CHILD DEVELOPMENT PERSPECTIVES

Statistical Language Learning in Infancy

Jenny R. Saffran

University of Wisconsin–Madison

ABSTRACT—***Research suggests that infants use statistical regularities in linguistic input to identify and learn a range of linguistic structures, from the sounds of language (e.g., native-language speech sounds, word boundaries in continuous speech) to aspects of grammatical structure (e.g., lexical categories like nouns and verbs, basic aspects of syntax). In this article, I review the literature on statistical language learning in infants and raise questions about why infants are sensitive to statistical regularities. In doing so, I consider the relationship among statistical learning, prediction, and reducing uncertainty in infancy.***

KEYWORDS—*infancy; language; statistical learning*

Over the past two decades, much research has focused on statistical learning: the ability to detect patterns and regularities in the environment. While statistical learning has been investigated across a range of ages, it is arguably most useful early in life; infants have the most to learn and the least prior knowledge to guide that learning. Studies suggest that infants are sensitive to statistical regularities across numerous domains, including speech, music, actions, and visuospatial patterns (for a recent overview of infant statistical learning across domains, see Saffran & Kirkham, 2018). Because statistical learning has been investigated most extensively in the domain of language, in this review, I focus primarily on the role of statistical learning in language development. However, while this review is focused on first language learning by human infants, much of the ensuing discussion is also relevant to research on statistical learning at other developmental times and in other domains and species.

INFANT STATISTICAL LANGUAGE LEARNING

Early studies of statistical learning in infants were designed to address a fascinating problem facing novice language learners: How do infants discover words in fluent speech? This is a notoriously difficult challenge because unlike written text, fluent speech lacks spaces between words. Listen to a speaker of a language one does not know or watch a speaker of a sign language one does not know and it quickly becomes clear that pauses and other acoustic cues do not reliably indicate word boundaries (even in infant-directed speech or sign language). Jusczyk and Aslin (1995) were the first to demonstrate that infants could recognize English words after hearing them embedded in sentences. This classic study provided the first evidence that infants could pick out the sounds of words from a continuous speech stream.

The results of Jusczyk and Aslin's (1995) seminal studies, along with other research in the mid-1990s, provided clear evidence that infants can detect words in fluent speech. Several cues might facilitate this process for infants, including acoustic cues (e.g., allophonic alternations), prosodic cues (e.g., lexical stress), and other boundary cues (e.g., words in isolation, presence of familiar words). Another informative type of cue may reside in the statistical regularities in speech: Syllables that are part of the same words tend to co-occur more reliably than syllables that span word boundaries.

In a series of experiments, we tested the hypothesis that infants are sensitive to the probabilities with which syllables co-occur as a way to break into the speech stream to find words (Saffran, Aslin, & Newport, 1996). Eight-month-olds were first exposed to a 2-min stream of syllables, presented in a monotone, with no pauses or other auditory cues to word boundaries. The infants were then tested to determine whether they could distinguish the words in the speech stream from syllable sequences spanning word boundaries. The results suggested that infants

Jenny R. Saffran, University of Wisconsin–Madison.

Preparation of this article was supported in part by grants from the National Institute of Child Health and Human Development to the Waisman Center (U54 HD090256) and the author (R37 HD037466). Thank you to Dick Aslin and Elissa Newport for their leadership in the development of this area of research, and to the students of the UW-Madison Infant Learning Lab for many helpful discussions of these issues over the past two decades.

Correspondence concerning this article should be addressed to Jenny R. Saffran, Department of Psychology, University of Wisconsin–Madison, Madison, WI 53706; e-mail: jsaffran@wisc.edu.

DOI: 10.1111/cdep.12355

were sensitive to the statistical regularities in the speech stream, distinguishing between words (in which syllables co-occurred with high probabilities) and sequences spanning word boundaries (marked by lower probabilities of syllables co-occurring). In subsequent studies, infants' statistical sensitivity was not limited to simple artificial languages; infants could also use statistical regularities to detect word-like units in natural speech input that was more ecologically valid (Pelucchi, Hay, & Saffran, 2009a, 2009b).

Infants are sensitive to many different types of statistical regularities in linguistic input (for a theoretical review, see Erickson & Thiessen, 2015). The studies mentioned in the previous paragraph focused on conditional statistics such as transitional probabilities: the likelihood that one event predicts another (e.g., that in English, the syllable *pre* tends to be followed by *ty*, *tend*, or *dict*). Infants are sensitive to these regularities in both directions; that is, they detect not only that *pre* tends to be followed by *ty* but also that *ty* is frequently preceded by *pre* (Pelucchi et al., 2009b). These types of patterns help infants discover word-like units in fluent speech, generating candidate words that are available for mapping to meaning (Graf Estes, Evans, Alibali, & Saffran, 2007; Hay, Pelucchi, Estes, & Saffran, 2011). The patterns can also be used to find nonadjacent relationships between words when other units intervene, at least under some circumstances (e.g., Gomez, 2002).

Infants are not limited to tracking conditional statistics but are also sensitive to distributional statistics. One language-relevant example involves determining whether a particular continuum of speech sounds corresponds to one or two categories of phonemes. Consider the continuum of sounds between/ra/ and/la/: Some languages (e.g., English) divide this continuum of sounds into two categories, whereas other languages (e.g., Japanese, Korean) treat the same continuum of sounds as members of a single category. How do infants determine whether they are hearing a language with one or two phoneme categories represented in a single continuum? One solution to this problem is for infants to detect whether there are one or two peaks in the histogram (bumps in frequency of occurrence) along the continuum, which corresponds to the number of underlying speech categories (e.g., Maye, Werker, & Gerken, 2002). This type of statistical regularity is particularly important for acquiring the native-language phonemic repertoire. Distributional statistics help inform infants about how their language breaks speech sounds into phoneme categories.

Infants also use statistical information in the context of mapping labels to referents. Figuring out which words refer to which objects or events is a notoriously challenging problem: Novel words frequently occur in the context of many possible referents. One potential solution is to track the correlations between the co-occurrences of labels and their potential referents. For example, consider the word *pineapple*. The first time a baby hears this word, it may be in the context of an array of unfamiliar tropical fruits, making it unclear which fruit the word *pineapple* refers to. With subsequent occurrences of the word *pineapple*, an infant may be able to track the correlation between the word and the items in her visual field, eventually narrowing the mapping between *pineapple* and the correct fruit. While any individual situation may be ambiguous with respect to label-object pairs, infants can use cross-situational statistics to cope with widespread ambiguity in the mappings between words and the world (e.g., Smith & Yu, 2008).

Distributions of words can also help infants discover lexical categories by tracking which words tend to co-occur with one another (e.g., Mintz, 2003). Lexical categories like nouns and verbs are not signaled transparently in the input; infants must use regularities in the speech they hear (or the signs they see) to determine which items belong to which lexical categories. Statistical regularities in word sequences are very informative for identifying these categories. For example, nouns in English are frequently preceded by words like *a* or *the*, whereas verbs are never preceded by these kinds of words. In studies of artificial language learning, infants are sensitive to these types of regularities when acquiring lexical categories (e.g., Lany & Saffran, 2010).

Perhaps most strikingly, infants can detect statistical relationships between lexical categories, suggesting that they can track patterns of abstract elements (Saffran et al., 2008). Grammatical regularities in natural languages frequently exist in the relationships between lexical categories; as skilled language users, we can determine the grammaticality of novel sentences because we are familiar with the patterns that connect lexical categories in our language. In one study, 12-month-olds were sensitive to the grammaticality of sentences in an artificial language in which the pertinent regularities were found at the level of lexical categories, not individual words (Saffran et al., 2008). While it is far from clear that infants can use statistical patterns to learn complex aspects of syntax, particularly those for which the evidence in the input is infrequent or nonexistent (Han, Musolino, & Lidz, 2016; Lidz, Waxman, & Freedman, 2003), it is important that infants appear to be able to learn relationships between lexical categories and abstract elements in laboratory learning tasks.

To summarize, infants appear to be sensitive to a wealth of different statistical regularities, at least in laboratory tasks. That said, many questions remain about the nature of these mechanisms, including their neural mechanisms, domain specificity, commonalities and differences, and relationship with other aspects of cognition, especially memory systems (for a recent theoretical approach to these issues, see Thiessen, 2017). In the next section, I turn to one such issue: the factors that lead infants to track statistical regularities.

WHY DO INFANTS TRACK STATISTICS?

As the evidence supporting the potential importance of statistical learning for theories of language development has grown, questions have emerged concerning *why* human infants learn in this way. Infants are neither obligated to track statistical regularities nor instructed to do so in the laboratory or in daily life. Moreover, most statistical learning tasks lack explicit rewards; infants detect these regularities in the absence of any external motivation to do so.

Why then do infants track statistical regularities? One possible explanation, at least in the domain of language, is that infants track linguistic patterns because they want to figure out how to communicate with their caregivers. While motivation to communicate undoubtedly plays an important role in infant language development, it seems unlikely to fully explain infants' sensitivity to statistical structure. Infants track many of the same types of patterns in both communicative and noncommunicative domains (e.g., musical tones, geometric shapes, computer alert sounds; for a review, see Saffran & Kirkham, 2018). Even newborn infants detect statistical regularities in speech long before they are plausibly engaged in communicative interactions (Fló et al., 2019). Nonhuman animals are also sensitive to some of the same statistical regularities as human infants (for a review, see Santolin & Saffran, 2018).

Next, I explore one hypothesis about why infants track statistical regularities: to generate expectations and predictions about their environment. I end with a related hypothesis: Learning itself is motivating, and infants are driven to attempt to reduce uncertainty—a process for which statistical information would be invaluable. These issues have been studied widely in research on both human adults and nonhuman animals, but they have received scant attention in developmental science, particularly in the domain of language.

STATISTICAL LEARNING AND PREDICTION

The literature on prediction is largely distinct from the literature on statistical learning. Infants and young children engage in predictive behavior across numerous domains. For example, young children listening to sentences use their knowledge about the relationships among words and word categories to anticipate which words are likely to come next: Hearing the word *pirate* leads to the expectation that *ship* is likely to occur soon thereafter (e.g., Borovsky, Elman, & Fernald, 2012). Similarly, toddlers who speak a language with grammatical gender (e.g., Spanish, French) can use the gender of the article preceding the target noun to anticipate the likely next word (e.g., Lew-Williams & Fernald, 2007). These behaviors can be considered predictive because children are generating and acting on an expectation about what noun is likely to follow the word or words they just heard. However, these behaviors are likely not governed solely by statistical regularities. Most examples of incremental language processing in childhood, like those just described, are not statistical per se; in these examples, listeners can use semantic information (e.g., the connection between *pirate* and *ship*) and grammatical information (e.g., whether a word is masculine, feminine, or neuter).

Language is a particularly informative domain in which to consider relationships between statistical learning and predictive processes: Natural languages contain rich statistical structure, and linguistic input unfolds in time. One way statistical regularities help infants generate expectations is in the realm of sound sequences. When infants are familiarized with a particular type of sound-pattern regularity, they are subsequently more successful at segmenting word forms consistent with that regularity from fluent speech (e.g., Saffran & Thiessen, 2003; Sahni, Seidenberg, & Saffran, 2010). Sound-pattern regularities also facilitate mapping sounds to meanings. For example, infants familiarized with word forms that contain a specific phonotactic structure (regularities about which sounds precede and follow which other sounds) are subsequently more successful at mapping novel word forms to referents when the novel forms are consistent with the familiarized forms (Breen, Pomper, & Saffran, 2019). These data suggest that infants generate expectations about likely word forms based on the regularities to which they have been exposed.

Studies in which infants are trained on novel label-object pairs following exposure to fluent speech streams (Graf Estes et al., 2007; Hay et al., 2011) provide further evidence of a relationship between statistical language learning and expectations. Infants are more likely to successfully map labels that were words from the fluent speech (strong internally cohesive statistical regularities) to novel referents than labels containing weaker statistical regularities. These data, along with word learning studies using labels that vary in their native-language statistical cohesion (Graf Estes & Bowen, 2013; Graf Estes, Edwards, & Saffran, 2011), support the view that statistical patterns influence infants' expectations about likely labels for objects.

Infant word learning is also influenced by the predictability of the events within which label-object pairings are situated. A tradeoff is apparent, with a label attached to an extremely unpredictable event (e.g., a ball floating in the air) enhancing word learning (Stahl & Feigenson, 2017) and a label attached to a moderately unpredictable event (e.g., a violation of a visual sequence) hindering word learning (Benitez & Saffran, 2018). Thus, the statistical regularities of the environment, whether derived from actual experience or from training in a laboratory, affect infants' expectations about what is likely to occur next, influencing the ease with which infants map labels onto the objects participating in those events.

This body of literature suggests that infants' linguistic expectations are influenced by the statistics of their environment. Looking ahead, studies should explore the degree to which infants' predictive behaviors are informed by the statistics of their linguistic environment. Researchers working in the visual domain have demonstrated that infants' visual behavior is influenced by the statistical structures to which they have been exposed. For example, infants use the previous locations of visual events to generate visual expectations about subsequent events (Romberg & Saffran, 2013; Tummeltshammer & Kirkham, 2013). Their skill in doing so is related to attainment of native-language vocabulary, suggesting some potential relationships between nonlinguistic statistical learning and language attainment (Reuter, Emberson, Romberg, & Lew-Williams, 2018). Toddlers can also make visual predictions based on the sequential statistics of human actions (Monroy, Gerson, & Hunnius, 2017). These studies raise important questions about analogous situations in language learning. For example, do sequential statistics—word combinations—influence infants' predictions about what word or words are likely to occur next? The literature about children's incremental language processing, discussed at the beginning of this section, suggests that infants can use semantic and grammatical information to make these predictions, but can they use statistical information either in lieu of or in addition to these other types of cues? Moreover, can infants learn from incorrect linguistic predictions, detecting their errors and updating their expectations, as observed with their visual predictions (Romberg & Saffran, 2013)? Addressing these types of questions will help integrate the research fields of statistical learning with those of prediction and expectation.

STATISTICAL LEARNING AND UNCERTAINTY

In addition to informing predictive processes, statistical learning may help infants figure out what to learn about. Researchers studying cognitive development have begun to focus on the potential role of uncertainty as a motivator for learning, providing new ways to think about seemingly abstract behaviors such as curiosity. Children are motivated by uncertainty as they decide how to allocate their attention to different aspects of their environment. For example, children's play behavior is influenced by opportunities to reduce uncertainty: Preschoolers preferentially select toys that do not operate in obvious ways, using the opportunity to isolate variables to discover underlying causes (e.g., Schulz & Bonawitz, 2007). In this situation, preschoolers' exploratory play suggests a preference for uncertainty, leading to subsequent behaviors that reduce uncertainty. In learning tasks, infants prefer stimuli that are neither too redundant nor too random, but that maximize opportunities for reducing uncertainty (e.g., Kidd, Piantadosi, & Aslin, 2012). Parallel literatures in developmental robotics and neuroscience support the efficacy of curiosity-driven learning procedures (e.g., Kidd & Hayden, 2015; Oudeyer & Smith, 2016). The emerging developmental literature suggests a view of infants as active learners who sample from their environment to maximize learning outcomes (e.g., Sim & Xu, 2017; Smith, Jayaraman, Clerkin, & Yu, 2018). Oudeyer and Smith (2016) summarize the concept this way:

> These theoretical advances lead to a definition of curiosity as an epistemic motivational mechanism that pushes an organism to explore activities for the primary sake of gaining information (as opposed to searching for information in service of achieving an external goal like finding food or shelter). (p. 493)

Statistical learning has a clear role in considerations of uncertainty as a motivator for attention and behavior. Uncertainty is related to experience with the statistics of the environment. Given infants' intense interest in linguistic stimuli, curiosity-based learning may be relevant to language acquisition. Indeed, several sources of evidence point to infants' active role in interactions with caregivers, whereby infants influence their own language input via the language-relevant behavior they elicit from their caregivers. For example, infants' pointing gestures signal their interest in learning: 19-month olds can more successfully learn labels for objects they had previously pointed to than for objects they had not pointed to (Lucca & Wilbourn, 2018). Similarly, infants' babbling behavior, when directed toward objects, increases caregivers' contingent responsiveness (Albert, Schwade, & Goldstein, 2017). These types of findings suggest that infants take an active role in shaping their language learning environment. However, they do not tell us much about the factors that motivate infants' choices of things to learn about, which may vary in terms of their novelty, salience, and affective significance.

I hypothesize that statistical regularities influence infants' decisions about which stimuli to sample from their environment. These regularities can be used to locate islands of uncertainty, permitting infants to subsequently take an active role in reducing that uncertainty. Despite the clear conceptual links among uncertainty, statistical learning, and language acquisition, these relationships have not been investigated. Part of the reason for this gap is methodological: How can infants' interest in learning about different types of stimuli be assessed directly? My colleagues and I recently developed infant-controlled eye-tracking methods that allow infants to select items for additional exposure during word learning tasks (Zettersten & Saffran, 2019). Our goal is to harness these methods to determine whether infants preferentially sample items about which they are more uncertain. More generally, we hope to explore the hypothesis that infants are motivated to track statistical regularities to identify and remediate uncertainty. These early processes may be among the general-purpose mechanisms at the roots of the development of curiosity (Kidd & Hayden, 2015).

CONCLUSION

More than two decades have passed since the first investigations of infant statistical language learning. Over the ensuing years, statistical learning perspectives have been integrated into many different approaches to the study of language development, as well as into myriad other aspects of child development, adult cognition and psycholinguistics, cognitive neuroscience, and cross-species comparative cognition. In this brief review, I pointed to areas that have been well studied (e.g., the use of statistical regularities to discover word boundaries in fluent speech), as well as other aspects of language development where statistical learning approaches are becoming entrenched.

In the latter part of the review, I connected statistical learning approaches to two other areas of intense current interest in language and cognitive development: prediction and uncertainty. Now that statistical learning approaches have matured to the point that we can begin to ask *why* questions, researchers can profitably engage in theory development that integrates insights from related areas of study. In particular, the contemporary literature suggests that infants are not passive sponges soaking up regularities in their environment. Instead, infants actively engage with the world to gather information about things that are interesting to them, thereby shaping their own environment (e.g., Lucca & Wilbourn, 2018).

In addition to expanding the theoretical reach of statistical learning accounts, it is also important to expand the methodological basis of research in this area. While experiments with highly simplified artificial languages have been informative in initiating research on statistical learning, it is crucial for theory building that researchers continue to expand their methodological armamentarium to use richer stimuli (for a recent critical review, see Frost, Armstrong, & Christiansen, 2019). It is also increasingly important to develop studies that tap many levels of language concurrently. Just as infants are not tasked with learning a single level of linguistic input at a time, our studies should connect learning at multiple levels. Language learning is a dynamic process; what a child can learn at any given moment depends on what she has already learned. With a richer base of experimental and computational results from which to draw, we can delineate more clearly both the promise and the limits of statistical learning approaches to language development.

REFERENCES

Albert, R., Schwade, J. A., & Goldstein, M. H. (2017). The social functions of babbling: Acoustic and contextual characteristics that facilitate maternal responsiveness. *Developmental Science*, *18*, e12641. https://doi.org/10.1111/desc.12641

Benitez, V., & Saffran, J. R. (2018). Predictable events enhance word learning in toddlers. *Current Biology*, *28*, 2787–2793. https://doi.org/10.1016/j.cub.2018.06.017

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, *112*, 417–436. https://doi.org/10.1016/j.jecp.2012.01.005

Breen, E., Pomper, R., & Saffran, J. (2019). Phonological learning influences label-object mapping in toddlers. *Journal of Speech, Language, and Hearing Research*, *62*, 1923–1932. https://doi.org/10.1044/2019_JSLHR-L-18-0131

Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review*, *37*, 66–108. https://doi.org/10.1016/j.dr.2015.05.002

Fló, A., Brusini, P., Macagno, F., Nespor, M., Mehler, J., & Ferry, A. L. (2019). Newborns are sensitive to multiple cues for word segmentation in continuous speech. *Developmental Science*, *22*, e12802. https://doi.org/10.1111/desc.12802

Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin*, *145*, 1128–1153. https://doi.org/10.1037/bul0000210

Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*, 431–436. https://doi.org/10.1111/1467-9280.00476

Graf Estes, K., & Bowen, S. (2013). Learning about sounds contributes to learning about words: Effects of prosody and phonotactics on infant word learning. *Journal of Experimental Child Psychology*, *114*, 405–417. https://doi.org/10.1016/j.jecp.2012.10.002

Graf Estes, K., Edwards, J., & Saffran, J. R. (2011). Phonotactic constraints on infant word learning. *Infancy*, *16*, 180–197. https://doi.org/10.1111/j.1532-7078.2010.00046.x

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, *18*, 254–260. https://doi.org/10.1111/j.1467-9280.2007.01885.x

Han, C. H., Musolino, J., & Lidz, J. (2016). Endogenous sources of variation in language acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 942–947. https://doi.org/10.1073/pnas.1517094113

Hay, J. F., Pelucchi, B., Estes, K. G., & Saffran, J. R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, *63*, 93–106. https://doi.org/10.1016/j.cogpsych.2011.06.002

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*, 1–23. https://doi.org/10.1006/cogp.1995.1010

Kidd, C., & Hayden, B. (2015). The psychology and neuroscience of curiosity. *Neuron*, *88*, 449–460. https://doi.org/10.1016/j.neuron.2015.09.010

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*, e36399. https://doi.org/10.1371/journal.pone.0036399

Lany, J., & Saffran, J. R. (2010). From statistics to meaning: Infants' acquisition of lexical categories. *Psychological Science*, *21*, 284–291. https://doi.org/10.1177/0956797609358570

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, *18*, 193–198. https://doi.org/10.1111/j.1467-9280.2007.01871.x

Lidz, J., Waxman, S., & Freedman, J. (2003). What infants know about syntax but couldn't have learned: Experimental evidence for syntactic structure at 18 months. *Cognition*, *89*, 295–303. https://doi.org/10.1016/S0010-0277(03)00116-1

Lucca, K., & Wilbourn, M. P. (2018). Communicating to learn: Infants' pointing gestures result in optimal learning. *Child Development*, *89*, 941–960. https://doi.org/10.1111/cdev.12707

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*, B101–B111. https://doi.org/10.1016/S0010-0277(01)00157-3

Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, *90*, 91–117. https://doi.org/10.1016/S0010-0277(03)00140-9

Monroy, C. D., Gerson, S. A., & Hunnius, S. (2017). Toddlers' action prediction: Statistical learning of continuous action sequences. *Journal of Experimental Child Psychology*, *157*, 14–28. https://doi.org/10.1016/j.jecp.2016.12.004

Oudeyer, P. Y., & Smith, L. B. (2016). How evolution may work through curiosity-driven developmental process. *Topics in Cognitive Science*, *8*, 492–502. https://doi.org/10.1111/tops.12196

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009a). Statistical learning in a natural language by 8-month-old infants. *Child Development*, *80*, 674–685. https://doi.org/10.1111/j.1467-8624.2009.01290.x

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009b). Learning in reverse: Eight-month-old infants track backward transitional probabilities. *Cognition*, *113*, 244–247. https://doi.org/10.1016/j.cognition.2009.07.011

Reuter, T., Emberson, L., Romberg, A., & Lew-Williams, C. (2018). Individual differences in nonverbal prediction and vocabulary size in infancy. *Cognition*, *176*, 215–219. https://doi.org/10.1016/j.cognition.2018.03.006

Romberg, A. R., & Saffran, J. R. (2013). Expectancy learning from probabilistic input by infants. *Frontiers in Developmental Psychology*, *3*, 610. https://doi.org/10.3389/fpsyg.2012.00610

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Saffran, J., Hauser, M., Seibel, R., Kapfhamer, J., Tsao, F., & Cushman, F. (2008). Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition*, *107*, 479–500. https://doi.org/10.1016/j.cognition.2007.10.010

Saffran, J. R., & Kirkham, N. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*, 181–203. https://doi.org/10.1146/annurev-psych-122216-011805

Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental Psychology*, *39*, 484–494. https://doi.org/10.1037/0012-1649.39.3.484

Sahni, S. D., Seidenberg, M., & Saffran, J. R. (2010). Connecting cues: Overlapping regularities support cue discovery in infancy. *Child Development*, *81*, 727–736. https://doi.org/10.1111/j.1467-8624.2010.01430.x

Santolin, C., & Saffran, J. R. (2018). Constraints on statistical learning across species. *Trends in Cognitive Sciences*, *22*, 52–63. https://doi.org/10.1016/j.tics.2017.10.003

Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology*, *43*, 1045–1050. https://doi.org/10.1037/0012-1649.43.4.1045

Sim, Z. L., & Xu, F. (2017). Infants preferentially approach and explore the unexpected. *British Journal of Developmental Psychology*, *35*, 596–608. https://doi.org/10.1111/bjdp.12198

Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences*, *22*, 325–336. https://doi.org/10.1016/j.tics.2018.02.004

Smith, L. B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*, 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Stahl, A. E., & Feigenson, L. (2017). Expectancy violations promote learning in young children. *Cognition*, *163*, 1–14. https://doi.org/10.1016/j.cognition.2017.02.008

Thiessen, E. D. (2017). What's statistical about learning? Insights from modelling statistical learning as a set of memory processes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*, 20160056. https://doi.org/10.1098/rstb.2016.0056

Tummeltshammer, K. S., & Kirkham, N. Z. (2013). Learning to look: Probabilistic variation and noise guide infants' eye movements. *Developmental Science*, *16*, 760–771. https://doi.org/10.1111/desc.12064

Zettersten, M., & Saffran, J. R. (2019). *Active sampling by infant learners*. Unpublished manuscript.