

Children's Use of Syntax in Word Learning

Jeffrey Lidz¹
University of Maryland

Abstract: This chapter investigates the role that syntax plays in guiding the acquisition of word meaning. I review data that reveals how children can use the syntactic distribution of a word as evidence for its meaning and discuss the principles of grammar that license such inferences.

Keywords: syntactic bootstrapping, word learning, thematic relations, attitude verbs, action verbs

¹ Many thanks to Alexander Williams, Valentine Hacquard and Laurel Perkins, all of whom have helped to clarify my thinking on these issues over the past 5-10 years. This work was supported in part by NSF grant BCS-1551629.

One might think that it is too obvious to mention that when it comes to language use, humans are free to say whatever they want to, or, in the case of children and politicians, whatever they can get away with. What we say is determined by our communicative goals - whether we are trying to provide information, get information, lie or mislead, issue commands, provide warnings, make promises or simply to speculate about the future (Austin 1962, Searle 1969, Murray & Starr 2020; see also Schwarz & Zehr, this volume). Our communicative goals, in turn, are determined by a host of factors relating to our beliefs and desires, including our beliefs about and desires for the people we are speaking to, in hopelessly complex ways.

With this complexity in mind, it is somewhat odd that our naïve intuitions about word learning seem to emanate from the idea that speakers operate in a descriptive mode. Following this intuition, we say the word “dog” when there are dogs around, and hence a child who doesn’t yet know the meaning of this word-form merely has to look around and notice the contingency between the form and the thing. My own experiences as a parent reveal the hopelessness of this naïve theory. My children use the word “dog” in a never-ending harangue about why we do not have one. Consequently, a dog novice would have a hard time figuring out the meaning of “dog” based on observations of what’s happening in the physical world when people in my family use that word. Utterances of the word “dog” are conditioned not by dogs *per se*, but by thoughts of dogs and the utility of getting our interlocutors to have dog-thoughts given our communicative goals.

Learning the meaning of a word, then, must involve sifting through all of that psychology. If word learners were psychics, word learning would be trivial. But given

that word learners are just like the rest of us in lacking the ability to read minds, there must be some kind of evidence that they could rely on that could at least point them in the right direction. In this chapter, we consider one kind of evidence that serves to focus child word learners in on the relevant dimensions of meaning – syntactic distribution. We will see that a word's syntax provides evidence about its meaning (and why it does so) and that from the earliest stages of word learning, children are sensitive to that information. This “syntactic bootstrapping” theory of word learning does not solve the problem of word learning, but it does place some constraints on the learner that, when paired with other capacities, including our ability to estimate the goals underlying people's actions and utterances, make significant contributions to our ability to acquire word meanings.

1. The puzzle of word 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 (cf. Chomsky 1959). Even if a speaker is talking about the here and now, and even if learners could somehow zero in on the part of the world being picked out by a novel word, recognizing, for example, that the speaker was attending to the rabbit, this reference would be consistent with a broad range of concepts with overlapping extensions (e.g., Peter Rabbit, animals, mammals, rabbits in the yard, rabbits or black holes, rabbits with more than one ear, things physically identical to rabbits, etc.; Goodman 1955). Indeed, it was precisely this kind of

indeterminacy that led Quine to reject the very idea of word meanings being constant across speakers.

We can respond to these philosophical problems by noting that humans have perceptual, conceptual and linguistic abilities that limit the space of plausible word meanings (Chomsky 1971, Fodor 1975). These abilities lead learners to experience the world through the same conceptual apparatus as the speakers producing those words, making meanings like “rabbit or black hole” or “rabbit in the yard” unlikely candidates for word meanings (Gleitman 1990, Spelke 1990, Markman 1990, Waxman & Lidz 2006, Carey 2009). Moreover, these conceptual burdens may be further 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 (see also Jackendoff, this volume). And, as Landau & 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 et al 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. Overall, participants were remarkably bad at this task. For nouns, they 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 (e.g., Smith 2000). Pull on that thread and the meaning will be revealed. However, participants in these experiments actually had 6 attempts to make their guesses, and improvement across trials was very weak at best. 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. This fact could reasonably lead to the expectation 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. And empirically, there is little evidence that supports the idea of cross-situational learning and lots of evidence against it (see Gleitman & Trueswell, this volume, for review).

But perhaps more importantly, once we recognize that we only rarely use sentences to describe the events before our eyes (afterall, the person we are talking to can see what's going on as well as we can), the expectation that we learn the words by placing them in correspondence with the bits of the world that they refer to seems like an idea we never should have taken all that seriously in the first place.

2. Why syntax might help

Landau and Gleitman (1985) described the language acquisition of a blind child, noting that she learned the meanings of *look* and *see*, despite lacking visual inputs. She had learned them as perception verbs, but mapped their meanings onto haptic rather than visual perception. For example, she responded to the request to "Look up!" by reaching upwards with her hands, whereas sighted children wearing a blindfold turned their

heads upwards. So, Landau & Gleitman initially reasoned that “look” must have been used by the child's parents when the object to be apprehended was near to the child, and hence available for manual inspection. But dividing verbs along this situational dimension did not distinguish these verbs from other verbs, such as *touch* or *hold*, that had nothing to do with perception. So instead, they hypothesized that the syntactic distribution of *look* compared to *touch* and *hold* drives this difference in interpretation. *Look* occurs in different syntactic environments than *touch* or *hold*, as shown in (1-3).

- (1) a. *Look where I'm going*
b. **Touch/hold where I'm going*
- (2) a. *Look at that picture*
b. **Touch/hold at that picture*
- (3) a. *Look down*
b. **Touch/hold down*

Perhaps the blind child in this study first used the syntactic distributions to distinguish perception verbs from other verbs (using features like compatibility with a clausal complement), and then to use the near/far distinction to explain how these verbs came to be associated with manual inspection. Thus, the syntactic context provided the information that the verbs had something to do with perception and the extralinguistic context provided further information about the kind of perception that was relevant.

This kind of learning procedure, using syntax to zero on the semantic domain and then combining this information with an understanding of the extralinguistic context, could generalize beyond blind children. Any learner armed with knowledge of how

meaning projects onto syntax could use the syntax of a word to zero in on a specific semantic domain and then use the extralinguistic context to further specify the word's meaning. The extralinguistic context provided information only after the initial syntactic/semantic partition. This initial partitioning is what has come to be known as *syntactic bootstrapping* (Landau & Gleitman 1985, Gleitman 1990).

3. Two dimensions of syntactic bootstrapping

The term “syntactic bootstrapping” actually covers two closely related ideas. The first idea was that syntax is a constraining factor in narrowing reference. Given the lack of information in a scene by itself about a verb’s meaning, the syntactic structure of a clause could direct a learner to have a particular perspective on the scene. That perspectival information would then allow the learner to identify the event labeled by the sentence and hence the meaning of the verb in that sentence. 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 second idea was that semantic properties of a verb’s meaning stand in a principled relation to its syntactic distribution (Fillmore 1970, Jackendoff 1972, Levin 1993, see also Jackendoff, this volume). Consequently, observations of a verb used in a range of syntactic environments could help the learner identify the semantic properties of that verb’s meaning.

These two ideas about syntactic bootstrapping are linked together by the fundamental idea that the less information there is in the world about a verb’s meaning, the more information will be found in the syntactic distribution (Snedeker & Gleitman 2004). At the far extreme are mental state verbs, whose meanings are hidden from

observers because their contents reside inside the mind. For such verbs, the syntactic environment is highly informative both as evidence about the verb in a single sentence and about the verb across many different sentence contexts (Papafragou, Cassidy & Gleitman 2007, Hacquard & Lidz 2019).

4. Syntax and Event Reference

As noted above, the meaning of a sentence/verb reflects the perspective of the speaker on an event. Events are not defined simply by their physical properties, but rather by a perspective. For example, a scene in which a bunny runs and follows the path of an elephant ahead of him might be described in either of the two ways in (4).

- (4) a. The bunny is chasing the elephant
- b. The elephant is fleeing the bunny

As physical events, chasings and fleerings are indistinguishable. Where they differ is in the perspective the speaker takes towards them. A transitive clause that identifies the agent and patient as subject and object can therefore provide a learner with the appropriate perspective on the event to make the clause and verb meanings more accessible. Hearing a sentence like (4a) prior to knowing the verb's contribution to meaning would allow the learner to adopt a perspective on the event that makes the bunny the agent. Hence, the verb in that sentence should describe the event as a chasing. By the same token, hearing (4b) would provide a perspective in which the elephant is the agent, hence a fleeing but not a chasing (Fisher et al 1994, Nappa et al 2009).

In support of this theory, numerous studies have shown that infants can use the syntactic structure of a sentence as evidence about what event it describes, and

consequently what the verb in that sentence means. For example, Naigles (1990) presented 25-month-olds with a novel verb in the context of a complex scene with two parts: a causal part in which a duck pushes a bunny over, and a non-causal part in which the duck and bunny each wheel their arms independently. While they watched this scene, the children heard a novel verb used either transitively (The duck is gorp-ing the bunny) or intransitively (The duck and the bunny are gorp-ing). Naigles then separated these two parts of the scene into two different videos each showing only one of the two parts. She measured infants' looking preferences when they were asked to "find gorp-ing" as a function of whether they were initially familiarized to the novel verb in a transitive or intransitive clause. Infants who had heard the transitive looked longer at the pushing scene, and infants who heard the intransitive looked longer at the arm-wheeling scene. Infants were thus sensitive to the syntactic frame of the novel verb, inferring that *gorp* in a transitive frame was more likely to label the causal event, whereas *gorp* in an intransitive frame was more likely to label the non-causal event.

This basic finding has been reproduced in various ways, confirming that infants as young as 22 months are sensitive to transitivity, and will reliably infer that a novel transitive verb labels a causal event (Arunachalam & Waxman, 2010; Arunachalam & Dennis 2018; Brandone et al., 2006; Fisher et al., 2010; Noble et al., 2011; Pozzan et al., 2015; Yuan & Fisher, 2009; Yuan et al., 2012).

Moreover, children are able to draw this inference on the basis of distributional information alone. Yuan & Fisher (2009) familiarized 28-month-olds with short dialogues containing novel transitive or intransitive verbs, without any informative visual context. At test, infants were then asked to identify the referent of the novel verb (e.g. *Find*

blicking!) while viewing two candidate events, one causative and one non-causative.

Infants who had heard the transitive dialogues looked longer at the causative event than infants who had heard the intransitive dialogue. This indicates that they had tracked the syntactic properties of the novel transitive verb and used those properties to draw inferences about its possible meanings, even without the support of referential context.

However, beyond Naigles' (1990) seminal study, further work has found inconsistent behavior with intransitive clauses. Infants who hear novel verbs in intransitive frames do not show a reliable preference for events intended to be viewed with one participant as opposed to two (e.g. Arunachalam & Waxman, 2010; Noble, Rowland, & Pine, 2011; Yuan et al., 2012). Several methodological explanations have been proposed to account for these variable results with intransitive clauses. First, many studies use intransitive sentences with conjoined subjects (e.g. *The duck and the bunny are gorging*) in order to control the number of nouns across conditions. It is possible that infants may not reliably perceive these sentences as intransitive: if they mistake the conjoined subject for two separate arguments, this might lead them to infer a causative meaning for the verb (Gertner & Fisher, 2012; Yuan et al., 2012). Alternatively, it is possible that infants do not reliably perceive the presented scenes under the intended event representation. If infants conceptualize a scene of one actor pushing another as an event of two actors playing, then they might consider the intended "two-participant" scene a good referent for a novel intransitive verb (Brandone et al., 2006; Pozzan et al., 2015).

4.1 Explaining the effects of transitivity: one-to-one matching vs. thematic linking

Results showing that infants reached different interpretations of verbs presented in transitive vs. intransitive clauses have been used to support an influential hypothesis about *how* infants use the syntactic contexts that verbs appear in to draw inferences about their meanings. Under this hypothesis, infants take the noun phrases in a clause to be arguments, and expect the number of arguments in a clause to match one-to-one the number of participants in the event the clause describes (Fisher, 1996; Gleitman, 1990; Naigles, 1990; Lidz, Gleitman & Gleitman 2003, Fisher et al 2010). Thus, a transitive clause with two arguments should label an event perceived with two participants, whereas an intransitive clause with only one argument should label an event perceived with one participant. This is a potentially powerful learning strategy for infants at early stages of syntactic development because it requires very little syntactic knowledge. In order to narrow down the candidate events that a clause might refer to, infants need only to identify the number of nouns or noun phrases in the clause, and do not need to identify their thematic roles or hierarchical position in the clause.

On the other hand, there are several reasons to question whether this simple matching strategy is the appropriate way to capture effects of transitivity on verb meaning. First, it is not generally true of languages that participants and arguments stand in one-to-one correspondence. Verbs that describe events that entail three participants often only require two arguments:

- (5) a. Jesse robbed the train
- b. Bonnie stole the money
- c. Clyde took the jewels

In these examples, the events seem to have 3 main participants (the thief, the loot and the victim) but these verbs are simple transitives, regularly occurring in clauses with only 2 arguments. Thus, a child guided by a one-to-one matching strategy might be expected to learn meanings for these verbs that entail only two participants.

It is worth pausing for a moment to ask how one would know whether these examples are problematic for a bootstrapping theory based in one-to-one correspondence. Is it relevant that robbings entail loot if sentences with the verb *rob* do not require this participant to be named? The answer to this question depends on the event concepts that such events are viewed under, independent of language. If the loot in a robbery is entailed but not foregrounded as a participant in our event representations, then the fact that *rob* takes two syntactic arguments is not a problem. But if the loot is foregrounded as a participant, then the one-to-one-correspondence theory would predict that infants would have difficulty acquiring such a verb, since the two argument syntax would not align with the three argument concept. Moreover, it is not sufficient to note that *rob* is a transitive verb to decide that the loot is not a participant in a conceptual representation, as this would be question begging. One cannot take transitivity as evidence for conceptual adicity and also argue for a bootstrapping theory based in one-to-one correspondence; doing so presupposes the conclusion that it argues for. Thus, one would ultimately want independent evidence for the conceptual representations in order to test whether the one-to-one correspondence theory was correct (Williams 2015, Wellwood et al 2015).

Second, cross-linguistically speaking, these kinds of mismatches between participant relations and argument relations are common. As an extreme example,

nearly every St'át'imcets verb can occur in an intransitive clause (Davis 1997, Davis & Demirdache 2000, Davis 2010). Even verbs that seem to entail 3 participant roles, like the verb in (6), can occur in intransitive clauses without losing those entailments:

- (6) Q'amt kwskwim ,cxen
hit.with.projectile det.NAME
'Kwim ,cxen got beaned.'

The sentence in (6) is a basic intransitive clause, with neither null arguments nor valency reducing morphology like passive or middle. A child expecting one-to-one correspondence between participants and arguments would thus be severely misled by (6) about the meaning of the verb in that clause (Williams 2015). Because the one-to-one matching heuristic does not reflect a basic fact about the languages of the world, or about verbs in general, a bootstrapping theory using it as a basis will require learners to abandon it at some point in development as they acquire a theory of linking that is more richly structured. So, if the one-to-one matching heuristic is correct, we will need an additional theory detailing how it is abandoned in development, and on what basis.

Further, infants seem to have rich syntactic representations of clause structure and the relation between argument position and thematic relation from as young as 16-months. Lidz, White & Baier (2017) asked what infants know about the mapping between the syntactic position of an NP and its thematic interpretation. They presented events in which an agent used an instrument to affect an object. For example, they saw events in which a hand used a ruler to tap a traffic cone. While they saw these events, they heard either a simple transitive clause containing a novel noun in the direct object

position (*She's hitting **the tam***) or an intransitive clause containing a novel noun inside an instrumental Prepositional Phrase (*She's hitting with **the tam***). After several exposures to these sentences containing the novel noun, they were then shown the two objects (i.e., the ruler and the cone) and were asked "which one is the tam?". Sixteen-month-olds looked more at the cone in the transitive condition and they looked more at the ruler in the Prepositional Phrase condition. These results indicate that by 16-months, infants know how to identify the thematic role of a NP based on its syntactic position. Because infants drew different conclusions about the referent of the novel NP (and hence the meaning of the novel noun) as a function of its syntactic position, we can conclude that they build syntactic representations that contain more information than simply the number of nouns in the clause. In turn, this conclusion suggests that effects of clause structure in verb learning experiments might be driven by a richer representation of subject and object and the thematic consequences of these representations.

Perkins (2019) spells this idea out more fully, suggesting that the effects of transitivity could be explained by infants' initial expectations about thematic linking. Specifically, infants might begin language learning with an expectation that subjects of transitive clauses label agents and objects of transitive clauses label patients (Baker 1988, Dowty 1991, Fillmore 1970, Jackendoff 1972, Pinker 1984). If infants could identify the subject and object in a transitive clause, they could then infer that the clause labels not just any event seen as having two participants, but one in which the referent of the subject is the agent and the referent of the object is the patient. This kind of learning mechanism would support acquisition in the same way, independent of whether

the language being acquired was like English or St'át'imcets. And, it would allow for a continuous theory of the relation between argument structure and interpretation across development.

The idea that early verb learning is driven by expectations about thematic linking is supported by several empirical results. Gertner, Fisher, and Eisengart (2006) tested 24- and 21-month-olds' abilities to link syntactic position to thematic relation using a preferential looking task. Infants heard a transitive sentence (e.g. *The duck is gorging the bunny*) in the context of two causative scenes: one in which a duck pushed a bunny, and one in which the bunny pulled the duck. Both groups of infants looked preferentially at the scene in which the duck was the agent, indicating that they knew that the subject of a transitive clause labels the agent rather than the patient of a causal event. Furthermore, infants preferred the duck-agent and bunny-patient event even for sentences like *He is gorging the bunny*: here, they could only rely on the referent of the object because the subject does not identify a unique referent in the discourse. This indicates that infants knew that the object of a transitive clause labels the patient rather than the agent of a causal event. These infants were able to exploit relationships between argument position (subject vs. object) and argument roles (agent vs. patient) in order to constrain the inferences they draw about transitive verb meanings.

For intransitive verbs these relationships are more complicated: the subject of an intransitive clause can label either an agent (e.g. *John baked*) or a patient (e.g. *The bread rose*). These sub-classes of intransitives also display differences in meaning: intransitives whose subject is an agent tend to label actions of that agent, whereas intransitives whose subject is a patient tend to label changes undergone by that patient

(e.g. Fillmore, 1970; Levin & Hovav, 2005; Williams, 2015). Another line of work has asked whether children can draw these finer-grained inferences about verb meanings on the basis of the animacy and thematic role of the intransitive subject (Bunger & Lidz, 2004, 2008; Naigles, 1996; Scott & Fisher, 2009). For example, Bunger and Lidz (2004) familiarized 24-month-old infants to an event in which a girl bounced a ball with a tennis racquet, while they heard one of 4 types of linguistic input: transitive (the girl is pimming the ball), unaccusative (the ball is pimming), multiple frame (the girl is pimming the ball, the ball is pimming) or a no-word control (hey, look at that). At test, they were asked “where’s pimming now” while seeing the event broken into two parts – the girl patting the ball (but with no bouncing) or the ball bouncing on its own. In the transitive and no-word conditions, infants showed no preference at test. This suggests that the transitive clause by itself is not sufficient to identify the event as a contact event (hitting) or a change of state event (bouncing). However, in the unaccusative and multiple frame conditions, infants looked more at the bouncing event. This suggests that infants know that an intransitive clause with an inanimate subject is likely to label an event describing a change to that argument.² Thus, infants seem to be aware of the thematic relations associated with arguments in different syntactic positions.

Scott & Fisher (2009) found a similar pattern. They familiarized 28-month-olds with a dialogue in which a novel verb alternated between transitive and intransitive uses. Infants either heard the intransitive with an animate subject (e.g. *Matt dacked the*

² One should be cautious here about the identification of unaccusative clauses. Obviously, unaccusative verbs can take both animate and inanimate subjects and there is no guarantee that an intransitive clause occurring with an inanimate subject has an unaccusative verb. Nonetheless, subject animacy is strong probabilistic cue to the classification of intransitive verbs (Scott & Fisher 2009, Becker 2014).

pillow. He dacked), cueing unergativity, or an inanimate subject (e.g. *Matt dacked the pillow. The pillow dacked*), cueing unaccusativity. At test, infants heard the verb in a transitive frame in the context of two causative scenes: a caused-motion event in which a girl pushes a boy over, or a contact-activity event in which the girl dusts the boy with a feather duster. Infants who were exposed to the animate-subject intransitive dialogue preferred to look at the contact-activity event, whereas infants who were exposed to the inanimate-subject dialogue preferred to look at the caused-motion event. These infants were able to use cues to the thematic role of the intransitive subject, such as its animacy, to infer whether the novel verb labeled the action of an agent or the change undergone by a patient.

Perkins (2019) further distinguished this thematic linking hypothesis from the one-to-one matching hypothesis discussed earlier by exploring whether infants would allow a three participant event to be described with a two argument description. As noted above, in order for such an event to be a useful probe into these alternative bootstrapping theories, it is important to have an independent measure of the number of participant relations in the event concept. So, prior to this verb learning study, Perkins et al (2018) explored the event representations of 10-month-olds. Infants were habituated to an event in which a woman picks up a truck in an arcing motion while a man sits idly near the truck. After habituation, infants saw one of two changes to the event: a participant change or a manner change. In the participant change, the man has his hands on the truck prior to the woman picking it up. In the manner change, the woman slides the truck into her possession rather than picking it up. The physical difference between the habituation and test events were larger in the manner change than in the

participant change. Infants showed greater dishabituation to the participant change than to the manner change, suggesting that they viewed the man as a participant in the event.

Perkins (2019) therefore used the event in which the woman takes the truck from the man as a three participant event in a verb learning study with 20-month-olds. This 'taking' event was labeled with a novel verb either in a transitive clause (she's gonna *pim* the truck) or in an intransitive clause (the truck is gonna *pim*). After familiarization, infants were shown two different events: one in which the girl takes the truck from the boy vs one in which she moves the truck in same way, but with no boy present. The critical test is the transitive condition. If children expect clausal arguments and event participants to align one-to-one, then we would predict that the infants in the transitive condition would not naturally link "pim" to the three-participant event concept under which they see the event. Instead, this theory predicts that they would take the verb to label a two-participant event concept, something like MOVE or SLIDE. Thus, both test events should be equally good exemplars of pimming, predicting no difference in the test condition. If, on the other hand, infants take the subject to be the agent of the pimming event and the object to be the patient of the three participant concept that they naturally see the event under, then, they should think that the taking event is the only exemplar of pimming, and hence look more to that event. This is indeed what happened. Infants performed in line with the predictions of the thematic linking hypothesis and against the predictions of the one-to-one matching hypothesis. In the intransitive condition, both hypotheses predict no preference between the two events, and this is what was found. The intransitive results also rule out the possibility that

infants in the transitive condition simply chose the more familiar video. If that were the explanation for the transitive condition, the same pattern would have been seen in the intransitive condition, contrary to fact.

In sum, infants before their second birthday show an impressive ability to use information about clause structure to provide them with a perspective on an event, and hence to zoom in on relevant features of a verb's meaning. These abilities are driven by an early appreciation of the link between grammatical relations like subject and object and the thematic relations borne by the noun phrases in those positions. Infants as young as 20-months are able to identify the subject and object of a clause, to appreciate the thematic relations borne by the NPs in those positions, and to use that information to identify the meaning of a novel verb in that clause.

5. Distributional profiles, semantic features & propositional attitude verbs

To this point, we have seen that infants can use observations about a verb's syntactic environment to help identify which event in the world the verb labels. However, many sentences are not used to label events. As noted in opening, speakers do more with language than simply describe the world around them. And even when they do describe what's going on, much of what they are talking about leaves no detectable trace in the physical world that would serve as the "referent" of the expression. The very same scene might elicit any of the following utterances:

- (12) a. April is walking quickly down the street.
b. April is going home.
c. April is late.
d. April wants to get home soon.

- e. I think April is late for dinner.
- f. They expected April to be home already.
- g. Remember when we went sailing?

The differences between these sentences have little to do with what's happening at the moment of utterance, but rather with what thoughts the physical event triggers in the speaker and what the conversational goals of the speaker happen to be in producing the sentence. Of particular interest are sentences like (12d) and (12e), with propositional attitude verbs like *want* and *think*.

Propositional attitude verbs describe the contents of peoples' minds such as beliefs or desires about possible states of affairs. Such verbs present a special challenge to learners because they name internal states of speakers' minds (Gleitman, 1990; Gleitman et al., 2005). A large body of literature has argued that young children may have difficulty acquiring attitude verbs because they lack the mental state concepts that these verbs label; in particular, children fail in certain tasks to demonstrate the ability to represent others' beliefs (the so-called developing Theory of Mind, e.g. Astington & Gopnik, 1991; Flavell, Green, & Flavell, 1990; Gopnik & Wellman, 1994; Perner, 1991). However, more recent work finds that children's failure on these tests may be due to experimental and pragmatic factors rather than immature belief representations (e.g. Hansen, 2010; Z. He, Bolz, & Baillargeon, 2012; Helming, Strickland, & Jacob, 2014; Lewis, Hacquard, & Lidz, 2017; Onishi & Baillargeon, 2005; Rubio-Fernández & Geurts, 2012).

But even if children do have the ability to represent speakers' mental states, learning which verbs label these mental states is still no trivial matter. It is difficult to tell

when mental states rather than actions are under discussion: if a speaker uses a new verb, how does a child know whether the verb labels what someone is feeling or what someone is doing? In the human simulation study by Gillette et al. (1999), adults were particularly bad at identifying attitude verbs from the situations in which they were uttered; they could occasionally identify action verbs like *hit*, but almost never identified attitude verbs like *think*, *know* and *want*. This difficulty may be caused by the low conversational salience of beliefs and desires (Papafragou, et al 2007, Dudley 2017)

Papafragou, Cassidy & Gleitman (2007) showed that false belief contexts increase the situational salience of beliefs and the likelihood that children and adults guess that a novel verb labels a belief. This result suggests that false belief uses would be particularly informative contexts for learning that a verb labels the belief concept. However, Dudley et al (2017) show that approximately 75% of uses of *think* in child-directed English occur with first person subjects in present tense (see also Diessel & Tomasello 2001), contexts which do not lend themselves to false belief uses. These uses are more commonly indirect assertions, whose conversational function is to proffer the content of the embedded clause, potentially hiding even further the meaning of the verb *think*. Similarly, the verb *know* is most often used in indirect questions like “do you know where my keys are?”, where the illocutionary force of the entire sentence is to ask about the location of the keys, not about the addressee’s knowledge.

Attitude verbs do have a reliable syntactic signal to their meaning, however. Such verbs take full clauses as complements, whereas action verbs do not:

- (13) a. *Kim thought that Chris liked her*
b. **Kim danced that Chris liked her.*

Therefore, even though children may have difficulty identifying attitude verbs from the situational contexts in which they are used, they might be able to identify them through their syntactic distribution—specifically, by paying attention to which verbs take clausal complements (Fisher, Gleitman, & Gleitman, 1991; Gleitman et al., 2005, Papafragou et al., 2007).

Furthermore, differences in the clausal complements of attitude verbs might help children tell certain attitude verbs apart from each other. Attitude verbs fall into two major classes: representational and preferential (Bolinger 1968). The representational verbs, such as *think*, *know*, or *say*, express judgments of truth and present a picture of the world. The preferential verbs, such as *want* or *demand*, convey preferences about how the world ought to be. Cross-linguistically, these two classes of attitude verbs also differ in the properties of their clausal complements (Bolinger 1968, Farkas 1985, Giannakidou 1997, Hooper 1975, Villalta 2008). In English, this difference is reflected in the tense (finiteness) of the complement. Preferential attitude verbs like *want* tend to occur with nonfinite complements:

- (14) a. *I want Jo to be at home*
b. **I want that Jo is at home.*

By contrast, representational attitude verbs like *think* tend to occur with finite complements:

- (15) a. *I think that Jo is at home*
b. **I think Jo to be at home.*

For syntactic bootstrapping to work in the domain of attitude verbs, subclasses of attitude verbs must have different syntactic distributions. In addition, it must be that these

differences can be linked to cross-linguistically stable properties, so that it is possible for learners to link the distributional differences to those aspects of meaning that explain them. Finally, learners must be sensitive to the relevant distributional features and use them to make inferences about meaning. Each of these conditions appears to hold.

White *et al.* (2018a) tested whether subclasses of attitude verbs reliably show different distributional signatures. Building on a method pioneered by Fisher *et al.* (1991) these researchers collected two kinds of judgments. First, they collected syntactic acceptability judgments for a set of 30 attitude verbs in 19 syntactic environments, which were used to identify subclasses of verbs based on the similarity of judgments across all 19 environments. Second, they collected semantic similarity judgments in sets of 3 for all 30 verbs. Putting these together, they showed that the verb similarities identified in the semantic similarity task are highly predictive of the similarities identified in the syntactic acceptability judgment task. Together, these results indicate that attitude verbs with similar meanings show similar syntactic distributions.

In addition, these patterns of distribution relate to a principled feature of the syntax-semantics mapping. Although languages differ in the particular features that distinguish belief verbs from desire verbs, there is a higher order generalization that links these classes to the syntax. Specifically, representational attitude verbs take complements that have syntactic features of declarative main clauses (Hacquard & Lidz 2019). The specific features distinguishing declarative main clauses from other kinds of clauses vary from language to language.³ In English, these features include finite tense, lack of a

³ It is worth keeping in mind that the representational/preferential split does not exhaust the subclasses of attitude verbs. There are many further subclasses, which also display systematic

complementizer, presence of subject/verb agreement, obligatory overt subjects, among others. In German, finiteness is a less good cue, but verb-second word order is a more reliable indicator. Similarly, in French finiteness is not a good indicator of declarative main clauses, but indicative mood is. When the features of declarative main clauses occur in embedded clauses, they are good predictors of the kind of attitude verb that embeds them. Thus, the abstract link between declarative main clauses and the complement of belief verbs appears to be cross-linguistically stable, though a more thorough investigation of the range of variation, drawing from a wider sample of languages, remains to be undertaken.

The link between representational attitudes and declarative main clauses may be further explained by the pragmatic function to which declarative clauses are used. Declarative main clauses are used to make assertions and when representational attitude verbs are used to make indirect speech acts, these are indirect assertions. Recognizing the similarity between direct and indirect assertions and the formal similarity between declarative main clauses and the complements of representational attitudes may play a key role in explaining why the abstract connections between verb meaning and complement type hold (see Hacquard and Lidz 2019 for elaboration).

In order for the link between complement type and attitude verb meaning to be useful in acquisition, though, learners must be able to identify the relevant features from their input, as they acquire attitude verbs. White et al (2018b) built a computational model showing that it is possible to identify these features and consequently to learn which

relations between meaning and syntactic distribution (see Hacquard & Wellwood 2012, White et al 2018a, White & Rawlins 2019, Djarv 2019, Wurmbrand & Lohninger 2020).

attitude verbs are belief verbs and which are desire verbs. Their model identifies the morphosyntactic features of declarative main clauses. The model then looks for these features inside complement clauses to classify the embedding verbs. This model successfully figures out how to divide attitude verbs into representational and preferential subclasses in English. Crucially, it does so not by relying on specific morphosyntactic properties, but rather, via more abstract expectation about verb classes, whose expression can be discovered depending on the surface features of the language being acquired. Huang et al (2018, 2020) successfully extend this analysis to Mandarin, a language with a sparser set of morphosyntactic cues.

Finally, children are sensitive to the relevant features in their acquisition of attitude verbs. Harrigan, Hacquard and Lidz (2019) tested whether four-year-olds use the finite/non-finite complement distinction as evidence about the meaning of attitude verbs. They probed four year olds' understanding of *hope*. *Hope* is relatively uncommon in child-directed speech; it should thus be familiar enough to four-year-old children for them to know that it is an attitude verb, but perhaps not enough for them to be sure about its meaning. Hope is also relevant because it can occur with both finite and nonfinite complements, and also shows properties of both representational and preferential attitudes (Portner 1992, Scheffler 2008, Anand & Hacquard 2013). For example, while a hope is a kind of desire, one cannot hope for things that they know to be incompatible with truth (16).

(16) a. Kim knows it is snowing but wants it to be sunny.

b. *Kim knows it is snowing but hopes it is raining.

The experimental setup in Harrigan et al. made both the beliefs and desires of a puppet, Froggy, salient, and tested whether the syntactic shape of the complement influences children's interpretation of *hope*.

In this game, the child and one experimenter are behind an occluder, while Froggy is on the other side. In front of the child is a box with 40 wooden hearts and stars, which are either red or yellow. Color is predictive of shape: 15 of the hearts are red, 5 are yellow, and 15 of the stars are yellow, 5 are red. The child and the experimenter pull shapes out of the box to show Froggy, and every time the shape is a heart, the child gives Froggy a sticker. Froggy likes getting stickers, therefore his desire on every trial is that the shape be a heart. On each trial, before Froggy sees what the shape is, the child and the experimenter show him a 'clue,' which is ambiguous as to shape but not color, by inserting a point (of the heart or the star) through an opening in the occluder (see Figure 1). Thus, on every trial, Froggy has both a desire about shape (he always wants the shape to be a heart), and a belief about shape (when it is red, he always guesses that it's a heart and when it is yellow, that it's a star). Another puppet, Booboo, whom the child is told is "silly and wants to learn how to play the game, but often gets things mixed up," utters test sentences about what Froggy wants (17), thinks (18), or hopes (19-20). The child's task is to say whether Booboo is right.

(17) Froggy wants it to be a heart/star

(18) Froggy thinks that it's a heart/star

(19) Froggy hopes to get a heart/star

(20) Froggy hopes that it's a heart/star

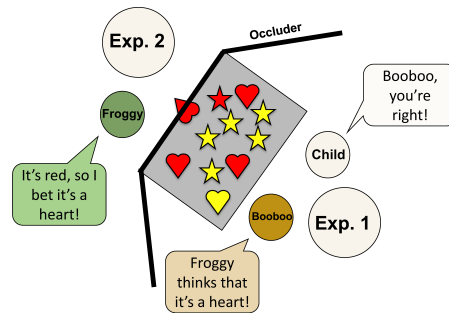


Figure 19.1: Experimental set up of Harrigan et al (2019)

The results reproduced the traditional split in performance with *think* and *want*. Children correctly judge *want* sentences even when the reported desire conflicts with reality, but they tend to incorrectly reject *think* sentences that report false beliefs. Crucially, children's responses to *hope* sentences differ depending on the syntactic frame in which they are presented. With a finite complement, their responses pattern like their responses to *think* sentences. However, with a nonfinite complement, their responses pattern like their responses to *want* sentences.

These results reveal that children use complement syntax to identify aspects of an attitude verb's meaning: four-year-olds treat *hope* as a preferential attitude verb when it takes a nonfinite complement, and a representational attitude verb when it takes a finite complement. One potential concern with this finding is that it depends on children's performance with a real verb (though one that they likely have very little experience with) and uses the false belief error as evidence of the verb's representationality.

To get a more direct demonstration of the role of syntax in learning attitude meanings, Lidz *et al.* (2017) tested four-year-olds' understanding of a *novel* attitude verb, in an experiment building on Asplin (2002). In a representative story, Dad (the DISCOVERER)

leaves chicken legs on the kitchen table and goes to work. Fido (the ACTOR) arrives hungry but can't find any food. Jimmy (the ENTICER) arrives and **ENTICES** Fido to eat the chicken legs, puts them in Fido's bowl and leaves. Fido decides to eat the chicken. When Dad returns, he finds his chicken legs gone and chicken bones in Fido's bowl, thereby **DISCOVERING** that Fido ate his chicken. The enticer has a desire (for Fido to eat the chicken), while the discoverer has a belief (that Fido ate the chicken). After each story, two puppets deliver test sentences as descriptions of the story using the same novel verb and complement syntax (finite vs. nonfinite), but with different subjects and participants are asked to choose the puppet who said something true about the story:

(21) FINITE CONDITION

Puppet 1: *Dad* gorped that Fido ate the chicken.

Puppet 2: *Jimmy* gorped that Fido ate the chicken.

(22) NONFINITE CONDITION

Puppet 1: *Dad* gorped Fido to eat the chicken

Puppet 2: *Jimmy* gorped Fido to eat the chicken.

Preferential attitudes (e.g., enticings in this context) take nonfinite complements in English; representational attitudes (e.g., discoveries in this context) take finite complements in English. If children are sensitive to this mapping, they should pick the discoverer when they hear a sentence with a finite complement, but the enticer when they hear a sentence with a nonfinite complement. This is what we find. In the finite condition, children were significantly more likely to pick the puppet that mentioned the discoverer, but in the nonfinite condition, they were significantly more likely to pick the one that mentioned the enticer.

In this section, we have seen that different classes of attitude verbs show different distributional profiles, that these distributional profiles are related to a cross-linguistically stable, and hence principled, mapping between syntax and semantics, and that children are sensitive to the relevant features of syntax in acquiring attitude verbs. More generally, we've seen that syntactic structure and syntactic distribution play two important roles in the acquisition of verb meanings. First, the syntactic structure of a sentence can help a learner to identify a perspective on an event, allowing them to fix the referential intentions of a speaker and to identify the meaning of an unknown verb in that sentence. Second, the syntactic distribution of a word is related to its meaning in principled ways, allowing learners to identify semantic properties of a word's meaning by observing relevant syntactic features.

6. Syntactic bootstrapping beyond verbs

While syntactic bootstrapping effects have been predominantly associated with verb learning, syntactic environment is a cue to word meaning across several classes. The syntactic category of a word in many cases is sufficient to restrict its possible meaning. For example, Waxman & Booth (2006) show that infants as young as 14-months treat a novel word presented as an adjective as referring to an object property but treat a novel noun as referring to an object kind. They presented infants with four objects that were alike in both their kind (horses) and an accidental property (purple). One group of infants was introduced to each of these objects with a novel noun (this is a blicket). Another group was introduced to them with a novel adjective (this is a blickish one). They were then given a choice either between two objects that shared a kind, but differed in color (yellow horse vs. purple horse) or between two objects that shared a

color but differed in kind (purple plate vs. purple horse). They found that 14-month-olds treated nouns as a label for object kind. When asked to find “another blicket”, they chose the purple horse over the purple plate, but showed no preference in the case of two horses. And, they found a different pattern with adjectives. Infants who were asked to find “another blickish one” chose the purple horse over the yellow horse, and the purple horse over the the purple plate.

Similarly, Fisher et al (2006) show that 26-month-old children treat a novel preposition as labeling a spatial configuration, but they treat a novel noun in the same context as labeling an object (cf. Landau & Stecker 1990). Children in this study were shown a duck on top of a box and heard either “this is a corp” (noun condition) or “this is a corp my box” (preposition condition). They were then shown either a new duck on top of the box or a new duck and a pair of glasses on top of the box and were asked “what else is a corp (my box)?” Children in the noun condition looked more at the new duck by itself whereas those in the preposition condition looked more at the scene with the duck and glasses. This suggests that they learned that the novel noun labeled the duck and that the novel preposition labeled the spatial configuration.

Finally, Wellwood, Gagliardi & 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. Here, participants were shown several scenes in which there were two sets of cows, one set by the barn and one set in the meadow. The cows by the barn were always more numerous than the cows in the meadow. And, the cows by the barn had more spots on them than the cows in the meadow. As they saw these scenes, children were told either

that a picky puppet liked it better when “gleebest of the cows are by barn” (determiner condition) or that he liked it better when “the gleebest cows are by the barn.” (adjective condition) Then, they were asked about scenes in which the numerosity of cows and the numerosity of spots were dissociated, so that the cows by the barn were more numerous but less spotted, or less numerous and more spotted. The children in the determiner condition learned that the puppet likes it when the larger set of cows is by the barn, whereas those in the adjective condition learned that the puppet likes it when more spotted (but less numerous) cows are by the barn.⁴ Again, syntactic category has a powerful influence on the kinds of meanings that learners assign to novel words.

In each of these cases, patterns of extension for a novel word were determined by the syntactic category of the novel word, an illustration of how syntactic information can shape the perspective that a learner takes on a scene in learning a novel word. This effect of syntax has been studied most widely in the domain of verb learning, but the potential for learners to exploit information about syntax is present in any area of grammar where there are systematic relations between a word’s syntactic (sub)category and its meaning.

7. Origins of syntactic bootstrapping

While the primary purpose of this chapter is to show that syntactic distribution provides information about a novel word’s meaning and that children are sensitive to this information, it is important to also ask about the origins of these abilities. There are two

⁴ This effect is not likely to be an analogy to *most*. The earliest reported success with *most* is at age 3;6 (Halberda, Taing & Lidz 2008) and many studies do not reveal knowledge of *most* until much later (Barner et al 2009, Papafragou & Schwarz 2005, Sullivan et al 2019). Moreover, since *most* is grammatical in both frames, the different meanings assigned in the different frames cannot be achieved by drawing a simple analogy to the distribution of *most*.

important questions to ask in this domain. First, to what extent do children's abilities to use syntactic features as evidence for word meanings reflect architectural properties of the language faculty as opposed to learned generalizations? Second, given that children are not born knowing the syntax of their language, how do they acquire enough syntax to allow these kinds of bootstrapping effects to manifest? We take up these questions briefly, in turn.

Let us first consider the possibility that children learn some words in a distributional category via some non-syntactic method and then treat other words in that category as being semantically similar to those. On this view, studies showing that children use syntax as a cue to meaning reveal statistical generalizations that are acquired only after the meanings and syntactic properties of some small subset of words in that category are acquired. There are several reasons to be skeptical towards such an account.

First, Lidz, Gleitman & Gleitman (2003) examined learners of Kannada in a task where known verbs were put into syntactic environments that do not occur in speech to children. They found that learners extended the meanings of such verbs based on features related to grammatical architecture rather than to more reliable statistical cues in the language. This suggests that the role that syntax plays in cueing meaning does not simply derive from statistical features of the environment, but rather from properties internal to the child (see also Trueswell et al 2012, Pozzan & Trueswell 2016).

Second, recall the effect of syntactic structure in cueing thematic relations in Lidz, White & Baier (2017). 16-month-olds who heard a novel NP in the direct object position interpreted that NP as referring to the patient of the event, whereas those who heard the novel NP as a prepositional object interpreted it as referring to the instrument of the

event. These authors found that the effect of syntactic position is negatively related to children's knowledge of specific verbs. As verb vocabulary goes up, children's ability to use the syntactic position as evidence of the thematic relation goes down. Lidz et al. argue that this effect results from the lexical statistics of the verb making an independent contribution to parsing and understanding. Because the verbs used in this study occurred in speech to children with direct objects roughly 80% of the time, this statistical information overpowered the bottom-up information derived from the syntax. In the current context, what this means is that the effect of syntax cannot be derived from the lexical statistics, since the children who know the lexical statistics use this information at the expense of syntactic information in guiding their interpretations.

Finally, a theory in which syntactic cueing of meaning begins with first acquiring some meanings via some method *other* than syntactic bootstrapping, presumably by perceiving the meaning directly. However, as reviewed in the beginning of this chapter, sentence meanings are exceedingly difficult to observe directly. This difficulty derives from the fact that sentences are not mere descriptions of the events around us and the fact that perspectives on events are often not shared among people, especially caregivers and children, in a given situation. Support for this difficulty comes from the human simulation studies, finding that both adults and children are extraordinarily bad at guessing what someone is likely to say based only on extralinguistic information (Gillette et al 1999, Piccin & Waxman 2007).

The alternative view is that some syntactic bootstrapping effects reflect architectural features of the language faculty. Bootstrapping effects are about learners taking

advantage of relations between representations, in this case relations between syntactic distribution and meaning.

In the domain of action verbs, the relevant features are about the relations between syntactic positions and thematic relations. A child who can use the syntactic position of arguments to make inferences about their likely thematic relations can then use the thematic relations of those arguments to identify the event described by a sentence. And knowing the event that the sentence describes goes a long way towards identifying the meaning of the verb in that sentence.

In the domain of attitude verbs, learners would need the following expectations: (a) that some verbs embed clauses and that when they do, they report a relation between the subject and some state of affairs described by the complement, (b) that certain clause types are strongly associated with certain speech acts (e.g., declarative clauses with assertions), (c) that speech acts can be indirect, and (d) that the type of complement an attitude takes is predictive of its meaning in matching the clause type of the canonical speech act that the verb lends itself to. Armed with these expectations, a learner would need to identify the surface hallmarks of various clause types in the particular language being acquired. Having done so, this information, in combination with the expectations (a-d) would allow them to classify novel attitude verbs as being either representational or preferential.

It is worth noting, however, that while these prior expectations about the mapping between syntax and semantics allow learners to make rough initial classifications, there is nonetheless a broad range of verb meanings that will not be acquirable on the basis of this information. Finer subcategories of both action verbs (Levin 1993) and attitude

verbs (White & Rawlins 2016) show systematic relations between form and meaning, but at least some of these relations appear to be idiosyncratic and language particular. For such cases, it is likely that learners' initial expectations play less of a role in guiding acquisition. Instead, we can envision a learning model more along the lines of how gender classes are acquired (Gagliardi & Lidz, 2014), where semantic features are only probabilistically associated with syntactic features. On such a view, the semantic features would have to be acquirable based on the initial classification of the verb in concert with the rich understanding of linguistic and extralinguistic context that comes with being an experienced language user. In other words, the fact that there are biases concerning the syntax-semantics mapping that provide a foothold for children's first steps into meaning does not preclude the possibility that later acquisitions can be driven in part by previously acquired knowledge.

The second issue concerning the origins of syntactic bootstrapping effects concerns the initial acquisition of syntax. If children need syntactic information to guide the acquisition of word meaning, we are forced to the question of how children learn the syntax to begin with. One possibility that has been considered widely is that meaning can provide a foothold into the acquisition of syntax (Pinker 1984, 1989). But here we run into potential problems of circularity – if you need semantics to acquire syntax and syntax to acquire semantics, how can you ever get started?

There are two classes of proposals for dealing with the initial steps into syntax, both of which help to avoid the potential circularity, and which are not mutually exclusive. One class of proposals holds that certain words can be acquired without the help of syntax and that these then function as anchors for subsequent category building and for

the discovery of syntax (Gleitman & Trueswell 2020, Gutman et al 2014, Fisher et al 2020). A second class of proposals suggests that prosodic cues to syntactic structure, in concert with the identification of functional vocabulary at the edges of prosodic constituents, allow learners to build an initial syntactic skeleton containing the information needed for identifying core syntactic categories and the subject-predicate divide (Christophe et al 2008, de Carvalho et al 2019). This information would provide sufficient information to drive further acquisition of syntax and to guide early syntactic bootstrapping.

8. Conclusions

We began this chapter noting that the world of perception and the world of language are not always in alignment. This misalignment has many sources, ranging from the diverse perspectives we can take on events, through to the wide range of conversational goals that we use language to achieve. In the worst case, utterances may achieve the same conversational goals despite having very different forms. A parent sending his children to bed might just as well say any of the following sentences:

- (23)
- a. Go to bed
 - b. It's time for bed
 - c. I want you to go to bed
 - d. I think it's bedtime

Moreover, these same sentences might be used to achieve different conversational goals. If, for example, my friends invite me out for a drink after a long day at work, I might use (23d) to decline the invitation. One might think that the gulf between what we say, what we mean and what is happening around us might make language learning

next to impossible. But, as we've seen in this chapter, there are systematic relations between a word's syntactic distribution and its meaning that can help bridge the gap. These systematic relations are principled, allowing them to serve as an inductive base to guide learning. They are reliably expressed in speech to children, allowing them to be detected. And, children appear to be quite capable of making the inferences from syntactic form to lexical meaning. These inferences play a foundational role in word learning across the lexicon and guide children's earliest steps into word learning.

References

- Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, 114(3), 442–446.
- Arunachalam, S. & S. Dennis (2018) Semantic detail in the developing verb lexicon: An extension of Naigles & Kako (1993). *Developmental Science*.
- Asplin, K. (2002). *Can complement frames help children learn the meaning of abstract verbs?* Ph.D. Dissertation, University of Massachusetts, Amherst.
- Astington, J. W., & Gopnik, A. (1991). Theoretical explanations of children's understanding of the mind. *British Journal of Developmental Psychology*, 9(1), 7–31.
- Austin, J.L. (1962) *How to do things with words*. Oxford University Press.
- Baker, M. C. (1988). *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.
- 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: The MIT Press.
- Brandone, A., Addy, D. A., Pulverman, R., Golinkoff, R. M., & Hirsh-Pasek, K. (2006). One-for-one and two-for-two: anticipating parallel structure between events and language. In *Proceedings of the 30th Annual Boston University Conference on Language Development*. Cascadilla Press Somerville, MA.
- Bunger, A., & Lidz, J. (2004). Syntactic bootstrapping and the internal structure of causative events. In *Proceedings of the 28th Annual Boston University Conference on Language Development*. Cascadilla Press.
- Bunger, A., & Lidz, J. (2008). Thematic relations as a cue to verb class: 2-year-olds distinguish unaccusatives from unergatives. *University of Pennsylvania Working Papers in Linguistics*, 14(1), 4.
- Carey, S. (2009) *Origins of Concepts*. Oxford University Press.
- Chomsky, N. (1959). Review of Verbal Behavior by B.F. Skinner. *Language*, 35(1), 26-58.
- Chomsky, N. 1971. *Problems of Knowledge and Freedom*. Pantheon: NY.
- Clark, E. V. & Amaral, P. M. (2010). Children build on pragmatic information in language acquisition. *Language & Linguistics Compass*, 4(7), 445–457.
- Davis, H. (1997). Deep unaccusativity and zero syntax in St'a't'imcets. *Anuario del Seminario de Filologia Vasca "Julio de Urquijo"*, 55–96.
- Davis, H. (2010). *A Teaching Grammar of Upper St'at'imcets*. ms., University of British Columbia and The Upper St'a't'imc Language Culture and Education Society.
- Davis, H. and H. Demirdache 2000. On Lexical Verb Meanings: Evidence from Salish.

- In J. Pustojevsky and C. Tenny (eds.), *Events as Grammatical Objects: the Converging Perspectives of Lexical Semantics and Syntax*. CSLI: Stanford University Press, 97-142.
- Diessel, H. and Tomasello, M. (2001). The acquisition of finite complement clauses in english: A usage based approach to the development of grammatical constructions. *Cognitive Linguistics*, 12:97–141.
- Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 547– 619.
- Dudley, R. (2017) *The role of input in discovering presupposition triggers: figuring out what everybody already knew*. Ph.D. dissertation, University of Maryland.
- Fillmore, C. 1970. The grammar of hitting and breaking. in R. Jacobs & P. Rosenbaum (eds). *Readings in Transformational Grammar*, 120-133. Ginn: Waltham, MA.
- Fisher, C., Gertner, Y., Scott, R. M., & Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(2), 143–149.
- Fisher, C. H., Gleitman & Gleitman, L. R. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology*, 23, 331–392.
- Fisher, C., D.G. Hall, S. Rakowitz, L. Gleitman (1994) When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. *Lingua*, 333-375.
- Fisher, C., S.L. Klingler & H.J. Song (2006) What does syntax say about space? 2-year-olds use sentence structure to learn new prepositions. *Cognition*, 101, B19-29.
- Flavell, J. H., Green, F. L., & Flavell, E. R. (1990). Developmental changes in young children's knowledge about the mind. *Cognitive Development*, 5(1), 1–27.
- Fodor, J. (1975) *The language of thought*. Harvard University Press.

- Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: abstract knowledge of word order in early sentence comprehension. *Psychological Science*, 17(8), 684–691.
- Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, 73, 135–176.
- Gleitman, L.R. (1990) The structural sources of verb learning. *Language Acquisition*, 1, 3-55.
- Gleitman, L., Cassidy, K., Nappa, R., Papafragou, A. & Trueswell, J. (2005). Hard words. *Language Learning and Development*, 1(1), 23–64.
- 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.
- Gleitman, L.R. & J.C. Trueswell (2020) The easy words: Referent resolution in a malevolent referent world. *Topics in Cognitive Science*.
- Goodman, N. (1955). *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard University Press.
- Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: domain specificity in cognition and culture* (pp. 257–293). New York, NY: Cambridge University Press.
- Hacquard, V. & J. Lidz (2019) Children's attitude problems: Bootstrapping verb meaning from syntax and pragmatics. *Mind and Language*, 34, 73-96.
- Harrigan, K., V. Hacquard & J. Lidz (2019) Hope for syntactic bootstrapping. *Language*, 642-682.

- Hansen, M. B. (2010). If you know something, say something: young children's problem with false beliefs. *Frontiers in Psychology*, 1, 23.
- He, Z., Bolz, M., & Baillargeon, R. (2012). 2.5-year-olds succeed at a verbal anticipatory-looking false-belief task. *British Journal of Developmental Psychology*, 30(1), 14–29.
- Helming, K. A., Strickland, B., & Jacob, P. (2014). Making sense of early false-belief understanding. *Trends in Cognitive Sciences*, 18(4), 167–170.
- Huang, N., C.-H. Liao, V. Hacquard & J. Lidz (2018). Learning attitude verb meanings in a morphosyntactically poor language via syntactic bootstrapping. In A. Bertoni & M. Kaplan, eds. *Proceedings of the 42nd Boston University Conference on Language Development*. Somerville, Mass.: Cascadia Press.
- Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge: MIT Press.
- Landau, B. & Gleitman, L. (1985). *Language and Experience. Evidence from the blind child*. Cambridge, MA: Harvard University Press.
- Landau, B. & D. Stecker (1990) Objects and places: Geometric and syntactic representations in early lexical learning. *Cognitive Development*, 5, 287-312.
- Levin, B. (1993) *English verb classes and alternations*. University of Chicago Press.
- Levin, B., & Hovav, M. R. (2005). *Argument realization*. Cambridge: Cambridge University Press.
- Lewis, S., Hacquard, V., & Lidz, J. (2017). “Think” pragmatically: children's interpretation of belief reports. *Language Learning and Development*, 1–23.
- Lidz, J., Dudley, R. & Hacquard, V. (2017). Children use syntax of complements to

- determine meaning of novel attitude verbs. *Paper presented at BUCLD 41.*
- Lidz, J., Gleitman, H. & Gleitman, L.R. (2003). Understanding how input matters: verb learning and the footprint of universal grammar. *Cognition*, 87, 151-178.
- 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.
- Murray, S.E. & W.B. Starr (2020) The structure of communicative acts. *Linguistics and Philosophy*. Article in press.
- 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.
- Nappa, R., Wessel, A. McEldoon, K.L., Gleitman, L.R. & Trueswell, J.C. (2009). Use of speaker's gaze and syntax in verb learning. *Language Learning and Development*, 5, 203-234.
- 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.
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308(5719), 255–258.
- Papafragou, A., Cassidy, K. & Gleitman, L. (2007). When we think about thinking: The acquisition of belief verbs. *Cognition*, 105, 125–165.

- Perkins, L. (2019) *How grammars grow: Argument structure and the acquisition of non-basic syntax*. Ph.D. Dissertation, University of Maryland.
- Perner, J. (1991). *Understanding the Representational Mind*. Cambridge, MA: MIT Press.
- Pinker, S. (1984). *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.
- Pozzan, L., Gleitman, L. R., & Trueswell, J. C. (2015). Semantic ambiguity and syntactic bootstrapping: The case of conjoined-subject intransitive sentences. *Language Learning and Development*.
- Quine, W.V.O (1960) *Word and object*. MIT Press.
- Rubio-Fernández, P., & Geurts, B. (2013). How to pass the false-belief task before your fourth birthday. *Psychological Science*, 24(1), 27–33.
- Scott, R. M., & Fisher, C. (2009). Two-year-olds use distributional cues to interpret transitivity-alternating verbs. *Language and cognitive processes*, 24(6), 777–803.
- Searle, J. (1969) *Speech Acts*. Cambridge University 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.
- Snedeker, J., & Gleitman, L. R. (2004). Why It Is Hard to Label Our Concepts. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (p. 257–293). MIT Press.
- Spelke, E. S. (1990). Principles of object perception. *Cognitive Science*, 14(1), 29–56.

- 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 Kuhn, D. & Siegler, R. (Eds). *Handbook of Child Psychology: Cognition, Perception and Language*. New York, NY: John Wiley & Sons, 299–335.
- 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. (2018a). Semantic information and the syntax of propositional attitude verbs. *Cognitive Science*, 42(2), 416–456.
- White, A. S., Hacquard, V. & Lidz, J. (2018b). Main clause syntax and the labeling problem in syntactic bootstrapping. In K. Syrett & S. Arunachalam (Eds.), *Semantics in acquisition*. TiLAR. Amsterdam: John Benjamins.
- Williams, A. (2015) *Arguments in syntax and semantics*. 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