guide decisions about whether such phrases modify a preceding noun (4b) or are attached to the main verb and might be arguments (4a) (e.g., Snedeker & Trueswell, 2004). In addition, modifiers as in (4b) are expected, and therefore easier to parse, if the discourse context includes multiple referents of the same kind (two puppies), and thus a modifier is needed to identify the intended referent. Adults use such referential–context constraints, whether provided in the physical setting (two puppies in view) or in a linguistic discourse context (two puppies under discussion), to make attachment decisions under ambiguity (e.g., Altmann & Steedman, 1988; Trueswell, Sekerina, Hill, & Logrip, 1999). Preschoolers typically fail to use the number of possible referents in the physical setting to make parsing decisions (e.g., Trueswell et al., 1999), but readily use the linguistic discourse context to do so. For
example, 4- and 5-year-olds were more likely to arrive at a modifier interpretation of an ambiguous prepositional phrase if a preceding question highlighted the need to differentiate between multiple referents, as in (6), relative to a more general question ("What happened?"; Trueswell & Gleitman, 2004).

(6) Question: Which cat did the turtle tickle?
Answer: I know! The turtle tickled the cat on the fence.

Thus, the biases of known verbs and the semantic constraints of the discourse affect preschoolers’ parsing under ambiguity. This raises the intriguing possibility that similar influences might shape learning about argument structure early in acquisition. A learner who gathers distributional evidence about the syntactic-semantic combinatory privileges of unknown verbs, and who interprets each sentence relative to a model of the discourse, might sometimes distinguish (4a) from (4b) even before the meanings of these verbs are known. Distinguishing arguments from adjuncts requires tracking probabilities at multiple linguistic levels, to estimate the combinatory privileges of particular verbs relative to verbs in general, as discussed earlier. Roughly speaking, adjuncts combine freely with many verbs, retaining their meanings as they do so (e.g., “with Susie” in (5b), “on Tuesday”), whereas the occurrence and interpretation of arguments is predicted by particular verbs. Distributional criteria based on this intuition can be used to automatically classify phrases as arguments or adjuncts (e.g., Merlo & Ferrer, 2006). Linking the distinction between arguments and adjuncts to verb-bias learning in this way is consistent with evidence that argument status is gradient rather than categorical (Rissman, Rawlins, & Landau, 2015).

Though the data required to make good on these final speculations remain to be gathered, these considerations suggest additional ways in which young learners go beyond “counting the nouns” to estimate the argument structures compatible with each verb. This estimate should guide verb learning despite the ambiguity of sentences, which might display too few noun phrases or too many.

Acknowledgments

This research was supported by grants from the National Science Foundation (BCS 1348522) and the Eunice Kennedy Shriver National Institute of Child Health and Human Development (HD054448). We thank Renée Baillargeon, Jesse Snedeker, and an anonymous reviewer for helpful comments.

Notes

1. We use the term "argument" for both syntactic arguments and semantic arguments. We use "participant roles" in the sense described by Williams (2015), as a
privileged set of the roles implied by a conceptualization of an event (the agents and patients of events are privileged relative to their locations and times, for example). This characterization raises many questions about which roles are privileged and why. We will say a bit more about this later in the paper, but we generally assume that the most satisfying answers for a theory of acquisition will emerge from the study of infant event concepts and how infants’ preferred conceptualizations of events influence sentence comprehension (e.g., Gordon, 2003; Lakusta, Wagner, O’Hearn & Landau, 2007; Wellwood, He, Lidz, & Williams, 2015). 2

2. We use terms such as "agent" and "patient" as a short-hand for categories of thematic roles (semantic roles that are relevant to the grammar), intending no commitment to any particular view of how these roles are represented (as atomic notions; as clusters of entailments determining the relative prominence of verb-specific roles, Dowty, 1991; or as positions in an event structure anchored by primitive predicates such as CAUSE and BECOME; Jackendoff, 1990; Levin & Rappaport Hovav, 2005). 3

3. Note that this first inference is a form of semantic bootstrapping, using links from semantics to syntax to infer the syntactic category of known words (Pinker, 1984). Semantic and syntactic bootstrapping rely on the same tight links between syntax and semantics but deploy them in different directions; these are probably best seen as complementary rather than as contradictory learning procedures (e.g., Gleitman, 1990). 4

4. Ural et al. (2009) calculated the average number of nouns per sentence, and they found that this measure was of some use in classifying verbs as typically transitive or intransitive in Turkish. This raises many questions about how learners might track the statistics of their input. For example, the average number of nouns may be of less use than an estimate of the maximum number of nouns that appears with a verb, at least when this estimate is shaped by the learner’s developing expectations about the syntactic-semantic combinatorial biases of each verb, relative to the behavior of verbs in general, as discussed in this section. See Gleitman et al. (2005) for discussion. 5

5. The transitive and intransitive dialogues also differed in noun phrase animacy. As exemplified in Fig. 4, the topic established in the context was always human (e.g., Grandma, Daddy, Aunt), whereas the new noun phrase introduced in the answer to this question across the transitive dialogues included humans (boy, girl, baby), other animates (puppy, raccoon, turtle), and inanimates (teddy bear, pillow). This difference was introduced to avoid including many human direct objects without accusative case markers. This is important because Korean shows probabilistic differential case marking (Lee, 2006): Case markers are optional in Korean, but nouns whose properties make them atypical for their grammatical role tend to receive overt case marking. Thus, nouns high in animacy are typically case-marked when used as direct objects, and nouns low in animacy are case-marked when used as subjects.
References


