

Active Learning in Language Development

Ruthe Foushee^{1,✉}, Mahesh Srinivasan², and Fei Xu²

¹University of Chicago, Chicago, IL, United States of America

²University of California, Berkeley, Berkeley, CA, United States of America

Despite evidence that children play an active role in their own learning in other domains (e.g., causal learning), the tendency in language development has been to study children as passive recipients of adult guidance. Here, we argue that emphasizing the child’s ‘receipt’ of linguistic input and adult support overlooks language development as a fruitful domain in which to explore learners’ active, self-directed learning. We define active language learning contexts as ones where the learner *selects the linguistic information they want to receive, in order to enhance their own learning*. We review how children select the information they want to receive by (1) rationally deploying their attention among potential sources of language input, (2) ‘tuning in’ to ambient language to reduce uncertainty, (3) eliciting labels from knowledgeable interactants, and (4) eliciting and evaluating evidence to bear on linguistic hypotheses. Notably, all cases implicate some implicit sense on the part of the child of their own uncertainty and linguistic capability. We suggest that reframing the child as an active language learner introduces novel explanations for phenomena in language development, and argue that using language as a test domain for rational learning accounts can provide researchers with complex, ecologically valid learning tasks.

Keywords: *language acquisition* | *active learning* | *self-directed learning* | *ecological learning* | *rational attention* | *lexical development*

Introduction

Children are famously scrappy learners. And yet within the study of language development—a complex social system that children are highly motivated to master—researchers often discuss children as the passive recipients of adult guidance. The idea that children learn language *passively* might be intuitive, as adults often appear to be *teaching* children to talk, by engaging them with simplified, exaggerated speech (Soderstrom, 2007). Moreover, the public sphere is rife with examples that highlight the importance of language directed to children, including messages that liken talking to children to feeding them (e.g., Zauche et al., 2017). Here, we argue that this emphasis on the child’s ‘receipt’ of linguistic input and adult support overlooks language development as a fruitful domain in which to explore learners’ direction of their own learning. We define an active learner as a learner who selects the information they want to receive in order to enhance their own learning (Gureckis & Markant, 2012). Below, we show how adopting a more active view of children’s language learning can yield insights and new research directions.

Prior research shows that children demonstrate active learning in diverse ways from early in life. Infants’ attention to a stimulus is reliably related to its novelty, complexity, or learnability (e.g., Gerken et al., 2011; Kidd et al., 2012). Toddlers explore to reduce uncertainty (e.g., about

how a novel toy functions; Sim & Xu, 2017a). Preschoolers conduct impromptu experimental tests of their hypotheses (Sim & Xu, 2017b), and school-age children ask increasingly strategic verbal questions as they mature (Ruggeri & Lombrozo, 2015). Adults show enhanced learning and memory when they select the information they want to receive, instead of having it selected for them (Kachergis et al., 2013; Markant & Gureckis, 2014). Yet this vision of an intrepid, rational child learner—which has been central in research on causal and ecological learning—has been largely absent from research in language development. How might children’s apparent self-directed learning prowess in non-linguistic domains translate to language learning?

Because languages are culturally-transmitted systems of communication, there are limits on what children can learn independently (e.g., a child cannot learn that dogs are called “dog” in English without exposure to English). At the same time, language is a good candidate for self-directed learning. For one, children grasp the utility of linguistic communication from infancy (Martin et al., 2012), suggesting that they will be intrinsically motivated to learn language to communicate. For another, language development is resilient: children learn language across diverse contexts, from multi-family dwellings where young children are primarily cared for by slightly older peers, to nuclear family units in the Western, industrialized middle class (Hoff, 2006; Schieffelin & Ochs, 1987). Children may learn across these different contexts by playing a driving role in the learning process, taking advantage of available sources of language input, and seeking out relevant information as needed (Bloom, 2000).

In this review, we explore how adopting a view of the child as an *active language learner* might give us purchase on basic questions about how language development unfolds, and on what language input is effective for learning. In using the term “active learning” in the language domain, we refer to contexts where the learner *selects the linguistic information they want to receive, in order to enhance their own learning*. We focus on the child’s developing lexicon to illustrate active learning in language development for two reasons: (1) word forms are largely arbitrary conventions, therefore they must be learned, and learned *from* someone; (2) although some of the words that children know (e.g., “boo-boo”) likely come from language directed to them by caregivers, others are more likely to have been picked up from language directed to others (e.g., obscenities). Below, we review how children select the information they want to receive by (1) rationally deploying their attention, (2) tuning in to ambient language to reduce uncertainty, (3) eliciting input from knowledgeable interactants, (4) eliciting and evaluating evidence to bear on

linguistic hypotheses.

Children efficiently allocate their attention among potential sources of language input

Take a scenario that may be familiar: having just finished reading a story to your 3-year-old for the 12th time, they say, “Again! Again!” Research tying children’s attention to stimulus learnability raises the possibility that children ask to hear the same story again and again in part because there is still something for them to learn from it.

Do active language learners select the language data they want to receive so that their learning will be the most efficient? In support of this idea, Gerken and colleagues (2011) showed that the amount of attention infants paid to artificial language stimuli in the lab depended on the learnability of the artificial grammar. The *unlearnable* grammatical stimuli contained a conflicting blend of grammatical gender markings — half of exemplars conformed to one set of rules, and half conformed to another. In contrast, *learnable* grammatical stimuli all conformed to the same set of rules. Interestingly, 17-month-old infants looked away more quickly when listening to the unlearnable grammar, and stayed looking longest when the grammar they heard was *subjectively learnable* — inferred to be such because a previous sample of same-age infants was able to learn the rule, given a similar period of familiarization.

More recently, we directly tested the link between preschoolers’ self-directed attention to a stimulus — here, spoken English — and their learning (Foushee et al., 2021). Children (4–6 years) listened to a story narrated at either a Simple (using age-appropriate words) or Complex (using later-acquired words) level, while an eyetracker captured their visual attention across a storybook display. The audio narration for each page looped such that the story progressed faster if the child lost interest in the story and attended to a distractor instead. We expected that if a child’s attention to spoken language reflects its subjective complexity — i.e., how understandable or learnable it is for them — the same objective degree of complexity should be experienced differently by younger and older children, due to their different levels of linguistic competence. Indeed, children’s age predicted the degree to which the Complex versus Simple speech elicited and maintained their attention. In the Complex condition, older children were more likely than younger children to continue listening to the complex speech. However, the opposite was true in the Simple condition: younger children were more likely to continue listening than older children. This is the pattern of results that one would expect if a child’s attention to spoken language, in and out of lab, is responsive to how much they can learn from it. Indeed, individual children’s story comprehension and novel word learning, tested after the story, were systematically related to their attention to the speech.

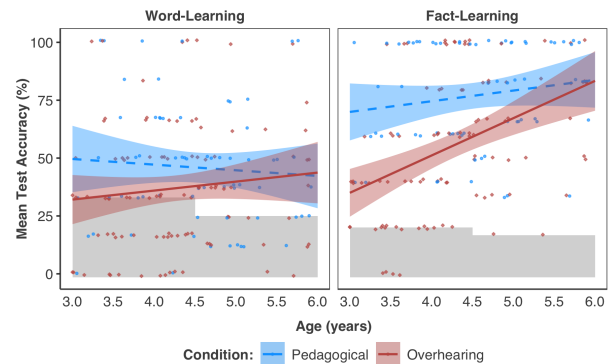


Fig. 1. Age Effect in *Foushee, Srinivasan, & Xu, 2021*. Mean test accuracy for two types of novel linguistic information corresponding to a set of unfamiliar toys: words (e.g., *pimwit*; left panel) and facts (e.g., *the one my sister loves*; right panel). Shaded region indicates accuracy at or below chance.

Children learn from ambient language in the absence of adult support

The internet is littered with vivid demonstrations of language that children are likely to have learned “actively.” For example, “where did my toddler learn to swear?” returns billions of search results, and clips of young children surprising us with their perfect mimicry of adult verbal behavior are often among the videos that go viral. In both cases, children are evincing knowledge of language that is unlikely to have been directed to them.

In a recent study, we assessed whether preschoolers tune in to overheard speech, across two experimental conditions that varied dramatically in terms of the demands they place on self-directed learning (Foushee et al., 2021). In both conditions, children (3–6 years) played with a mixed set of familiar and novel toys, and they had the opportunity to learn a set of new words and facts about the objects, embedded in dense, naturalistic speech. In the Pedagogical conditions, an experimenter cued the child’s attention to each object as it was discussed, while in the Overhearing conditions, an experimenter spoke on a phone call nearby, looking at neither the objects nor the child (Fig. 2A). Our results indicate that, with age, children learned to coordinate their attention between the overheard speech and the unfamiliar set of objects, and also became more and more likely to demonstrate robust word-learning from the overheard speech (Fig. 1); Strikingly, older (4.5–6-year-old) children learned four novel object labels equivalently from just one minute of overheard as from one minute of child-directed speech.

Together, the studies reviewed here and in the preceding section may help resolve an apparent paradox in the language development literature: although even toddlers are able to learn new words from overheard speech in experimental studies in simplified conditions, there is surprisingly little evidence that children learn words from overheard speech in their natural language environments (e.g., Shneidman & Goldin-Meadow, 2012). As rational learners “in the wild,”

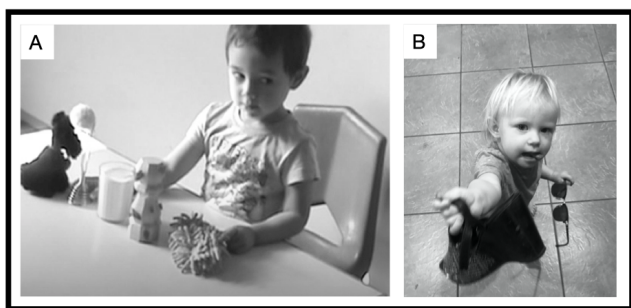


Fig. 2. Active learning behaviors in language development. A: Children ‘tune in’ to relevant ambient language. B: Children elicit language from caregivers through bids.

children may monitor all potential sources of language in order to learn most efficiently. In contexts where simplified child-directed speech is available, children may be less likely to attend to and learn from overheard speech that is relatively more complex; in contexts where child-directed speech is rare, children may adapt by attending to and learning more from overheard speech. This hypothesis may help explain why children appear to reach linguistic milestones on similar timetables cross-culturally (e.g., Casillas et al., 2020).

Children elicit labels from adults

Take another scenario: A toddler waddles over carrying an unfamiliar object (as in Fig. 2B) and asks some version of, “What’s this?” The caregiver replies with a label. As conventions, words only have value as much as they are agreed-upon by others. Thus, a child who wants to be able to talk about animals, plants, or *this-thing-I-just-found* will need to solicit help finding the right words.

A combination of experimental and observational evidence illustrates how children expand their vocabularies by eliciting linguistic information from those around them. Laboratory studies show that infants use pointing to request information from knowledgeable adults (Begus & Southgate, 2012), and that preschoolers both know when they don’t know what something is called (Lipowski et al., 2013), and who to explicitly ask for the word (Koenig & Harris, 2005). Indeed, many of the questions that pepper children’s early productions represent requests for linguistic information: in an analysis of four children’s spontaneous speech, for example, between 28% and 65% of the questions children asked between their first and second birthdays were requests for labels (Chouinard et al., 2007). Along these lines, children’s “active language learning” may be part of the explanation for why the onset of locomotion is so often correlated with a marked increase in vocabulary growth (He et al., 2015; Walle & Campos, 2014). With a newfound visual perspective (Kretch et al., 2014), autonomy (and freer hands), walkers encounter new and different things to name, and become more likely to make the sorts of attentional bids that elicit verbal responses from caregivers (Karasik et al., 2011). This of course remains speculative: unlike infants’ points (Begus & Southgate, 2012), and children’s questions (Chouinard et

al., 2007; Jimenez et al., 2018), evidence for children’s intentional role in eliciting information through bids—as opposed to merely sharing attention—remains indecisive.

Children test and refine their hypotheses about word meanings through feedback

Children use other, more capable language users not only to get new words, but also to refine their hypotheses about word meanings: e.g., a four-year-old in 2020 asks their mother, “is *coronavirus* really popular right now?” This question looks like an attempt to triangulate a new word’s meaning: what does *coronavirus* refer to, such that all adults are suddenly talking about it?

As when children in other domains selectively seek help based on task difficulty (Goupil et al., 2016), children seek linguistic information from the social world in a way that implicitly reflects their relative degree of uncertainty about how to use or interpret a given word. Children recruit help with how to define or interpret a word explicitly, as in the question above (Jimenez et al., 2018). They also engage in subtler information-seeking behaviors: preschool-aged (2–5 years) children scan an adult experimenter’s face more when the experimenter makes a referentially ambiguous request, e.g., asking them to give them the “modi” between two novel objects, relative to when the experimenter asks for the “dawnoo” between a novel object and a toy dinosaur (Hembacher et al., 2020).

Additional evidence for how young children monitor and reduce their relative uncertainty about different word meanings comes from a cross-situational word-learning study. Children were shown novel and familiar object-word pairs. Some of the novel word-object pairs were ambiguous (e.g., the same two novel objects always co-occurred with the same two novel words), while others could be disambiguated via an inference (i.e., if the words are “leemu” and “dog,” “leemu” must refer to the non-dog object). When later given the opportunity to learn more about specific objects, children (3–8 years) preferentially chose to sample referents whose labels remained ambiguous, and became even more likely to do so with age (Zettersten & Saffran, 2021). Thus, learners track not only their hypotheses about potential word meanings, but also the strength of their evidence, and actively seek additional information to reduce their uncertainty.

Discussion

The preceding sections have reviewed how an active learning framework can be extended to explain children’s remarkable language learning success. Several key outstanding questions remain, which we hope will inform future research:

- (1) How can we characterize the mechanisms of active learning in language development, and do these same mechanisms apply in other domains? Our explanations for how children sample linguistic information rely heavily on notions of *complexity*, *learnability*, and *uncertainty*, but what are children tracking, implicitly,

that makes their behavior explicable via these variables (e.g., compression rate? prediction error? entropy?), and does that signal have analogies in other domains?

- (2) Where can we see active language learning of other aspects of language knowledge? We have focused on the utility of active learning for acquiring the lexicon, but there are no doubt analogies for grammar-learning. For example, is there some ‘just right’ level of syntactic complexity to which children at a given stage of language development would be most attentive?
- (3) How does affect intersect with cognitive motivations for language learning? One promising area of research concerns the ways in which children’s lexical development reflects their *interests* (e.g., dinosaur names; Mani & Ackermann, 2018). On functional accounts of emotions (Barrett & Campos, 1987) children’s attention to ‘just right’ learning opportunities may be driven by positive affective experience, rather than metacognitive insight.

Conclusion

In our view, there is great potential in a research program at the intersection of active learning and language development. — especially one with an eye toward ecologically valid demonstrations of children’s abilities. Diverse empirical questions lie at this intersection. As the work reviewed here reveals, reframing the child as an *active* language learner introduces novel explanations for phenomena in the development of language. At the same time, using language as a test domain for formal accounts of rational learning can provide researchers with complex learning tasks that make sense to children, and are informative of how children navigate complexity within their own lives. Finally, applying the active learning framework to language development presents an opportunity to make our science more inclusive: the one-on-one pedagogical contexts that research and public policy tend to emphasize represent only a sliver of the language learning contexts that young children experience over the course of the day, across different households, and across different cultures. That children across diverse milieus become capable adult speakers may reflect children’s active role in getting the information they need.

References

- Barrett, K. C., & Campos, J. J. (1987). Perspectives on emotional development II: A functionalist approach to emotions. *Handbook of infant development*, 2nd ed. (pp. 555–578). John Wiley & Sons.
- Begus, K., & Southgate, V. (2012). Infant pointing serves an interrogative function. *Developmental science*, 15(5), 611–617.

- Bloom, L. (2000). The Intentionality Model of Word Learning: How to Learn a Word, Any Word. In R. Golinkoff, K. Hirsh-Pasek, L. Bloom, L. Smith, A. Woodward, N. Akhtar, M. Tomasello, & G. Hollich (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 19–50). NY: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195130324.003.002>
- Casillas, M., Brown, P., & Levinson, S. C. (2020). Early Language Experience in a Tzeltal Mayan Village. *Child Development*, 91(5), 1819–1835. <https://doi.org/10.1111/cdev.13349>
- Chouinard, M. M., Harris, P. L., & Maratsos, M. P. (2007). Children’s questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development*, i–129.
- Foushee, R., Srinivasan, M., & Xu, F. (2021). Self-directed learning by preschoolers in a naturalistic overhearing context. *Cognition*, 206(July 2020), 104415. <https://doi.org/10.1016/j.cognition.2020.104415>
- Gerken, L. A., Balcomb, F. K., & Minton, J. L. (2011). Infants avoid ‘labouring in vain’ by attending more to learnable than unlearnable linguistic patterns. *Developmental Science*, 14(5), 972–979. <https://doi.org/10.1111/j.1467-7687.2011.01046.x>
- Goupil, L., Romand-Monnier, M., & Kouider, S. (2016). Infants ask for help when they know they don’t know. *Proceedings of the National Academy of Sciences of the United States of America*, 113(13), 3492–3496. <https://doi.org/10.1073/pnas.1515129113>
- Gureckis, T. M., & Markant, D. B. (2012). Self-Directed Learning: A Cognitive and Computational Perspective. *Perspectives on Psychological Science*, 7(5), 464–481. <https://doi.org/10.1177/1745691612454304>
- He, M., Walle, E. A., & Campos, J. J. (2015). A cross-national investigation of the relationship between infant walking and language development. *Infancy*, 20(3), 283–305. <https://doi.org/10.1111/inf.12071>
- Hembacher, E., DeMayo, B., & Frank, M. C. (2020). Children’s Social Information Seeking is Sensitive to Referential Ambiguity. *Child Development*, 91(6), e1178–e1193. <https://doi.org/10.1111/cdev.13427>
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26(1), 55–88. <https://doi.org/10.1016/j.dr.2005.11.002>
- Jimenez, S., Sun, Y., & Saylor, M. M. (2018). The process of active word learning. *Active learning from infancy to childhood* (pp. 75–93). Springer.
- Kachergis, G., Yu, C., & Shiffrin, R. M. (2013). Actively learning object names across ambiguous situations. *Topics in Cognitive Science*, 5(1), 200–213.
- Karasik, L. B., Tamis-LeMonda, C. S., & Adolph, K. E. (2011). Transition from crawling to walking and infants’ actions with objects and people. *Child Development*, 82(4), 1199–1209. <https://doi.org/10.1111/j.1467-8624.2011.01595.x>

- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE*, 7(5). <https://doi.org/10.1371/journal.pone.0036399>
- Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development*, 76(6), 1261–1277. <https://doi.org/10.1111/j.1467-8624.2005.00849.x>
- Kretch, K. S., Franchak, J. M., & Adolph, K. E. (2014). Crawling and walking infants see the world differently. *Child Development*, 85(4), 1503–1518. <https://doi.org/10.1111/cdev.12206>
- Lipowski, S. L., Merriman, W. E., & Dunlosky, J. (2013). Preschoolers can make highly accurate judgments of learning. <https://doi.org/10.1037/a0030614>
- Mani, N., & Ackermann, L. (2018). Why Do Children Learn the Words They Do? *Child Development Perspectives*, 12(4), 253–257. <https://doi.org/10.1111/cdep.12295>
- Markant, D. B., & Gureckis, T. M. (2014). Is it better to select or to receive? learning via active and passive hypothesis testing. *Journal of Experimental Psychology: General*, 143(1), 94.
- Martin, A., Onishi, K. H., & Vouloumanos, A. (2012). Understanding the abstract role of speech in communication at 12months. *Cognition*, 123(1), 50–60. <https://doi.org/10.1016/j.cognition.2011.12.003>
- Ruggeri, A., & Lombrozo, T. (2015). Children adapt their questions to achieve efficient search. *Cognition*, 143, 203–216. <https://doi.org/10.1016/j.cognition.2015.07.004>
- Ruggeri, A., & Lombrozo, T. (2015). Children adapt their questions to achieve efficient search. *Cognition*, 143, 203–216.
- Schieffelin, B., & Ochs, E. (1987). *Language Socialization across Cultures*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511620898>
- Shneidman, L. A., & Goldin-Meadow, S. (2012). Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science*, 15(5), 659–673. <https://doi.org/10.1111/j.1467-7687.2012.01168.x>
- Sim, Z. L., & Xu, F. (2017a). Infants preferentially approach and explore the unexpected. *British Journal of Developmental Psychology*, 35(4), 596–608. <https://doi.org/10.1111/bjdp.12198>
- Sim, Z. L., & Xu, F. (2017b). Learning higher-order generalizations through free play: Evidence from 2- and 3-year-old children. *Developmental Psychology*, 53(4), 642–651. <https://doi.org/10.1037/dev0000278>
- Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, 27(4), 501–532. <https://doi.org/10.1016/j.dr.2007.06.002>
- Walle, E. A., & Campos, J. J. (2014). Infant language development is related to the acquisition of walking. *Developmental Psychology*, 50(2), 336.
- Zauche, L. H., Mahoney, A. E. D., Thul, T. A., Zauche, M. S., Weldon, A. B., & Stapel-Wax, J. L. (2017). The power of language nutrition for children's brain development, health, and future academic achievement. *Journal of Pediatric Health Care*, 31(4), 493–503.
- Zettersten, M., & Saffran, J. R. (2021). Sampling to learn words: Adults and children sample words that reduce referential ambiguity. *Developmental Science*, 24(3), 1–12. <https://doi.org/10.1111/desc.13064>