# TOPICS
## TOPICS IN COGNITIVE SCIENCE

# Learning, Memory, and Syntactic Bootstrapping: A Meditation

## Jeffrey Lidz

*Department of Linguistics, University of Maryland*

**Abstract**

Lila Gleitman's body of work on word learning raises an apparent paradox. Whereas work on syntactic bootstrapping depends on learners retaining information about the set of distributional contexts that a word occurs in, work on identifying a word's referent suggests that learners do not retain information about the set of extralinguistic contexts that a word occurs in. I argue that this asymmetry derives from the architecture of the language faculty. Learners expect words with similar meanings to have similar distributions, and so learning depends on a memory for syntactic environments. The referential context in which a word is used is less constrained and hence contributes less to the memories that drive word learning.

*Keywords:* word learning; syntactic bootstrapping; cross-situational learning; rationalism

## 1. Introduction

In this paper, I wish to make some remarks about a puzzle that emerges by putting together two parts of Lila Gleitman's work on word learning. This puzzle concerns the nature of the child language learner's memory for past experiences. Specifically, she observes an asymmetry in the information that is retained in memory for the purposes of word learning. On one hand, Gleitman and her colleagues have shown that learners do not retain facts about the extralinguistic context surrounding a word, other than their

Correspondence should be sent to Jeffrey Lidz, Department of Linguistics, University of Maryland, 1401 Marie Mount Hall, College Park, MD 20910. E-mail: jlidz@umd.edu

hypothesis about what the word means (Medina, Snedeker, Trueswell, & Gleitman, 2011; Trueswell, Medina, Hafri, & Gleitman, 2013). On the other hand, the theory of syntactic bootstrapping (Gleitman, 1990, Landau & Gleitman 1985) depends on the learner's ability to retain information about the syntactic context in which a word occurs. Syntax is a strong cue to word meaning only given whole patterns of distribution; a single sentence contains very little relevant information by itself. Examining the pattern of syntactic distribution allows learners to restrict their hypotheses about meaning when the verb is used in a specific context. So why is it that referential information from the context of word learning is forgotten, but information from syntactic context is retained?

Clearly, a learner should not remember everything. As Gleitman herself puts it, "An observer who notices everything can learn nothing" (Gleitman, 1990). Consequently, only some aspects of experience enter our memories in a way that allows them to be used by a word learning mechanism. And research on word learning indicates that what is retained is the syntactic environment of an utterance, but not its extralinguistic environment. I will speculate that this asymmetry in what gets retained in memory follows from the architecture of the language faculty. Learners expect words with similar meanings to have similar distributions and so learning depends on a memory for syntactic environments. The referential context in which a word is used is less constrained. Consequently, it is rational to depend on the principled information over the arbitrary. But I will argue that a learner cannot arrive at this rational stance except by fiat. The fact that only some aspects of experience enter into long-term memory and feed forward for subsequent learning reveals the fundamental bias of the learner to favor specific kinds of information over others in a problem-specific way. But to see how, we must first explore the nature of the asymmetry between referential and syntactic information.

## 2. The failure of cross-situational learning

Quine (1960) famously observed that the extralinguistic environment accompanying the use of a word leaves open the particular concepts in the mind of the speaker that condition that use. A situation in which a speaker uses a word to refer to a rabbit might be identical to one in which he intended to refer to the rabbit's fur, the speed at which it is moving, or to the memory of a delicious stew. Indeed, the fact that we are free to talk about anything that crosses our minds, and not just our immediate environment, makes a learning procedure based on pairing words with the contingencies of use of questionable utility (Chomsky, 1959). Finally, even if a speaker is talking about the here and now, and even if a learner could pinpoint precisely the part of the world being picked out by a novel word, recognizing, for example, that the speaker was referring to the rabbit, this reference would be consistent with a broad range of concepts with overlapping extensions (e.g., Pete the rabbit, rabbits or black holes, rabbits with more than one ear, things physically identical to rabbits, etc.; Goodman, 1955).

Linguists and cognitive scientists have responded to these philosophical problems by noting that humans have perceptual, conceptual, and linguistic abilities that can help.

These abilities lead learners to parse the world into the same pieces and to relate those pieces to the same concepts as the speakers producing those words (Carey, 2009; Gleitman, 1990; Markman, 1990; Spelke, 1990; Waxman & Lidz, 2006). Moreover, these conceptual burdens may be significantly reduced by learners' ability to track the goals and intentions of their interlocutors (Baldwin, 1991; Bloom, 2001; Clark & Amaral, 2010).

But even granting learners these constraining capacities, the world still vastly underdetermines the concepts that learners might invoke in explaining the use of a novel word. And, as Landau and Gleitman (1985) observed, learners somehow acquire meanings even for words whose content is absent from their sensory experience, for example when blind children learn meanings for words like *look* and *see*. The insufficiency of the world is perhaps made most clear experimentally in the Human Simulation Paradigm (Gillette, Gleitman, Gleitman, & Lederer, 1999). In this paradigm, adults are shown videos of parent–child interactions with the sound turned off and are then asked to guess what word the parent said at a particular point in the video. The results were not encouraging. For nouns, participants guessed the correct word 44% of the time on average, and for verbs, they guessed correctly only 15% of the time.

One might expect that the weakness of the world as an information source could be overcome by multiple exposures, over which the learner could find a common thread. Pull on that thread and the meaning will be revealed. However, participants in these experiments actually had six attempts to make their guesses, and improvement across trials was very weak. This lack of improvement suggests that integrating information across occurrences does not help much. Indeed, the lack of similarity across contexts led to a significant rise in guesses like "toy" or "look" as the trials progressed. One might expect, then, that a learning procedure based in cross-situational comparison would end with every word having an extremely weak meaning so that all contexts would fall under its extension.

In more recent work, in collaboration with John Trueswell and others, Gleitman has shown not just that cross-situational comparison is a bad idea, but also that learners do not do it. The idea of cross-situational word learning is that learners retain associations between a word and all possible referents for that word, with the expectation that the true association will emerge over many exposures. This kind of theory predicts that learners gradually zero-in on a word's meaning, getting closer and closer across time. Medina et al. (2011) looked for precisely this improvement in an experimental context. Participants were shown natural scenes in which a parent uttered a word to their child and were asked to guess the word's meaning after each scene. They found no improvement across trials, and hence no evidence of cross-situational comparison. Instead, their data were more consistent with a theory in which learners guessed a meaning on the first trial and then sought to confirm that guess.

This theory was further tested in Trueswell et al. (2013). Here, participants were shown several images while they heard a novel word and were asked to guess the meaning of the novel word. A cross-situational learner would predict gradual improvement across trials and would predict no real difference in the information retained between trials in which learners guessed correctly and those in which they guessed incorrectly.

These researchers found that if participants had guessed correctly, then on subsequent trials, they were above chance at choosing the correct object. However, if they guessed incorrectly, they remained at chance, showing no evidence of having retained relevant information about the context, such as the referent that they did not choose. In sum, learners did not show any evidence of retaining information about un-chosen candidate word meanings. In all of these manipulations, the key finding was that learners do not retain lots of information about the context of use. Instead, they make a targeted guess about the word's meaning and forget everything else about the context.

The lesson of this stream of research is that learners do not retain lots of information about the context of use surrounding a word that they are learning. We can interpret this in the following way. Learners fail to encode in long-term memory many aspects of the nonlinguistic context surrounding the use of a word. Consequently, that conglomeration of information cannot contribute (except in the moment) to the word learning procedure, whatever that procedure is. Before we can see how this kind of restricted memory[1] encoding impacts the overall theory of word learning, we must first highlight some important facts about how syntax impacts word learning.

## 3. On syntactic bootstrapping

Given the insufficiency of the extralinguistic context as an information source about word learning, some additional source of information is required to provide the critical information. It is syntax to the rescue (Gleitman, 1990; Landau & Gleitman, 1985). In Gleitman's words, "the structure of the sentence that the child hears can function like a mental zoom lens that cues the aspect of the scene the speaker is describing" (Gleitman & Gleitman, 1992). The idea here is that there are principles of grammar that relate verb meaning and verb distribution in systematic ways (Fillmore, 1970; Gallistel, 1990; Levin, 1993). A learner equipped with these principles can then use the syntactic structure in which a verb occurs as evidence about the kind of meaning it is likely to have.

One influential variant of this idea (Fisher, Gertner, Scott, & Yuan, 2010; Gleitman, 1990; Lidz, Gleitman, & Gleitman, 2003; Naigles, 1990) posits (among other things) the correlation in (1), and says that this guides children around the age of 2 in acquiring the meanings of action verbs.

(1) Participant-to-Argument Matching (PAM): The core Noun Phrase (NP) dependents in a clause (subject, object, indirect object) correspond one-to-one with the participant roles in the event concept it expresses.

An intransitive clause, such as (2), has one core NP dependent, or argument, while a transitive, such as (3), has two. According to PAM, therefore, intransitive and transitive clauses express event concepts with, respectively, one and two "participant roles."

(2) Lee swam.
(3) Lee carried Mo.

The term "participant role" has no agreed-upon definition, but for the purposes of this paper, it suffices to have a rough idea that accords with its uses in the relevant literature. The participant roles for a verb, V, are a subset of those entailed by the event concept it expresses; they are those that are foregrounded psychologically whenever we view an event as a V'ing. Any event that verifies (2), for example, is related to a time, a place, a body of water, a swimmer, and many other things. These roles are entailed. But only its relation to the swimmer would ordinarily be called a participant role, as this role seems to enjoy some psychological privilege, in being foregrounded when we view an event as a swimming. Thus, participant roles are like the lines in a quick sketch of a face. The sketch represents the whole face, complete with pupils, ears, and wrinkles, but the lines make just a few of those features explicit, perhaps the eyes, nose, and smile (Williams, 2015).

Now consider what PAM can do for the learner. She hears a sentence—more precisely, a string of sounds in her language—with a novel word, like *satch* in (4). Ideally, she recognizes that this is a clause with two NP arguments and that *satch* is its verb.

(4) Lee satched Mo.

Guided by PAM, she will then understand that (4) describes an event that is readily viewed under a two-participant concept. If just some of the events that are salient to her can readily be viewed in these terms, then only these can be what the sentence is about, and this will restrict what *satch* might mean. For example, suppose that only two events are salient: Lee hugging Mo, and Lee giving Mo a toy. Then, if the learner views the hugging but not the giving under a two-participant concept, she can make a useful inference—huggings but not givings are included among the satchings, limiting the meaning of *satch*.

Much prior work has tested PAM in children between 18 and 40 months (Arunachalam & Waxman, 2010a, 2011; Brandone, Addy, Pulverman, Golinkoff, & Hirsh-Pasek, 2006; Fisher, 1996, 2002; Fisher, Gertner, Scott, & Yuan, 2010; Golinkoff, 1996; Hirsh-Pasek et al. 1996, Lidz et al., 2003; Naigles, 1990, 1996; Naigles, 1996; Noble, Rowland, & Pine, 2011; Yuan & Fisher, 2009; Yuan, Fisher, & Snedeker, 2012). The experiments are mostly variations on the same theme. A child is presented with a novel verb, like *dax* or *gorp*, in either an intransitive or a transitive clause. Yuan et al. (2012) use the stimuli in (5).

(5) a. He is gorping.
   b. He is gorping him.

Using measures such as duration of eye gaze, researchers measure the child's response to the display of two scenes, as a function of whether they hear the transitive or intransitive clause. The two scenes differ in the number of participants, two or one, implied by what seems to the experimenter like their most natural description. For example, the two-participant scene might have one person causing another to lean over by pushing against his back, while the one-participant scene just has people waving their arms, or leaning

over (Naigles, 1990; Yuan et al., 2012). Which scene is the child more likely to regard as verifying the given sentence?

The child guided by PAM, if she perceives the scenes under the intended descriptions, will assume that the transitive clause cannot be verified by the one-participant scene. It has more arguments than the scene, as we view it, has participants. Nearly all experimental results are consistent with this prediction.[2]

A theory of syntactic bootstrapping based in PAM holds that there is information in a single usage of a verb about its meaning, that the clause structure of a sentence carries information about the meaning of the verb in that sentence. This may be so, but this information is necessarily weak. It is weak because for aspects of meaning at a finer grain size than participant structure, the correlations between syntax and semantics are based on patterns that can only be seen across multiple frames (Fisher, Gleitman, & Gleitman, 1991; Levin, 1993; White, Hacquard, & Lidz, 2017).

Fillmore (1970) famously observed that there are systematic differences between verbs of contact and verbs of change of state. For example, contact verbs allow their instrument role to occur as the direct object, unlike change of state verbs:

(6) a. Kim hit the fence
    b. Kim hit the stick against the fence

(7) a. Kim broke the fence
    b. *Kim broke the stick against the fence (≠Kim broke the fence with the stick)

And change of state verbs allow their patient to be the subject of an intransitive, unlike contact verbs:

(8) a. The fence broke
    b. *The fence hit

These kinds of patterns can be found throughout the lexicon. Indeed, Levin (1993) identifies over 150 semantic classes based on patterns of distribution. Fisher, Gleitman, and Gleitman (1991) validated some of these patterns experimentally by asking participants to make semantic similarity judgments and acceptability judgments and found that verbs that were judged more similar also showed similarities in the sentence frames that they occurred in naturally. White et al. (2017) extended this method further by showing that even within a narrow range of verbs (the attitude verbs), meaning similarity was highly predictive of distributional similarity.

For attitude verbs (i.e., those referring to mental states), the extralinguistic context is perhaps even less informative, since the context of use provides little to no information about the particular mental states that one holds in relation to some state of affairs. People are always thinking (at least if they are awake), and the vast majority of utterances containing the word *think* might just as well have left it out. For example, the situation prompting me to say, "I think this paper is horribly written" is precisely the same situation in which I might say, "This paper is horribly written." Worse, one often uses mental

state verbs when the mental state is not even the point of the utterance. Thus, we say, "I think it's time for lunch," in order to communicate our hunger and get the meeting we are in to end. And even worse, different mental state verbs can lead to the same communicated content, despite differences in lexical meaning. Thus, the situations prompting a parent to say, "I think it's time for bed" are just the same as those prompting, "I want you to go to bed," and the message communicated to the child is the same, regardless of which mental state verb was used.

Here, again, syntax may be a helpful source of information allowing us to overcome the insufficiency of context (Gleitman, 1990). For example, Hacquard and Lidz (2018) argue that learners can use the form of a complement clause to infer something about its meaning. Verbs whose complements look like declarative main clauses in the language refer to representational mental states, that is, those that express judgments of truth (Bolinger, 1968). Verbs whose complements do not look like declarative main clauses are likely to pick out preferential mental states, that is, those relating to desires. Thus, a verb with a tensed complement in English is likely to label belief and a verb with an untensed complement in English is likely to label a desire. However, because the surface properties that distinguish declarative main clauses vary from language to language, a learner must simultaneously track properties of main clauses and properties of embedded clauses in order to discover the subclasses of attitude verbs, and to overcome the problems of inferring the content of a mental state from the extralinguistic situation.

Importantly, in all of this work the semantic classes are identified based on *patterns* of distribution. In the vast majority of cases, single frames carry very little information. For example, a simple transitive clause does not entail anything about the verb's meaning since verbs of contact, change of state, and perception are all compatible with this frame.[3] It is only in the context of additional frames that the informational content of the transitive can be identified (Levin, 1993). Similarly, in the domain of English attitude verbs, taking a tensed complement is compatible with both representational verbs and factive verbs, and the addition of embedded question complements will separate the factives from the nonfactives, independent of representationality (White et al., 2017). From the perspective of syntactic bootstrapping, it is only by seeing a verb in multiple syntactic environments that a learner can see the information about finer-grained semantic distinctions that are cued by the syntax.

Gleitman put it this way: "The range of subcategorization frames has considerable potential for partitioning the verb set semantically, and language learners have the capacity and inclination to recruit this information source to redress the insufficiencies of observation," (Gleitman, 1990, p.27). On this view of syntactic bootstrapping, the information about a verb's meaning is carried in the learner's memory of the set of frames it has previously occurred in. When children hear a verb used in several syntactic structures, they can use their knowledge about how these structures relate to the meaning in order to pinpoint the intended meaning of a given utterance. But this inference from structure to meaning is only possible if the learner has stored in an accessible format the set of frames associated with a given verb. The learner carries the history of a verb in memory and uses this history to select a preferred interpretation from among the several

possibilities made viable by observation of the extralinguistic contexts of use. The syntactic zoom lens comes not from single sentences, but from the history of sentences that the verb has occurred in previously (Landau and Gleitman, 1985, p. 114ff). That learners store these frames is evident in both adult and child psycholinguistic studies showing that incremental structure building is affected by knowledge of the frequencies with which verbs occur in particular frames (Lassota, Omaki, & Franck, 2016; Lidz, White, & Baier, 2017; Trueswell & Kim, 1998; Trueswell, Tanenhaus, & Kello, 1993).

## 4. The quandary of memory

So here is where we stand. Investigations of cross-situational learning indicate that when it comes to learning new words, we do not carry forward facts about the extralinguistic context of use. We do not assign some probability to possible meanings for the word that we do not choose in the context. Instead, we take a stab at the meaning and hope for the best. From the perspective of child learners, this seems like an excellent way to proceed. It allows them to avoid the problem of getting lost in the space of possible correlations between the word and the world. And it allows them to avoid an inevitable back off to general meanings like "thing" or "exist" when they find no reliable correlations between the world and the word at a finer grain size.

On the other hand, for a syntactic bootstrapping procedure for verb learning to work, it requires that we remember details of syntactic distribution across multiple uses. This requirement follows from the nature of the problem space. Single frames do not carry large amounts of information about verb meaning. Instead, relevant semantic classes hold together by virtue of distributional similarities among the members across a large set of frames. Thus, it is in the learner's interest to remember as much as possible about a verb's syntactic distribution. Only by doing so will she be able to capitalize on the information in that distribution in order to zero in on the meaning.

So, by virtue of what do child learners remember facts about the contexts in which a word is used? What causes them to forget many facts about the extralinguistic environment and to remember many facts about the linguistic environment, especially given that the meaning of a word resides, at least to some significant degree, in the way that that word applies to situations in the world? It strikes me that there are two kinds of answers to this question, but only one of them is plausible.

However, before we can address the cause of the asymmetry between referential and linguistic environment, it is important to see that this asymmetry is not merely the noun–verb distinction in disguise. It is true that the propose-but-verify hypothesis was developed primarily in the domain of noun learning and that syntactic bootstrapping has been deployed primarily in the domain of verb learning. However, the difference we are discussing cross-cuts that distinction.

First, many infant studies on the nature of syntactic bootstrapping are tasks about reference. That is, in studies where infants see a scene paired with a single linguistic description such as "She's gorping it," the tests typically involve identifying the referent of the

verb by seeing which of two scenes also counts as a "gorping." Just like the cross-situational learning tasks, these studies do not address the conceptual content of the acquired word or what range of events would count as a gorping. Instead, they simply ask which of the two previously viewed candidate referent events falls under that concept. In addressing the role of syntax in word learning, they ask whether the linguistic environment paired with a scene provides information about the referent event. At this stage it is unknown whether details of the extralinguistic environment are retained for subsequent refinement or rejection of the child's hypothesized meaning for verbs, but the very same issues arise for verb learning as in noun learning: When a scene provides evidence about a word's referent, how much of the extralinguistic context of use is retained in episodic memory for subsequent inspection in linking the word-form to a meaning?

These studies of how syntax contributes to identifying a referent event differ from those that examine the joint use of syntactic information to guide acquisition of word meaning. Studies emphasizing the role of multiple frames only occasionally have examined how multiple frames relate to identifying referent events (Bunger & Lidz, 2004; Naigles, 1996). Instead, where multiple frames arise, it is in the service of showing that semantic features are not carried by single syntactic frames, but rather by clusters of frames (Fisher, Gleitman, & Gleitman 1993; White et al., 2017). Thus, even in the domain of verb learning, we can differentiate studies about the role of syntax in identifying a candidate referent from those about the role of syntax in partitioning the class of verbs into syntactic and semantic subcategories. And so, the problem of identifying what information gets carried forward for learning is independent of the noun–verb distinction.

Second, while syntactic bootstrapping effects have been predominantly associated with verb learning, it is undoubtedly the case that syntactic environment is a cue to word meaning across several classes. Waxman and Booth (2001) show that infants as young as 14 months will treat a novel word presented as an adjective as referring to an object property but will treat a novel noun as referring to an object kind. Syrett and Lidz (2010) extend this finding to show that different adverbs lead to different inferences about novel adjective meaning in 30-month-olds. Wellwood, Gagliardi, and Lidz (2014) show that 3-year-olds treat a novel superlative presented as a determiner as having a number-based meaning, but they treat a novel superlative adjective as having a property-based meaning. He and Lidz (2017) have shown that by 18 months, infants will treat a novel verb as labeling an event, but a novel noun as labeling an object kind. And Fisher et al. (2006) show that infants will treat a novel preposition as labeling a spatial configuration, but treat a novel noun in the same context as labeling an object. Thus, the use of syntactic environment as cue to meaning cross-cuts a wide variety of grammatical categories and subcategories. Again, the difference between referential information and syntactic information does not reduce to the noun–verb distinction.

So, if the asymmetry between linguistic and extralinguistic information in word learning is not just an artifact of the noun–verb distinction, then what is the cause of this asymmetry? On one view, learners *learn* to ignore facts about the world and to capitalize on syntactic distribution (Smith, 2000). Here, in classic associationist fashion, the learner begins by looking for correlations between anything and everything. Upon discovering

that linguistic environment is a good cue to meaning and that the extralinguistic environment is not so trustworthy, they downweight the signal from the nonlinguistic context and upweight the information in the linguistic context. Learners thus discover what they can safely ignore. This view has the virtue (if you think it is a virtue) of endowing learners only with the ability to notice features of the linguistic and nonlinguistic environment and to tune their sensitivity to those features for solving whatever learning problems they need to solve. On the other hand, it does not really help with the original problem of the insufficiency of context. How will the child come to associate certain syntactic features with the relevant semantic features, if the meanings of the words are not themselves directly observable? The utility of syntactic bootstrapping lies in the fact that the extralinguistic environment is a poor source of information about meaning. But to discover links between syntax and semantics without being given clues to those links in advance would require the learner to be able to identify the meanings of many words via some route other than syntactic bootstrapping. It thus seems safe to reject the associationist explanation for the asymmetry between contextual and syntactic memory.

The alternative view may not surprise you. Learners value syntactic information over extralinguistic information because they expect this information to be relevant. On this view, children simply expect that facts about syntactic distribution will bear on questions of meaning. This expectation could be quite specific, with precise principles of mapping making up the hypothesis space for how syntax and semantics link together (Lidz, Gleitman, & Gleitman, 2004). Such principles might include the expectations that patients occur in object position, that agents occur in subject position (Lidz et al., 2017; Williams, 2015), or that belief-type verbs take complement clauses that resemble main clauses (Hacquard & Lidz, 2018). Of course, there may be alternatives, but the fundamental feature of this approach is a rationalist view of learning, in which learners' antecedent knowledge of the problem space shapes the way they take in information (Chomsky, 1965).

This rationalist approach to the asymmetry between contextual and syntactic memory acknowledges that aspects of grammatical architecture are intrinsically related to how we learn specific languages. The approach leaves as a matter of investigation just how richly such architectural information is represented. And it is independent of the particular memory representations that are implicated, beyond the distinction between memory for episodic details and memory for linguistic context. How linguistic memory is encoded and how it is partialed out from nonlinguistic memory is an important research question, but immaterial to the ideas presented here.

More important, this way of thinking about syntactic bootstrapping connects to a rich set of discoveries about how learning works in other creatures, beginning with von Frisch (1967). Throughout the animal kingdom, learning is guided by mechanisms that anticipate the structure of the learning problems that animals will face (Chomsky, 1975; Gallistel, 1990; Gould, 1986; Tinbergen, 1951). Bees know innately that color is a better cue to a food source than shape is (Menzel, Erber, & Mashur, 1974). Rats know that food can make them sick, but noises cannot (Garcia & Koelling, 1966). And humans know that syntactic distribution is a good cue to verb meaning (Gleitman, 1990). Because human

children are built to expect meaning to be reflected in syntactic structure, they know what information they can safely purge from their memories in trying to identify the meaning of a word. And quite surprisingly, it is the referential information that is the least likely to be retained.

## Notes

1. In some circumstances, memories for extralinguistic context can help learning. For example, if a word is always used at the zoo, then it is a good bet that it refers to an animal (Dautriche & Chemla 2014, Roy, Frank & Roy 2012). This kind of data shows that the episodic memories surrounding the use of an unknown word are not completely empty. Such facts do not argue against the view described above whereby the details of episodic memory are restricted in the influence they have over word learning.
2. We leave aside for the purposes of this paper discussion of what happens in the intransitive case (results are mixed) and whether this kind of data is uniquely supportive of PAM, as opposed to some alternative mapping principles. See Perkins, He, and Lidz (2017) for discussion.
3. Of course, saying that no specific semantic features are entailed by a transitive clause does not mean that learners should not make inferences about likely meanings given a transitive clause.

## References

Arunachalam, S., & Waxman, S. R. (2010a). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, *114*(3), 442–446.

Arunachalam, S., & Waxman, S. R. (2011). Grammatical form and semantic context in verb learning. *Language Learning and Development*, *7*(3), 169–184.

Baldwin, D. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, *62*, 875–890.

Bloom, P. (2001). *How children learn the meaning of words*. Cambridge, MA: MIT Press.

Bolinger, D. (1968). *Aspects of language*. New York: Harcourt, Brace, and World.

Brandone, A., Addy, D.A., Pulverman, R., Golinkoff, M., & Hirsh-Pasek, K. (2006). One-for-one and two-for-two: Anticipating parallel structure between events and language. In D. Bamman, T. Magnitskaya & C. Zaller (Eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development* (pp. 36–47). Cambridge, UK: Cascadilla Press.

Bunger, A., & Lidz, J. (2004). Syntactic bootstrapping and the internal structure of causative events. In A. Brugos, L. Micciulla & C. Smith (Eds.), *Proceedings of 28th Annual Boston University Conference on Language Development* (pp. 74–85). Cambridge, UK: Cascadilla Press.

Carey, S. (2009). *Origin of concepts*. Oxford, UK: Oxford University Press.

Chomsky, N. (1959). *Review of verbal behavior. Language*, *35*(1), 26–58.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1975). *Reflections on language*. New York: Pantheon.

Clark, E. V., & Amaral, P. M. (2010). Children build on pragmatic information in language acquisition. *Language & Linguistics Compass*, *4*(7), 445–457.

Dautriche, I., & Chemla, E. (2014). Cross-situational word learning in the right situations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 892–903.

Fillmore, C. (1970). The grammar of hitting and breaking. In R. A. Jacobs, & P. Rosenbaum (Eds.), *Readings in English transformational grammar*. Waltham, MA: Ginn & Co.

Fisher, C. (1996). Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. *Cognitive Psychology*, *31*(1), 41–81.

Fisher, C. (2002). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science*, *51*, 55–64.

Fisher, C., Gertner, Y., Scott, R., & Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(2), 143–149.

Fisher, C., Gleitman, H., & Gleitman, L. R. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology*, *23*, 331–392.

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.

Garcia, J., & Koelling, R. A. (1966). The relation of cue to consequence in avoidance learning. *Psychonomic Science*, *4*, 123–124.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*, 135–176.

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

Gleitman, L. R., & Gleitman, H. (1992). A picture is worth a thousand words, but that's the problem: The role of syntax in vocabulary acquisition. *Current Directions in Psychological Science*, *1*, 31–35.

Golinkoff, R. (1996). *The intermodal preferential looking paradigm: A window onto emerging language comprehension*. Cambridge, MA: MIT Press.

Goodman, N. (1955). *Fact, fiction, and forecast*. Cambridge, MA: Harvard University Press.

Gould, J. L. (1986). The biology of learning. *Annual Review of Psychology*, *37*, 163–192.

Hacquard, V., & Lidz, J. (2018). Children's attitude problems: bootstrapping verb meaning from syntax and pragmatics. *Mind and Language*, *28*, https://doi.org/10.1111/mila.12192.

He, A. X., & Lidz, J. (2017). Verb learning in 14- and 18-month-old English learning infants. *Language Learning and Development*, *13*, 335–356.

Hirsh-Pasek, Kathy, & Golinkoff, Roberta (1996). *The intermodal preferential looking paradigm: A window onto emerging language comprehension*. Cambridge: MIT Press.

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

Landau, B., & Gleitman, L. (1985). *Language and experience. Evidence from the blind child*. Cambridge, MA: Harvard University Press.

Lassota, R., Omaki, A., & Franck, J. (2016). Developmental changes in misinterpretation of garden-path wh-questions in French. *Quarterly Journal of Experimental Psychology*, *69*, 829–854.

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

Lidz, J., Gleitman, H., & Gleitman, L. (2003). Understanding how input matters: Verb learning and the footprint of universal grammar. *Cognition*, *87*(3), 151–178.

Lidz, J., Gleitman, H., & Gleitman, L. (2004). Kidz in the hood. In S. R. Waxman & G. Hall (Eds.), *Weaving a lexicon*. Cambridge, MA: MIT Press.

Lidz, J., White, A. S., & Baier, R. (2017). The role of incremental parsing in syntactically conditioned word learning. *Cognitive Psychology*, *97*, 62–78.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, *14*, 154–173.

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences*, *108*, 9014–9019.

Menzel, R., Erber, J., & Mashur, T. (1974). Learning and memory in the honeybee. In L. Barton-Browne (Ed.), *Experimental analysis of insect behavior*. Berlin: Springer.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, *17*(02), 357–374.

Naigles, L. R. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition*, *58*(2), 221–251.

Noble, C. H., Rowland, C. F., & Pine, J. M. (2011). Comprehension of argument structure and semantic roles: Evidence from English-learning children and the forced-choice pointing paradigm. *Cognitive Science*, *35*(5), 963–982.

Perkins, Williams, He, Bjornsdottir, & Lidz, J. (2017). A new test of one-to-one matching between arguments and participants in verb learning. *Presented at Boston University Conference on Language Development*.

Roy, B. C., Frank, M. C., & Roy, D. (2012). Relating activity contexts to early word learning in dense longitudinal data. Proceedings of the 34th Annual Meeting of the Cognitive Science Society.

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Smith, L. (2000). Learning how to learn words: An associative crane. In R. Golinkoff & K. Hirsh-Pasek (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 51–80). New York: Oxford University Press.

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science*, *14*(1), 29–56.

Syrett, K., & Lidz, J. (2010). 30-month-olds use the distribution and meaning of adverbs to interpret novel adjectives. *Language Learning and Development*, *6*, 258–282.

Tinbergen, N. (1951). *The study of instinct*. Oxford, UK: Oxford University Press.

Trueswell, J., & Kim, A. (1998). How to prune a garden path by nipping it in the bud: Fast priming of verb argument structure. *Journal of Memory and Language*, *39*, 102–123.

Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology*, *66*(1), 126–156.

Trueswell, J. C., Tanenhaus, M., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning Memory and Cognition*, *19*(3), 528–553.

von Frisch, K. (1967). *The dance language and orientation of bees*. Cambridge, MA: Harvard University Press.

Waxman, S. R., & Booth, A. (2001). Seeing pink elephants: 14-month-olds' interpretations of novel adjectives. *Cognitive Psychology*, *43*, 217–242.

Waxman, S. R., & Lidz, J. (2006). Early word learning. In D. Kuhn, & R. Siegler (Eds.), *Handbook of child psychology: Cognition, perception and language* (pp. 299–335). New York, NY: John Wiley & Sons.

Wellwood, A., Gagliardi, A., & Lidz, J. (2014). Syntactic and lexical inference in the acquisition of novel superlatives. *Language Learning and Development*, *12*, 262–279.

White, A. S., Hacquard, V., & Lidz, J. (2017). Semantic information and the syntax of propositional attitude verbs. *Cognitive Science*, *42*, 416–456.

Williams, A. (2015). *Arguments in syntax and semantics*. Cambridge, UK: Cambridge University Press.

Yuan, S., & Fisher, C. (2009). Really? She blicked the baby? Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science*, *20*(5), 619–626.

Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: Simple structural cues to verb meaning. *Child Development*, *83*(4), 1382–1399.