Advanced Review

Word learning mechanisms

Angela Xiaoxue He* and Sudha Arunachalam

How do children acquire the meanings of words? Many word learning mechanisms have been proposed to guide learners through this challenging task. Despite the availability of rich information in the learner's linguistic and extralinguistic input, the word-learning task is insurmountable without such mechanisms for filtering through and utilizing that information. Different kinds of words, such as nouns denoting object concepts and verbs denoting event concepts, require to some extent different kinds of information and, therefore, access to different kinds of mechanisms. We review some of these mechanisms to examine the relationship between the input that is available to learners and learners' intake of that input—that is, the organized, interpreted, and stored representations they form. We discuss how learners segment individual words from the speech stream and identify their grammatical categories, how they identify the concepts denoted by these words, and how they refine their initial representations of word meanings.

How to cite this article:
WIREs Cogn Sci 2017, 8:e1435. doi: 10.1002/wcs.1435

INTRODUCTION

A child's first word is a celebrated milestone. Though the process of language acquisition begins long before this milestone, and even before birth, it overtly signals the child's entrance into her social milieu—in which language is the primary communicative system—and marks the beginning of a rapid addition of new words into her productive vocabulary; by age 18 months, children produce on average 50 words and comprehend over 200.[1]

Where do these words come from? Naturally, children acquire words of the language(s) to which they are exposed. And the frequencies with which children use specific words, as well as the ways in which they use them, are related to how and how often they are used in their environment.[2] But children's language competence does not consist of simple memorization and repetition of the incoming linguistic stream. To acquire an individual word, children segment its phonological form from the undemarcated auditory stream, map the form to a candidate conceptual referent, and continually refine until the extension of the word's meaning matches the usage of other members of the linguistic community. These tasks are not trivial. For example, the situation in which a word is uttered offers in principle an unbounded hypothesis space about the concept it denotes, as famously noted by Quine—saying 'gavagai' in synchrony with a rabbit running yields an indeterminate space of possible meanings (Does it mean *rabbit*? *Run*? *Tail*?).[3]

Thus, though the availability of language input is critical for children to acquire the language of their community, exposure to input alone is insufficient. Importantly, properties of the learning situation and of the learner together determine what precisely the learner absorbs from it, that is, the learner's 'intake.' Learning situations may differ in the degree to which they unambiguously distinguish competing hypotheses about word meaning. And learners may differ in their abilities to discover and utilize the cues available in the input. Therefore, in constructing a model of lexical development, the learner's 'input' must be considered in tandem with the concept of 'intake.'[4]

In this review we discuss some of the word learning mechanisms that have been hypothesized to guide learners as they acquire words. We limit the scope of our discussion to the acquisition of open class words, particularly focusing on nouns and verbs

*Correspondence to: angelahe.axh@gmail.com

Department of Speech, Language, and Hearing Sciences, Boston University, Boston, MA, USA

Conflict of interest: The authors have declared no conflicts of interest for this article.

because these constitute the majority of children's early receptive and expressive vocabularies and are the best understood with respect to learning mechanisms. In what follows, we first address the problem of *finding the words*, which include segmenting words from the continuous speech stream and identifying their grammatical categories; then, we turn to the issue of *identifying the concept*, discussing what mechanisms the learner makes use of in assigning meanings to words; and lastly, we discuss *learning over multiple situations*, through which the learner refines her initial hypothesis with additional exposure.

FINDING THE WORDS

Initial steps in forming a representation for a new word are segmenting individual word forms from the continuous speech stream and identifying their grammatical categories (e.g., noun, verb, adjective). Although language does not come to the child in neatly separated lexical units, the linguistic stream includes statistical regularities that help learners identify word boundaries. For example, the transitional probabilities between syllables—that is, the probability with which a particular syllable occurs given a preceding one—are generally higher within words, and lower across word boundaries. By eight months of age, infants can track such regularities and use them to identify word boundaries.[5] This can be an effective strategy, because in natural languages, syllables that frequently occur together do often make up words. But of course, this strategy may occasionally lead to faulty word recognition: Ngon et al. found that 11-month-olds acquiring French grouped disyllabic sequences based on their frequency of occurring together even when these sequences did not form words.[6] Learners must be able to reparse the speech stream and reestablish lexical form representations in order to recover from such errors.

Graf Estes et al. followed the teaching phase of a statistical segmentation study with a switch task for word learning in which word forms are paired with novel object referents,[7] finding that although 17-month-olds could map these newly-segmented forms onto meanings, they only did so for sound sequences that had strong internal structures (i.e., high transitional probabilities between syllables), but not those with equal frequencies but weak internal structures. Shukla et al. found that this ability is even present in 6-month-olds, although these infants could not map all statistically defined words onto referent objects, but rather only those that were aligned with prosodic constituency boundaries.[8] These findings suggest that input shapes learning, but not through mere frequency or familiarity; rather, it is the learner's intake from the input—that is, structured representations of sound sequences—that feeds the next learning task.

There may also be limitations in learners' abilities to use regularities in the input because of learner characteristics such as memory and processing abilities. For instance, Johnson and Tyler showed that infants were not able to identify word boundaries in a more complex situation than those studied previously, in which words varied in syllable number.[9] And a recent study using Bayesian modeling showed that learners more constrained in memory and processing abilities were consistently outperformed by ideal learners.[10] Taken together, these studies point to a discrepancy between the information available about word boundaries in the input and infants' abilities to use this information effectively—which may be limited by their developing memory and processing skills as well as their abilities to revise initially incorrect segmentations.

Distributional cues in the input can also support young learners' abilities to assign a phonological form to its grammatical category. Although languages differ in the precise types of cues they offer, learners are able to use the distribution of novel phonological forms with respect to free and bound function morphemes to perform this assignment. For example, 12-month-old English infants can use a preceding auxiliary or infinitival marker (e.g., can *gorp,* to *gorp*) to categorize novel words as verbs[11]; 14-month-old French infants and 12-to-16-month-old German infants can use determiners to categorize novel nouns[12,13]; and by 18 months, infants can recognize familiar words only when they are used in their expected nominal or verbal distributions (e.g., *they eat* as opposed to **an* eat).[14] Phrasal prosody, too, can support categorization. By 18 months of age, infants infer whether a novel word is a noun or verb, based on phrasal prosody (e.g., [la petite *bamoule*] (the small *bamoule*) vs. '[la petite] [*bamoule*]' (the small (one) *bamoules*), with '[]' indicating phrasal prosodic boundaries).[15] Once a novel word's grammatical category is identified, the learner can narrow down their hypotheses about its meaning by utilizing relationships between grammatical and conceptual categories—for example, nouns typically name object kinds and verbs name event categories.[15–22]

IDENTIFYING THE CONCEPT

Armed with the segmented form of a new word, the learner must identify the specific concept it denotes.

This is no trivial task. Naturalistic learning situations will typically be cluttered environments yielding a large hypothesis space from which learners must identify the target concept.[23] Good learning situations may contain rich cues to help narrow this hypothesis space, including the observational context (e.g., what entities and actions are observable in the scene), the social context (e.g., what the speaker's communicative intent appears to be), and the linguistic context (e.g., syntactic structures and familiar words with which the word appears). But can learners make use of such contextual information? Fortunately, learners are well equipped—with cognitive biases that lead them to entertain certain hypotheses over others, with social-pragmatic abilities that help them identify a speaker's intended meaning, and with language parsing abilities that allow them access to the word's surrounding linguistic context. These strategies allow learners to transform the unstructured and unorganized input into systematized intake representations, eventually guiding them out of the word meaning jungle.

Conceptual Knowledge and Cognitive Biases

Infants come to the task of word learning with a rich store of representational means (e.g., perceptual, sensory-motor, conceptual) that enables them to glean information from the environment in a structured way, rather than experiencing it as a 'blooming, buzzing confusion.'[24] Infants view objects as bounded and enduring over time and space[25,26]; and they distinguish objects based on features like animacy, solidity, and numerosity.[27–29] They also encode individualized events that instantiate relations such as causation, figure and ground, manner and path of motion, and probable endstates.[30–34] In recent work, 13-month-olds perceived tools as essential to bringing about action endstates, but only when the endstates could be interpreted as the goals of the actions.[35] Infants are thus able not only to represent such individual features of entities and events but to encode relationships between them.

On one hand, these representational means ensure that infants have access to the conceptual possibilities to which they can map a new word. On the other hand, the richness of these means may also expand the hypothesis space, given that, as Gleitman pointed out, 'an observer who notices everything can learn nothing.'[36] Fortunately, infants also appear to have access to biases or heuristics to constrain their initial hypothesis space when they encounter a new word and seek to identify its referent. These biases include the *whole object assumption*—the assumption that a word labels a whole object rather than part of an object[37,38]; the *shape bias*—the bias to generalize a word to other objects with the same shape[39]; the *taxonomic assumption*,[40] or *noun-category bias,* under which a word is hypothesized to extend to other members of the same category (e.g., 'dog' refers to the family dog and the neighbor's dog, rather than the family dog and thematically related objects like bones)[41]; and *mutual exclusivity*, under which children assign a new word to a referent for which they do not already have a basic-level label.[38]

Of these, mutual exclusivity has been the best studied, earning many additional names and descriptions (e.g., Principle of Contrast[42]; Novel Name-Nameless Category principle[43]). These different proposals have important theoretical differences—for example, they differ in whether the driving force is essentially social-pragmatic in nature,[42,44] such as an assumption that speakers will use conventional terms and that if they introduce a new term it must have a different referent, or whether it is essentially a novelty preference.[45] Halberda has suggested that learners engage in syllogistic reasoning to reject the familiar referent and instead select the novel referent, a proposal which is compatible with some but not others of the existing proposals.[46] Numerous experimental conditions manipulating how novel or familiar the candidate referents are to the child or to the speaker, and what kind of evidence the speaker provides about the intended referent, have revealed that children are sensitive to a variety of cues both to support a mutual exclusivity assumption as well as to override it when necessary, that is, when two labels are in fact assigned to the same object referent.[47] Thus, young learners' strategies not only include learning heuristics, but also conditions under which the heuristics are deployed flexibly.

The extent to which these constraints or biases are innate or learned, domain-specific or domain-general, has been hotly debated.[28,48–50] For example, some have argued that the shape bias is likely to be learned rather than innate,[51,52] by documenting that by the time learners have demonstrated this bias at 2 years of age, they have already learned words for objects that come from categories that are well defined by shape. While specific strategies like the shape bias may be learned, it may still, of course, be the case that innate knowledge systems ('core knowledge') underlie the acquisition of these other skills.[53] Importantly, regardless of their origin, these lexical constraints must be considered violable, if default, assumptions that can be overridden to permit children to acquire words such as homophones and labels for parts of objects.[54] Indeed, to acquire homophones, 5-year-olds use information about the sampling of exemplars and separately cluster the two meanings of a homophone in their conceptual space.[55] With these cognitive and conceptual strategies, child learners organize the otherwise unstructured observational input into semi-structured mappings between forms and meanings—these representations constitute their intake of the observational input.

Social-pragmatic Skills

Children are exquisitely sensitive to the social nature of language use. They expect language to be used communicatively—and in turn, expect that communicative signals are linguistic: In a recent study, 6-month-olds hearing beeps emitted from the mouths of human actors who were ostensibly engaged in conversation subsequently interpreted beeps as linguistic signals in an object categorization task, suggesting that they inferred from the conversational setting that even beeps can be a kind of language-like communicative signal.[56]

Such early social communicative sensitivity is likely to generate a set of extremely useful strategies for narrowing down the hypothesis space for word meaning. A speaker's direction of gaze or pointing, for example, are often indicative of the referents of the words they use. Infants are able to use speaker's gaze to determine the object a speaker likely intends to label,[57] though they appear to be sensitive to the communicative intent in doing so— 6-month-olds only follow an adult's direction of gaze when the adult signals an intention to communicate,[58] and 18-month-olds follow a robot's gaze but do not expect it to be indicative of meaning.[59]

Young learners are also sensitive to other manifestations of a speaker's communicative intent. In an early demonstration of inference of communicative intent in word learning, 2-year-olds mapped novel nouns to objects that the actor was searching for, but not to those that were rejected during searching, and mapped new verbs only onto intentional but not accidental events.[60] Two-year-olds are also able to use an adult's affect such as surprise or excitement to infer that a novel word is being used to label an object that the adult has not seen before.[61] By the middle of the second year, children can use actors' gaze and/or gesture to distinguish similar-looking events that are distinguished by intention (e.g., tiptoe vs. tiptoe to pursue).[62] By preschool age, child learners develop more sophisticated social-pragmatic skills, taking into consideration a speaker's knowledge state and reliability when learning new words from them.[63,64]

Indeed, in some cases social interaction appears to be *crucial* for successful learning; recent studies on children's word learning from media, in particular, have suggested that media is only a successful teacher in social contexts—such as when it includes reciprocal social interaction or when parent scaffolding is also available.[65–67] Other studies, however, have documented children's abilities to learn word meanings in markedly less social contexts, such as overhearing others' conversations, whether live or videotaped, or hearing ambient linguistic streams.[7,68–70] The nature and difficulty of the learning task may determine what kind of scaffolding is necessary, again highlighting the fact that input may be available to the learner in the form of speech that is not overtly directed to them, but the learner may not always be able to utilize this input—it may not ultimately be part of the learner's intake.

Linguistic Cues

Social-communicative cues may be particularly useful in highly transparent learning situations, in which the correct hypothesis about word meaning is much more salient than other hypotheses. For example, if the word 'cup' is said when the speaker holds and gazes at a cup, the learner is likely to guess the word's meaning correctly. But when new words are offered in learning situations of lower transparency (e.g., when the observational context is cluttered, such as on a playground), social-communicative cues may fall short. And in many cases, such cues are absent—Iverson et al. found that only 15% of parents' utterances were accompanied by referential gestures such as pointing.[71] In such situations, tapping into an unfamiliar word's linguistic context can be a particularly powerful cue to its meaning, as linguistic context can point the learner to hypotheses about meaning that are not identifiable on the basis of pure salience or attentional focus.

Gillette et al. provided support for this notion using what they called the 'Human Simulation Paradigm' (so named to humorously parallel the then exploding literature on computer simulations). Adult participants watched muted videos of parent-child interactions and were asked to guess what word the parent uttered to the child at various points in the video, based on observational and social context (e.g., the parent's pointing or gesturing to an object).[23] Participants performed poorly; from a single such instance, they guessed noun referents correctly less than one-third of the time, and when the word the parent had uttered happened to be a verb, they guessed the word less than 10% of the time. But when also given access to information about the linguistic contexts in which the words appeared, participants' performed much better.

Although linguistic context can assist with noun learning (by helping the learner determine whether a novel noun's referent is animate or inanimate, among other things[72]) its role is best established with respect to verb acquisition.[36,73] Experimental research has documented young children's ability, known as *syntactic bootstrapping*, to use linguistic context to infer a novel verb's meaning. For example, 2-year-olds use the number of a novel verb's arguments to assign it a causative or noncausative meaning,[74,75] and 4-year-olds use the presence of sentential complements to infer that a verb denotes a mental state.[76,77] By 2.5 years, children are also able to use the *set* of syntactic structures in which a verb occurs to more precisely determine its meaning.[78,79] Importantly, maternal speech to young children contains the necessary information to support syntactic bootstrapping—that is, causative verbs often appear with direct objects,[80] and mental state verbs with sentential complements, and such patterns are robust across languages.[81,82]

Perhaps paradoxically, in order to take advantage of linguistic context in syntactic bootstrapping, children must already have a fair amount of lexical and grammatical knowledge in place, as well as practiced processing skills.[83–85] How do children develop these requisite abilities even as they use these very abilities to build up their lexical knowledge? Children seem to be able to form a rough parse of the linguistic context, based on their knowledge of function morphemes, prosodic cues, and minimal semantic knowledge.

For example, learners under 2 years of age can count the number of nouns in a linguistic context—taking sentences with two nouns to name two-participant events, with the first-mentioned noun naming the agent and the second-mentioned the patient.[86] This rough heuristic is an effective strategy in general, but may also lead to predictable errors. For example, 21-month-olds misinterpreted the novel verb in 'the boy and the girl are *gorp*ing' to mean an event where the boy is doing something to the girl.[87] Similarly, French-learning 28-month-olds disregarded the prosodic cue in right-dislocated sentences like 'il *dase*, the bébé (he is *das*ing, the baby; meaning 'the baby is *das*ing'), and misinterpreted the intransitive verb as transitive.[88] And recent work demonstrates that adults and infants conceptualize event participants such as instruments of actions as event participants, which should lead learners using a counting-the-nouns heuristic to the erroneous hypothesis that instruments *must* be encoded in a sentence containing the verb, though in fact they are optional (e.g., Emeril cut the tomato (with a knife)).[89] Therefore, the degree to which learners can make use of syntactic bootstrapping depends on their developing language skills as well as their conceptualization of the referent event.

In recent work we have argued that supportive contexts for verb learning must have both low processing demands and high information content. For example, Arunachalam and Waxman found that 2-year-olds learning English acquired a novel verb's meaning if it was flanked by content nouns (e.g., 'The boy is gonna pilk the balloon') but not if it was flanked by less informative pronouns (e.g., 'He is gonna pilk it')—despite that the visual scene consisted only of a boy and a balloon and thus had no other salient interpretations.[90] But modified subjects (e.g., 'the nice tall boy') pose too high of a processing demand for 2- and 3-year-olds.[85] For younger children, at age 22 months, even a single content noun may pose too great a processing demand; He and Lidz found that in a simpler learning situation children performed better with 'it' or 'that thing' in subject position of a novel verb than the more informative but more difficult-to-process 'the balloon.'[91] For older children, age 5 years, pronouns are sufficient, but contexts with no overt arguments (e.g., 'Pilking!') are not informative enough.[92] Similarly, in a different paradigm, pilot studies suggest that novel nouns replacing the content nouns (e.g., 'The dax is gonna pilk the blick') may not support 2-year-olds' acquisition of a novel verb,[93] but may suffice for 3.5-year-olds.[94] These studies point to a developmental trajectory along which less and less information from the linguistic context suffices as children get older, while at the same time children become able to process more and more information. Further systematic research on how the input is filtered through the parser will be crucial if we are to understand the tandem development of linguistic knowledge and language processing.

LEARNING OVER MULTIPLE SITUATIONS

Thus far we have considered what the child learner can glean from minimal exposure, as is commonly presented in most experimental studies. But often, one scenario is under-informative. For example, it may be unclear what 'a blick' means in one scenario in which a bear, a toy truck, and a lamp are present; but if across multiple scenarios, when 'a blick' is heard, the toy truck is always present, but not the other objects, then the learner can draw a statistically based inference that 'blick' refers to the truck. This cross-situational learning mechanism—making comparisons across multiple observational scenes—may be critical for learners to converge on a word's target meaning.[95] To deploy this mechanism, the learner must extract multiple hypotheses on every learning situation, store them in memory, and compare across learning situations to select the best hypothesis. But can learners do this—and even if they can, do they, in naturalistic learning situations? These questions continue to be hotly debated. Some studies find that adults and young children are sensitive to and track such statistical co-occurrence of words and referents,[95–98] but not without memory constraints.[99] Others find that learners do *not* store the entire referential set from an observational scene: In one study, Trueswell et al. found that even in a greatly simplified task with only two possible referents, learners only appeared to remember the referent they had initially hypothesized to be the referent of the novel word.[100] These authors proposed that learners use a mechanism they call 'propose-but-verify,' in which they store a single hypothesis in mind and then verify it against new learning situations.

Regardless of the specifics, it is clear that what learners take in from the input across situations is filtered through the learner's cognitive capacity (e.g., memory) and/or learning strategies (e.g., language users expect a speaker's utterance to have only one intended meaning). In other words, statistically based regularities may be available in the input, but biases, constraints, and learning heuristics are necessary for 'human-scale lexicons to be learned in human-scale time'[101] (also see Blythe et al.[102] and Reisenauer et al.[103]). For example, Frank et al.[104] and Yu and Ballard[105] showed that models incorporating social cues (e.g., speaker's intention) were more successful than purely statistical models.

Multiple encounters are also important for learners to generalize a newly-learned word appropriately beyond the referent with which it initially occurred to other members of the same category—the word 'dog' from the family poodle to all dogs, the word 'throw' from a pitch at Little League to all instances of throwing any kind of object. How do young learners achieve generalization at all, and the correct generalization at that? Variability serves as a strong source of evidence that a higher taxonomic level is the appropriate one; seeing multiple dogs all referred to as dogs is surely evidence that the word generalizes beyond a single dog, and should help the learner correctly identify the boundaries around the 'dog' category. Several studies have shown that variability aids the formation of word-object mappings.[51,106,107] Gentner, for example, has argued that analogical reasoning—or the ability to perceive the relational similarity between multiple exemplars—plays a fundamental role in the acquisition of word meanings, perhaps especially for words with inherently relational meanings such as verbs.[108] However, the role of variability in verb acquisition is as yet unclear, with studies conflicting as to whether the ability to compare across variable exemplars supports acquisition of novel verbs in the laboratory,[109] or whether less variability is better.[110] It is likely that factors such as the complexity of the event and the number of exemplars matter.[110–112]

Learning that a word refers to a lower taxonomic level may also require multiple encounters. For example, encountering a novel word 'blick' with a Dalmatian (subordinate-level) that is also a dog (basic-level), it may be hard to decide at which level the word refers. But if 'blick' is heard three times in the context of three different Dalmatians, the taxonomically more specific meaning—'Dalmatian'—is more likely, because it would be a 'suspicious coincidence' to observe only Dalmatians labeled 'blick' if the word actually referred to all kinds of dogs.[106] Whether such a generalization can be learned solely from the input[106] or whether it also rests on preexisting knowledge about the relationships between grammatical and conceptual categories (e.g., noun-object, adjective-property)[113] is currently under debate. Spencer et al. found that the 'suspicious coincidence' effect was reversed when objects were presented in sequences (as opposed to simultaneous presentation), suggesting an interaction of the statistical regularities in the input and learners' attention and memory constraints.[114] A recent study using computational modeling affirmed such an interaction.[115]

Thus, both 'fast mapping' and slower 'extended mapping' are mechanisms involved in word learning. Research on 'fast mapping' indicates that even on a single encounter with a new word, learners form some kind of initial representation of its meaning that can articulate their conceptual and lexical space.[116] With additional encounters with the word, learners can engage in a slower 'extended mapping' process of refining and consolidating their initial hypotheses about word meaning and integrating new lexical representations into their existing knowledge.[37,117]

CONCLUSION

Although the task of word learning is challenging, learners bring to bear numerous learning mechanisms. Interestingly, despite significant advances, both empirical and theoretical, in our understanding of how children use such mechanisms to acquire new words, many of the same debates continue to rage, including to what extent associationist mechanisms can explain behavior, to what extent the mechanisms that have been identified are specific to language or domain-general, and what the roles are of innate versus learned knowledge. Further, while most word learning research has focused on mechanisms for acquiring nouns that label objects, and slightly less research on verbs and adjectives, we know very little about acquisition of other kinds of words; this gap must receive more attention in future work. And compared to how much we know about the different mechanisms available to the learner, we know relatively less about how the learner selects which mechanism to apply, for a given word type, at a given developmental stage, under a given learning situation. This deserves more emphasis in future work.

Nevertheless, one area in which we are at an exciting new frontier is in understanding relationships between input and intake. As we have seen, while children must have access to environmental input to acquire the language of their community, this input is not veridically represented, but rather transformed into an organized, interpreted, and stored representation through interaction with the learner's linguistic and cognitive systems. Future work must continue to relate studies of the input with studies of children's linguistic and cognitive abilities. On both sides, more sophisticated methods have been developed to permit more sophisticated research. With respect to input, children's linguistic and extralinguistic environments can now be studied with dense sampling via LENA systems[118] or via video capture in the home.[119] Children's abilities are now being studied with neuroimaging, which, in combination with behavioral measures such as eye-tracking and language production, lends precision to our picture of what the relevant mechanisms are and how they work. But relating input and learning, too, has made strides, as we now have ever more nuanced discussions of children's intake of the input.[83,84,120] For example, mediation modeling has shown that language processing speed is a mediator in the relationship between the quality of maternal language input as measured in LENA recordings and children's vocabulary outcomes.[121] Serious research on individual differences in both input and intake is relatively new but will play a tremendously important role going forward. So too, will be greater emphasis on studies of how individual languages, cultures, and other environmental features shape the trajectory of lexical development.

ACKNOWLEDGMENTS

Research reported in this publication was supported by NIDCD of the National Institutes of Health under award number K01DC013306. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

1. Fenson L, Dale PS, Reznick JS, Bates E, Thal DJ, Pethick SJ, Tomasello M, Mervis CB, Stiles J. Variability in early communicative development. *Monogr Soc Res Child Dev* 1994, 59: i+iii-v+1–185. doi:10.2307/1166093.
2. Tomasello M. The item-based nature of children's early syntactic development. *Trends Cogn Sci* 2000, 4:156–163. doi:10.1016/S1364-6613(00)01462-5.
3. Quine WO. *Word and Object*. Cambridge, MA: MIT Press; 1960.
4. Gagliardi A. All input isn't equal: how the nature of the learner shapes language acquisition. *Stud Linguist* 2013, 67:68–81. doi:10.1111/stul.12005.
5. Saffran JR, Aslin RN, Newport EL. Statistical learning by 8-month-old infants. *Science* 1996, 274:1926–1928.
6. Ngon C, Martin A, Dupoux E, Cabrol D, Dutat M, Peperkamp S. (Non)words, (non)words, (non)words: evidence for a protolexicon during the first year of life. *Dev Sci* 2013, 16:24–34. doi:10.1111/j.1467-7687.2012.01189.x.
7. Graf Estes K, Evans JL, Alibail MW, Saffran JR. Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychol Sci* 2007, 18:254–260. doi:10.1111/j.1467-9280.2007.01885.x.

8. Shukla M, White KS, Asline RN. Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-month-old infants. *Proc Natl Acad Sci USA* 2011, 108:6038–6043. doi:10.1073/pnas.1017617108.
9. Johnson EK, Tyler MD. Testing the limits of statistical learning for word segmentation. *Dev Sci* 2010, 13:339–345. doi:10.1111/j.1467-7687.2009.00886.x.
10. Pearl L, Goldwater S, Steyvers M. Online learning mechanisms for Bayesian models of word segmentation. *Res Lang Comput* 2010, 8:107–132. doi:10.1007/s11168-011-9074-5.
11. Mintz TH. Finding the verbs: distributional cues to categories available to young learners. In: Hirsh-Pasek K, Golinkoff RM, eds. *Action Meets Word: How Children Learn Verbs*. New York: Oxford University Press; 2006, 31–63.
12. Shi R, Melançon A. Syntactic categorization in French-learning infants. *Infancy* 2010, 15:517–533. doi:10.1111/j.1532-7078.2009.00022.x.
13. Höhle B, Weissenborn J, Kiefer D, Schulz A, Schmitz M. Functional elements in infants' speech processing: the role of determiners in the syntactic categorization of lexical elements. *Infancy* 2004, 5:341–353.
14. Cauvet E, Limissuri R, Millotte S, Skoruppa K, Cabrol D, Christophe A. Function words constrain on-line recognition of verbs and nouns in French 18-month-olds. *Lang Learn Dev* 2014, 10:1–18. doi:10.1080/15475441.2012.757970.
15. de Carvalho A, He AX, Lidz J, Christophe A. 18-Month-olds use phrasal prosody and function words to constrain the acquisition of novel words, Under review.
16. Brown RW. Linguistic determinism and the part of speech. *J Abnorm Soc Psychol* 1957, 55:1–5.
17. Waxman SR, Markow DB. Object properties and object kind: twenty-one-month-old infants' extension of novel adjectives. *Child Dev* 1998, 69:1313–1329. doi:10.2307/1132268.
18. Hall DG, Graham SA. Lexical form class information guides word-to-object mapping in preschoolers. *Child Dev* 1999, 70:78–91. doi:10.1111/1467-8624.00007.
19. Bernal S, Lidz J, Millotte S, Christophe A. Syntax constrains the acquisition of verb meaning. *Lang Learn Dev* 2007, 3:325–341. doi:10.1080/15475440701542609.
20. Oshima-Takane Y, Ariyama J, Kobayashi T, Katerelos M, Poulin-Dubois D. Early verb learning in 20-month-old Japanese-speaking children. *J Child Lang* 2011, 38:455–484. doi:10.1017/S0305000910000127.
21. He AX, Lidz J. Verb learning in 14- and 18-month-old English-learning infants. *Lang Learn Dev*. In press.
22. Waxman SR, Booth AE. Seeing pink elephants: fourteen-month-olds' interpretations of novel nouns and adjectives. *Cogn Psychol* 2001, 43:217–242.
23. Gillette J, Gleitman H, Gleitman LR, Lederer A. Human simulations of vocabulary learning. *Cognition* 1999, 73:135–176. doi:10.1016/S0010-0277(99)00036-0.
24. James W. *The Principles of Psychology*. New York: Henry Holt and Co; 1890.
25. Baillargeon R. Object permanence in 3½-and 4½-month-old infants. *Dev Psychol* 1987, 23:655–664. doi:10.1037/0012-1649.23.5.655.
26. Spelke ES. Principles of object perception. *Cognit Sci* 1990, 14:29–56. doi:10.1016/0364-0213(90)90025-R.
27. Golinkoff RM, Harding CG, Carlson V, Sexton ME. The infant's perception of causal events: the distinction between animate and inanimate objects. *Adv Infancy Res* 1984, 3:145–151.
28. Soja NN, Carey S, Spelke ES. Ontological categories guide young children's inductions of word meaning: object terms and substance terms. *Cognition* 1991, 38:179–211. doi:10.1016/0010-0277(91)90051-5.
29. Carey S. Where our number concepts come from. *J Philos* 2009, 106:220–254.
30. Göksun T, Hirsh-Pasek K, Golinkoff RM, Imai M, Konishi H, Okada H. Who is crossing where? Infants' discrimination of figures and grounds in events. *Cognition* 2011, 121:176–195. doi:10.1016/j.cognition.2011.07.002.
31. Wagner L, Carey S. Individuation of objects and events: a developmental study. *Cognition* 2003, 90:163–191. doi:10.1016/S0010-0277(03)00143-4.
32. Leslie AM, Keeble S. Do six-month-old infants perceive causality? *Cognition* 1987, 25:265–288. doi:10.1016/S0010-0277(87)80006-9.
33. Wagner L, Carey S. 12-month-old infants represent probable endings of motion events. *Infancy* 2005, 7:73–83. doi:10.1207/s15327078in0701_6.
34. Pulverman R, Golinkoff RM, Hirsh-Pasek K, Buresh JS. Infants discriminate manners and paths in non-linguistic dynamic events. *Cognition* 2008, 108:825–830. doi:10.1016/j.cognition.2008.04.009.
35. Hernik M, Csibra G. Infants learn enduring functions of novel tools from action demonstrations. *J Exp Child Psychol* 2015, 130:176–192. doi:10.1016/j.jecp.2014.10.004.
36. Gleitman LR. The structural sources of verb meanings. *Lang Acquisit* 1990, 1:3–55. doi:10.1207/s15327817la0101_2.
37. Carey S. The child as word learner. In: Halle M, Bresnan J, Miller A, eds. *Linguistic Theory and Psychological Reality*. Cambridge, MA: MIT Press; 1978, 264–293.

38. Markman EM. Constraints children place on word meanings. *Cognit Sci* 1990, 14:57–77. doi:10.1016/0364-0213(90)90026-S.
39. Landau B, Smith LB, Jones S. Object shape, object function, and object name. *J Mem Lang* 1998, 38:1–27. doi:10.1006/jmla.1997.2533.
40. Markman EM, Hutchinson JE. Children's sensitivity to constraints on word meaning: taxonomic versus thematic relations. *Cogn Psychol* 1984, 16:1–27.
41. Waxman SR, Kosowski TD. Nouns mark category relations: toddlers' and preschoolers' word-learning biases. *Child Dev* 1990, 61:1461–1473. doi:10.1111/j.1467-8624.1990.tb02875.x.
42. Clark EV. The principle of contrast: a constraint on language acquisition. In: MacWhinney B, ed. *Mechanisms of Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1987.
43. Golinkoff RM, Mervis CB, Hirsh-Pasek K. Early object labels: the case for a developmental lexical principles framework. *J Child Lang* 1994, 21:125–155. doi:10.1017/S0305000900008692.
44. Diesendruck G, Markson L. Children's avoidance of lexical overlap: a pragmatic account. *Dev Psychobiol* 2001, 37:630–641. doi:10.1037/0012-1649.37.5.630.
45. Horst JS, Samuelson LK, Kucker SC, McMurray B. What's new? Children prefer novelty in referent selection. *Cognition* 2011, 118:234–244. doi:10.1016/j.cognition.2010.10.015.
46. Halberda J. Is this a dax which i see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cogn Psychol* 2006, 53:310–344.
47. Kalashnikova M, Mattock K, Monaghan P. Flexible use of mutual exclusivitity in word learning. *Lang Learn Dev* 2016, 12:79–91. doi:10.1080/15475441.2015.1023443.
48. Hollich GJ, Hirsh-Pasek K, Golinkoff RM, Brand RJ, Brown E, Chung HL, Hennon E, Rocroi C, Bloom L. Breaking the language barrier: an emergentist coalition model for the origins of word learning. *Monogr Soc Res Child Dev* 2000, 65 i-vi+:1–135.
49. Samuelson LK. Building biases: word learning as a developmental cascade. *Psychol Sci Agenda* 2011, 25. Available at: http://www.apa.org/science/about/psa/2011/02/building-biases.aspx. (Accessed December 30, 2016).
50. Spencer JP, Blumberg MS, McMurray B, Robinson SR, Samuelson LK, Tomblin JB. Short arms and talking eggs: why we should no longer abide the nativist-empiricist debate. *Child Dev Perspect* 2009, 3:79–87. doi:10.1111/j.1750-8606.2009.00081.x.
51. Perry LK, Samuelson LK, Malloy LM, Schiffer RN. Learn locally, think globally: exemplar variability supports higher-order generalization and word learning. *Psychol Sci* 2010, 21:1894–1902. doi:10.1177/0956797610389189.
52. Smith LB, Jones SS, Landau B, Gershkoff-Stowe L, Samuelson LK. Object name learning provides on-the-job training for attention. *Psychol Sci* 2002, 13:13–19.
53. Spelke ES, Kinzler KD. Core knowledge. *Dev Sci* 2007, 10:89–96. doi:10.1111/j.1467-7687.2007.00569.x.
54. Diesendruck G, Shatz M. Two-year-olds' recognition of hierarchies: evidence from their interpretation of the semantic relation between object labels. *Cogn Dev* 2001, 16:577–594. doi:10.1016/S0885-2014(01)00047-8.
55. Dautriche I, Chelma E, Christophe A. Word learning: homophony and the distribution of learning exemplars. *Lang Learn Dev* 2016, 12:231–251. doi:10.1080/15475441.2015.1127163.
56. Ferguson B, Waxman SR. What the *[beep]*? Six-month-olds link novel communicative signals to meaning. *Cognition* 2016, 146:185–189. doi:10.1016/j.cognition.2015.09.020.
57. Baldwin DA. Infants' ability to consult the speaker for clues to word reference. *J Child Lang* 1993, 20:395–418. doi:10.1017/S0305000900008345.
58. Senju A, Csibra G. Gaze following in human infants depends on communicative signals. *Curr Biol* 2008, 18:668–671. doi:10.1016/j.cub.2008.03.059.
59. O'Connell L, Poulin-Dubois D, Demke T, Guay A. Can infants use a nonhuman agent's gaze direction to establish word-object relations? *Infancy* 2009, 14:414–438. doi:10.1080/15250000902994073.
60. Tomasello M, Barton M. Learning words in nonostensive contexts. *Dev Psychobiol* 1994, 30:639–650. doi:10.1037/0012-1649.30.5.639.
61. Akhtar N, Carpenter M, Tomasello M. The role of discourse novelty in early word learning. *Child Dev* 1996, 67:635–645. doi:10.2307/1131837.
62. Poulin-Dubois D, Forbes JN. Toddlers' attention to intentions-in-action in learning novel action words. *Dev Psychobiol* 2002, 38:104–114. doi:10.1037//0012-1649.38.1.104.
63. Sabbagh MA, Baldwin DA. Learning words from knowledge versus ignorant speakers: links between preschoolers' theory of mind and semantic development. *Child Dev* 2001, 72:1054–1070.
64. Scofield J, Behrend DA. Learning words from reliable and unrelabile speakers. *Sci Direct* 2008, 23:278–290. doi:10.1016/j.cogdev.2008.01.003.
65. Krcmar M, Grela B, Lin K. Can toddlers learn vocabulary from television? An experimental approach. *Media Psychol* 2007, 10:41–63. doi:10.1080/15213260701300931.
66. Roseberry S, Hirsh-Pasek K, Golinkoff RM. Skype me! Socially contingent interactions help toddlers

learn language. *Child Dev* 2014, 85:956–970. doi:10.1111/cdev.12166.

67. O'Doherty K, Troseth GL, Shimpi PM, Goldenberg E, Akhtar N, Saylor MM. Third-party social interaction and word learning from video. *Child Dev* 2011, 82:902–915. doi:10.1111/j.1467-8624.2011.01579.x.
68. Arunachalam S. Two-year-olds can begin to acquire verb meanings in socially impoverished contexts. *Cognition* 2013, 129:569–573. doi:10.1016/j.cognition.2013.08.021.
69. Akhtar N, Jipson J, Callanan MA. Learning words through overhearing. *Child Dev* 2001, 72:416–430. doi:10.1111/1467-8624.00287.
70. Yuan S, Fisher C. "Really? she blicked the baby?" Two-year-olds learn combinatorial facts about verbs by listening. *Psychol Sci* 2009, 20:619–626. doi:10.1111/j.1467-9280.2009.02341.x.
71. Iverson JM, Capirci O, Longobardi E, Caselli MC. Gesturing in mother-child interactions. *Cogn Dev* 1999, 14:57–75. doi:10.1016/S0885-2014(99)80018-5.
72. Ferguson B, Graf E, Waxman SR. Infants use known verbs to learn novel nouns: evidence from 15- and 19-month-olds. *Cognition* 2014, 131:139–146. doi:10.1016/j.cognition.2013.12.014.
73. Hirsh-Pasek K, Golinkoff RM. *The Origins of Grammar: Evidence from Early Language Comprehension*. Cambridge, MA: MIT Press; 1996.
74. Naigles LR. Children use syntax to learn verb meanings. *J Child Lang* 1990, 17:357–374. doi:10.1017/S0305000900013817.
75. Fisher C. Structural limits on verb mapping: the role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Dev Sci* 2002, 5:55–64. doi:10.1111/1467-7687.00209.
76. Papafragou A, Cassidy K, Gleitman LR. When we think about thinking: the acquisition of belief verbs. *Cognition* 2007, 105:125–165. doi:10.1016/j.cognition.2006.09.008.
77. Harrigan K, Hacquard V, Lidz J, eds. Syntactic bootstrapping in the acquisition of attitude verbs: *Think*, *want* and *hope*. In: *33rd West Coast Conference on Formal Lingusitics*. Somerville, MA: Cascadilla Proceedings Project; 2016.
78. Naigles LR. The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition* 1996, 58:221–251. doi:10.1016/0010-0277(95)00681-8.
79. Scott RM, Fisher C. Two-year-olds use distributional cues to interpret transitivity-alternating verbs. *Lang Cogn Process* 2009, 24:777–803. doi:10.1080/01690960802573236.
80. Naigles LR, Hoff-Ginsberg E. Why are some verbs learned before other verbs? effects of input frequency and structure on children's early verb use. *J Child Lang* 1998, 25:95–120. doi:10.1017/S0305000997003358.
81. Lee JN, Naigles LR. The input to verb learning in Mandarin Chinese: a role for syntactic bootstrapping. *Dev Psychol* 2005, 41:529–540. doi:10.1037/0012-1649.41.3.529.
82. Ural AE, Yuret D, Ketrez FN, Koçbas D, Küntay AC. Morphological cues vs. number of nominals in learning verb types in Turkish: the syntactic bootstrapping mechanism revisited. *Lang Cogn Process* 2009, 24:1393–1405. doi:10.1080/01690960902775525.
83. Kidd E, Bavin EL, Brandt S. The role of the lexicon in the development of the language processor. In: Bittner D, Ruhlig N, eds. *Lexical Bootstrapping: The Role of Lexis and Semantics in Child Language*. Berlin: De Gruyter Mouton; 2013.
84. Trueswell JC, Gleitman LR. Learning to parse and its implications for language acquisition. In: Gaskell G, Altmann GTM, eds. *Oxford Handbook of Psycholinguistics*. Oxford: Oxford University Press; 2007, 635–657.
85. He AX, Kon M, Arunachalam S. How much information is too much? Informativity and processing cost in verb learning. Submitted for publication.
86. Yuan S, Fisher C, Snedeker J. Counting the nouns: simple structural cues to verb meaning. *Child Dev* 2012, 83:1382–1399. doi:10.1111/j.1467-8624.2012.01783.x.
87. Gertner Y, Fisher C. Predicted errors in children's early sentence comprehension. *Cognition* 2012, 124:85–94. doi:10.1016/j.cognition.2012.03.010.
88. Dautriche I, Cristia A, Brusini P, Yuan S, Fisher C, Christophe A. Toddlers Default to Canonical Surface-to-Meaning Mapping When Learning Verbs. *Child Dev* 2014, 85:1168–1180. doi:10.1111/cdev.12164.
89. Wellwood A He AX, Lidz J, Williams A, eds. Participant structure in event perception: towards the acquisition of implicitly 3-place predicates. University of Pennsylvania Working Papers in Linguistics. Philadelphia, PA: University of Pennsylvania; 2015.
90. Arunachalam S, Waxman SR. Let's see a boy and a balloon: argument labels and syntactic frame in verb learning. *Lang Acquisit* 2015, 22:117–131. doi:10.1080/10489223.2014.928300.
91. He AX, Lidz J. When one cue is better than two: trade-off between processing load and informativity in verb learning and verb extension tasks. In: *XIX Biennial International Conference for Infant Studies (ICIS)*, New Orleans, LA; 2016.
92. Imai M, Li L, Haryu E, Okada H, Hirsh-Pasek K, Golinkoff RM, Shigematsu J. Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. *Child Dev* 2008, 79:979–1000. doi:10.1111/j.1467-8624.2008.01171.x.

93. Sheline L, Arunachalam S. When the gostak distims the doshes: novel verb learning from novel nouns. In: *Cognitive Development Society Biennial Meeting*, Memphis, TN; 2013.
94. Love J, Fisher C. Learning from Jabberwocky: 3.5-year-olds use purely structural information to interpret verbs. Society for Research in Child Development Biennial Meeting, Seattle, WA; 2013.
95. Yu C, Smith LB. Rapid word learning under uncertainty via cross-stiuational statistics. *Psychol Sci* 2007, 18:414–420. doi:10.1111/j.1467-9280.2007.01915.x.
96. Roembke T, McMurray B. Observational word learning: beyond propose-but-verify and associative Bean counting. *J Mem Lang* 2016, 87:105–127. doi:10.1016/j.jml.2015.09.005.
97. Yurovsky D, Fricker DC, Yu C, Smith LB. The role of partial knowledge in statistical word learning. *Psychon Bull Rev* 2014, 21:1–22. doi:10.3758/s13423-013-0443-y.
98. Vouloumanos A, Werker JF. Infants' learning of novel words in a stochastic environment. *Dev Psychobiol* 2009, 45:1611–1617. doi:10.1037/a0016134.
99. Vlach HA, Johnson SP. Memory constraints on infants' cross-situational statistical learning. *Cognition* 2013, 127:375–382. doi:10.1016/j.cognition.2013.02.015.
100. Trueswell JC, Medina TN, Hafri A, Gleitman LR. Propose but verify: fast mapping meets cross-situational word learning. *Cogn Psychol* 2013, 66:126–156. doi:10.1016/j.cogpsych.2012.10.001.
101. Yurovsky D, Frank MC. An integrative account of constraints on cross-situational learning. *Cognition* 2015, 145:53–62. doi:10.1016/j.cognition.2015.07.013.
102. Blythe RA, Smith K, Smith ADM. Learning times for large lexicons through cross-situational learning. *Cognit Sci* 2010, 34:620–642. doi:10.1111/j.1551-6709.2009.01089.x.
103. Reisenauer R, Smith K, Blythe RA. Stochastic dynamics of lexicon learning in an uncertain and nonuniform world. *Phys Rev Lett* 2013, 110:258701. doi:10.1103/PhysRevLett.110.258701.
104. Frank MC, Goodman ND, Tenenbaum JB. Using speakers' referential intentions to model early cross-situational word learning. *Psychol Sci* 2009, 20:578–585. doi:10.1111/j.1467-9280.2009.02335.x.
105. Yu C, Ballard DH. A unified model of early word learning: integrating statistical and social cues. *Neurocomputing* 2007, 70:2149–2165. doi:10.1016/j.neucom.2006.01.034.
106. Xu F, Tenenbaum JB. Word learning as Bayesian inference. *Psychol Rev* 2007, 114:245. doi:10.1037/0033-295X.114.2.245.
107. Waxman SR, Klibanoff RS. The role of comparison in the extension of novel adjectives. *Dev Psychobiol* 2000, 36:571–581. doi:10.1037//O012-1649.36.5.571.
108. Gentner D. Structure-mapping: a theoretical framework for analogy. *Cognit Sci* 1983, 7:155–170. doi:10.1207/s15516709cog0702_3.
109. Childers JB. The structural alignment and comparison of events in verb acquisition. In: *30th annual Cognitive Science Society meeting*; 2008.
110. Maguire MJ, Hirsh-Pasek K, Golinkoff RM, Brandone AC. Focusing on the relation: fewer exemplars facilitate children's initial verb learning and extension. *Dev Sci* 2008, 11:628–634. doi:10.1111/j.1467-7687.2008.00707.x.
111. Childers JB, Parrish R, Olson CV, Burch C, Fung G, McIntyre KP. Early verb learning: how do children learn how to compare events? *J Cogn Dev* 2016, 17:41–66. doi:10.1080/15248372.2015.1042580.
112. Scott RM, Fisher C. 2.5-year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition* 2012, 122:163–180. doi:10.1016/j.cognition.2011.10.010.
113. Gagliardi A, Bennett E, Lidz J, Feldman N. Children's inferences in generalizing novel nouns and adjectives. In: *Proceedings of the 34th Annual Conference of the Cognitive Science Society*; 2012, 354–359.
114. Spencer JP, Perone S, Smith LB, Samuelson LK. Learning words in space and time: probing the mechanisms behind the suspicious coincidence effect. *Psychol Sci* 2011, 22:1049–1057. doi:10.1177/0956797611413934.
115. Grant E, Nematzadeh A, Stevenson S, eds. The interaction of memory and attention in novel word generalization: a computational investigation. In: *38th annual conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society; 2016.
116. Carey S, Bartlett E. Acquiring a single new word. *Papers Rep Child Lang Dev* 1978, 15:17–29.
117. Swingley D. Fast mapping and slow mapping in children's word learning. *Lang Learn Dev* 2010, 6:179–183. doi:10.1080/15475441.2010.484412.
118. Xu D, Yapanel U, Gray S. *Reliability of the LENA Language Environment Analysis System in Young Children's Natural Home Environment*. Boulder, CO: LENA Foundation; 2009.
119. Roy D. New horizons in the study of child language acquisition. In: *10th annual conference of the International Speech Communication Association*; 2009.
120. Omaki A, Lidz J. Linking parser development to acquisition of syntactic knowledge. *Lang Acquisit* 2015, 22:158–192. doi:10.1080/10489223.2014.943903.
121. Weisleder A, Fernald A. Talking to children matters - Early language experience strengthens processing and builds vocabulary. *Psychol Sci* 2013, 24:2143–2152. doi:10.1177/0956797613488145.

