



Non-Linguistic Grammar Learning by 12-Month-Old Infants: Evidence for Constraints on Learning

Chiara Santolin & Jenny R. Saffran

To cite this article: Chiara Santolin & Jenny R. Saffran (2019): Non-Linguistic Grammar Learning by 12-Month-Old Infants: Evidence for Constraints on Learning, Journal of Cognition and Development, DOI: [10.1080/15248372.2019.1604525](https://doi.org/10.1080/15248372.2019.1604525)

To link to this article: <https://doi.org/10.1080/15248372.2019.1604525>



Published online: 29 Apr 2019.



Submit your article to this journal [↗](#)



View Crossmark data [↗](#)



Non-Linguistic Grammar Learning by 12-Month-Old Infants: Evidence for Constraints on Learning

Chiara Santolin^a and Jenny R. Saffran^b

^aUniversity Pompeu Fabra, Spain; ^bUniversity of Wisconsin-Madison

ABSTRACT

Infants acquiring their native language are adept at discovering grammatical patterns. However, it remains unknown whether these learning abilities are limited to language, or available more generally for sequenced input. The current study is a conceptual replication of a prior language study, and was designed to ask whether infants can track phrase structure-like patterns from nonlinguistic auditory materials (sequences of computer alert sounds). One group of 12-month-olds was familiarized with an artificial grammar including predictive dependencies between sounds concatenated into strings, simulating the basic structure of phrases in natural languages. A second group of infants was familiarized with a grammar that lacked predictive dependencies. All infants were tested on the same set of familiar strings vs. novel (grammar-inconsistent) strings. Only infants exposed to the materials containing predictive dependencies showed successful discrimination between the test sentences, replicating the results from linguistic materials, and suggesting that predictive dependencies facilitate learning from nonlinguistic input.

ARTICLE HISTORY

Received 21 Nov 2018
Accepted 02 Apr 2019
Revised 20 Mar 2019

Introduction

One of the unique features of linguistic syntax is phrase structure. Although words in spoken languages occur serially, one after the next, in an apparent random order, the structure of language is not linear nor casual. Words are grouped into phrases according to specific orderings, and phrases are embedded into other phrases conferring hierarchical organization to the input. To learn phrase structure, language learners must acquire statistical regularities that cluster lexical categories (like nouns and verbs) into phrases, allowing learners to identify the ordering of sentential constituents. This is a nontrivial learning task, requiring the detection of lexical categories and the relationships between them.

One cue that may guide detection of phrase structure is the *predictive dependencies* amongst words and classes of words (Saffran, 2001, 2002). These statistical relations link lexical categories within a phrase, based on the fact that the presence of a member of one lexical category (e.g., a determiner) depends on the presence of a member of another lexical category (e.g., a noun). In English, for example, the presence of a determiner (*a* or *the*) necessitates a noun somewhere within a sentence, but a noun can stand alone without a

CONTACT Chiara Santolin  chiara.santolin@upf.edu  Centre for Brain and Cognition, University Pompeu Fabra, c/ Ramon Trias Fargas 25-27, Barcelona 08005, Spain

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hjcd.

© 2019 Taylor & Francis

determiner. Predictive dependencies link word classes that are not necessarily adjacent to one another, as when determiners and nouns are separated by adjectives.

By 7 months of age, infants recognize simple patterns matching familiar structures (e.g., ABA, ABB) instantiated in a novel vocabulary, generalizing beyond trained sequences (Marcus, Vijayan, Rao, & Vishton, 1999). This capacity extends to nonlinguistic domains (e.g., Bulf, Brenna, Valenza, Johnson, & Turati, 2015; Ferguson & Lew-Williams, 2016; Johnson et al., 2009; Saffran, Pollak, Seibel, & Shkolnik, 2007) although general perceptual and cognitive factors constrain learning (see Santolin & Saffran, 2017 for review). Indeed, even 3-month-old infants can generalize beyond familiar stimuli when regularities are displayed in a manner consistent with infants' perceptual requirements (Ferguson, Franconeri, & Waxman, 2018).

By the end of the first year, infants start to grasp the syntactic structure of their native language, and acquire complex patterns such as finite-state grammars and hierarchical phrase structure in laboratory tasks. For example, after familiarization to finite-state grammars generating linear sequences of words, 12-month-old infants distinguished grammar-consistent vs. inconsistent sentences containing violations at different locations in the grammars, even when sentences were formed by novel words (Gomez & Gerken, 1999). At this age, infants appear to be sensitive to cues to linguistic phrase structure as well. Saffran et al. (2008) created miniature grammars which either included predictive dependency cues to phrase structure, or which did not include these cues (see Figures 1 and 2). The grammars generated strings of nonsense words clustered in categories and organized into phrases such that some lexical categories predicted other lexical categories. After familiarization, infants discriminated grammar-consistent vs. inconsistent test strings violating predictive dependencies between words. Interestingly, however, learning was only observed for the materials containing predictive dependency cues to phrase structure. In the absence of these cues, infants failed to demonstrate evidence of learning.

This body of research has expanded our knowledge about the type of syntactic structures human infants can detect. Infants' success at learning in these tasks may reflect their inherent interest in, and massive prior exposure to, linguistic materials, and their failures may reflect innate constraints on possible linguistic structures. Alternatively, infants' patterns of success and failure with linguistic materials, as a function of their

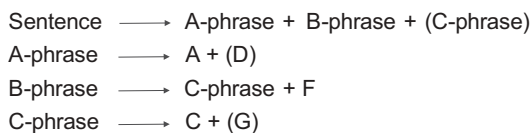


Figure 1. Predictive (P) Grammar. Each sound is denoted by a single letter (A, C, D, F, G); parentheses denote optional elements. Predictive dependencies are unidirectional e.g., D requires A but not vice versa.

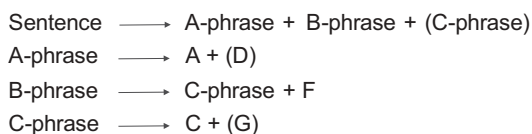


Figure 2. Non-Predictive (NP) Grammar.

statistical structure, may reflect the operation of domain-general learning mechanisms. Indeed, these data raise the tantalizing possibility that learning processes available across domains – not designed for language acquisition per se – have shaped the structure of natural languages (Christiansen & Chater, 2008, 2016; Chater & Christiansen, 2010; Saffran, 2001, 2002; see also Hawkins, 2004; O’Grady, 2005). If *detection of phrase structure via predictive dependencies across phrase-elements* is an ability that is not tailored specifically for learning linguistic structures, then we would expect similar patterns of successes and failures for learners acquiring nonlinguistic materials as well. This hypothesis was tested in a study of adults by Saffran (2002), who found that in both the auditory and visual domains, adults were better at learning nonlinguistic sequences that were organized with predictive dependencies than those that were not. Infant learners are also very good at tracking a variety of statistical patterns in nonlinguistic materials (see Saffran & Kirkham, 2018 for a recent review). However, infants’ ability to learn grammar-like nonlinguistic materials has not yet been tested. If human languages are shaped by general constraints on learning, then we would expect to see the same pattern of results – enhanced learning for materials containing predictive dependencies – for infants learning nonlinguistic structures.

The current study was designed to ask whether 12-month-old infants can track linguistic phrase-structure instantiated in nonlinguistic auditory input, as a conceptual replication of Saffran et al. (2008, Exp. 1). Importantly, the current study did not assess either generalization or category-based learning. Since no prior infant studies had used nonlinguistic auditory stimuli to test learning of hierarchical phrase structure, we decided to begin our investigation with the simplest version of the grammars, as in Saffran et al. (2008, Exp. 1). The materials consisted of nonlinguistic analogues of the artificial grammars used in the prior study. Predictive (P)-Grammar comprised predictive dependencies as cues to phrasal unit whereas the Non-Predictive (NP)-Grammar lacked these dependencies. In both cases, the “words” of each language consisted of nonlinguistic sounds i.e., highly discriminable computer alert sounds. The P-Grammar was head-final (unlike English, the participants’ native language) and mirrored basic natural language syntactic structures. The NP-Grammar violated typical structures of natural languages due to the greater optionality of its elements (the A-Phrase could include just the A-sound, just the D-sound or both the A- and D-sounds; see Figures 1 and 2). On other dimensions, the languages were as closely matched as possible. Both languages used the same vocabulary, and the training procedures were otherwise equivalent. Following exposure, infants were tested on the same set of familiar sentences, drawn from the exposure corpus, and novel (grammar-inconsistent) sentences. Note that successful learning did not require any form of generalization or category-based learning, since the familiar items were identical to the stimuli presented during exposure.

Our goal was to ask whether (a) infants could detect predictive dependencies characterizing phrase structure made up of nonlinguistic sounds, and (b) whether such predictive dependencies facilitated learning, as previously observed with linguistic materials. If infants simply learned the sound sequences presented during exposure, then they should show successful learning in both grammars. However, if infants can exploit predictive dependencies in nonlinguistic materials, we expected to see discrimination of familiar vs. novel (grammar-inconsistent) test sentences after exposure to the P-Grammar,

but not after the NP-Grammar – the same pattern observed in prior studies using these grammars.

Method

Participants

Participants were 57 infants (27 females) recruited from monolingual English-speaking homes in the Midwestern United States. The age range was 12.4–13.4 months (mean: 12.9), selected to be the same as Saffran et al. (2008, Exp. 1). Infants were randomly assigned to either the P-Grammar group (27 infants, 14 females; mean age 12.85 months) or the NP-Grammar group (30 infants, 13 females; mean age 12.88 months). Eleven additional infants were excluded from the final sample because their data did not meet the inclusion criterion set prior to data collection (at least 8 trials with looking times greater than 2 seconds). Twenty-six additional infants were excluded because of technical problems (14), fussiness (10), or crying (2). All participants were full term and had no history of hearing or vision problems. The study was conducted in accordance with ethical standards of the American Psychological Association. The protocol was approved by the local IRB; parents provided informed consent.

Materials

The materials were designed to be nonlinguistic analogues of the languages used in Saffran et al. (2008, Exp. 1). In both conditions (shown in Table 1), stimuli consisted of 8 strings generated by an artificial grammar, and produced using a “vocabulary” of 5 nonlinguistic sounds taken from the pool of Mac alert sounds (Mac OS Sierra, 10.12.5). The sounds were clearly discriminable from one another (*Glass, Basso, Hero, Ping, Sosumi*), and were intended to correspond to words in a linguistic grammar. The crucial difference between conditions was the structure of the grammar. In the P-Grammar condition, the sounds were concatenated into strings according to the grammar shown in Figure 1. Statistical dependencies between elements ensured that the presence of a given sound predicted the presence of another sound within the same phrase, not necessarily in an adjacent position. For instance, the A-Phrase required the presence of the A-sound whereas the D-sound was

Table 1. Familiarization and Test Strings.

Predictive (P) Grammar	Non-Predictive (NP) Grammar
Familiarization Strings	Familiarization Strings
ACF Glass-Basso-Hero	ACF Glass-Basso-Hero
ACGF Glass-Basso-Ping-Hero	ACGF Glass-Basso-Ping-Hero
ADCGF Glass-Sosumi-Basso-Ping-Hero	ADCGF Glass-Sosumi-Basso-Ping-Hero
ADCF Glass-Sosumi-Basso-Hero	ADCF Glass-Sosumi-Basso-Hero
ACFCG Glass-Basso-Hero-Basso-Ping	DCGF Sosumi-Basso-Ping-Hero
ADCFCG Glass-Sosumi-Basso-Hero-Basso-Ping	ADGF Glass-Sosumi-Ping-Hero
ACFC Glass-Basso-Hero-Basso	DCF Sosumi-Basso-Hero
ADCFG Glass-Sosumi-Basso-Hero-Ping	AGF Glass-Ping-Hero
Familiar (grammar-consistent) Test Strings	Novel (grammar-inconsistent) Test Strings
ACF Glass-Basso-Hero	ACDF Glass-Basso-Sosumi-Hero
ADCGF Glass-Sosumi-Basso-Ping-Hero	AGCF Glass-Ping-Basso-Hero
ADCF Glass-Sosumi-Basso-Hero	ADF Glass-Sosumi-Hero
ACGF Glass-Basso-Ping-Hero	ACGDF Glass-Basso-Ping-Sosumi-Hero

an optional element. However, if the D-sound was present, it was always preceded by A ($D|A = 1.0$). As the A-Phrase was always followed by the B-Phrase, both adjacent and nonadjacent conditional probabilities occurred between A and C sounds. Half of the P-Grammar strings contained the adjacent dependency AC, whereas the other half contained the nonadjacent dependency $\underline{A}DC$. In both cases, conditional probability of $C|A = 1.0$. In addition, the P-Grammar was characterized by a hierarchical structure: the C-Phrase was also embedded in the B-Phrase. In the NP-Grammar condition, the exposure language was formed by the same sounds but lacked predictive dependencies between sounds, and was thus characterized by overarching optionality in its structure (see [Figure 2](#)). The only statistical dependencies in the NP-grammar were negative: for example, if an A-sound was *not* present, a D-sound was present. Other regularities were shared by both grammars; for example, all strings contained at least three elements, and all strings began with the A-Phrase.

Strings generated by each grammar ranged from three to six sounds in length. Each sound lasted approximately 0.3 seconds; the duration of the concatenated sounds ranged between 1.8 and 3.3 seconds, with 0.5 seconds of silence between sounds within strings, and 1.5 seconds between strings. Sounds were concatenated in sequences using Praat version 6.0.20, and intensity was set at 60 dB.

The test items consisted of four familiar and four novel sound strings (see [Table 1](#)). Familiar strings were drawn from the pool of familiarization sequences and matched the familiarization grammar in both conditions, following regularities shared across P and NP grammars (e.g., all strings begin with A-Phrase). Novel, grammar-inconsistent strings were formed by re-combining familiar sounds into sequences that violate the grammar in both conditions. We used the same strings to test the two groups to control for idiosyncratic preferences for particular sound sequences. Each test item consisted of a single string of sounds repeated three times.

Procedure

The Headturn Preference Procedure was used to assess learning. Infants were seated on a caregiver's lap in a soundproof booth equipped with three computer screens placed in front of and on the two sides of the infant. Infants were familiarized with the language for 3 minutes, then received a 2-minute re-exposure, as in Saffran et al. (2008, Exp. 1). During familiarization, infants heard the stimuli and watched images presented on the central screen; the caregiver listened to music over headphones and provided snacks to keep infants engaged in the task. During re-exposure, which occurred immediately following familiarization, a blinking light was alternated contingent on infant looking behavior to familiarize infants with the methodology.

There were twelve test trials, three for each of the four test items, presented in random order. Half of the items were familiar (grammar-consistent), and half of them were novel (grammar-inconsistent). At the beginning of each trial, a pinwheel was displayed on the central screen until the infant fixated on it. At that point, the experimenter (blind to the audio stimuli) signaled the central pinwheel to extinguish and one of the side-pinwheels to pop up. When infants looked at the side pinwheel, one of the test items was repeated until the infant looked away for more than 2 seconds, or until 24 seconds had elapsed. Looking times were coded using custom-designed MATLAB software (R2010b, Mathworks, Inc.).

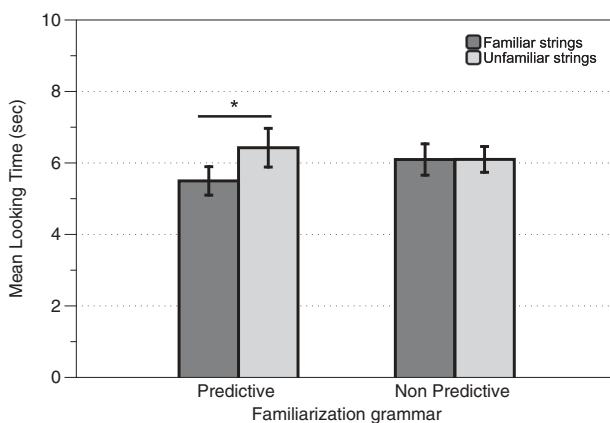


Figure 3. Looking times for familiar and novel strings (Y axis) in the two conditions (X axis). Vertical bars represent standard errors of the mean; asterisk indicates significant difference ($p < .05$) in looking time between test strings.

Results & discussion

As in Saffran et al. (2008), we ran matched-pairs t -tests to compare looking time for familiar versus novel stimuli in each condition. Following exposure to the P-Grammar, infants listened longer to novel than familiar strings ($t_{(26)} = 2.45$, $p = .021$, Cohen's $d = .47$, Bayes Factor in favor of alternative hyp. $BF_{10} = 2.5$). The average looking time was 5.49s for familiar strings and 6.42s for novel strings (Figure 3). Infants exposed to the NP-Grammar showed no preference ($t_{(29)} = .005$, $p = .995$, Bayes Factor in favor of null hyp. $BF_{01} = 5.1$); average looking time was 6.09s for both type of test strings. We performed a 2×2 repeated-measures ANOVA to examine the interaction between conditions (exposure to P- vs. NP-Grammar) and test strings (familiar vs. novel), which was not significant ($F_{(1,55)} = 2.36$, $p = .13$). Note that the same pattern of results, including a non-significant interaction, was also observed in Saffran et al. (2008, Exp. 1).

Only the infants familiarized to the P-Grammar were able to discriminate between test strings that followed the familiar structures from test strings that broke those structures. Recall that infants in both conditions received the same test; thus, differences in test materials cannot explain the observed differences in looking time behavior between the two groups of infants. Instead, the differences between groups are most likely attributable to differences in the structures of the two grammars.

Conclusions

As previously observed with linguistic materials, 12-month-old infants showed better learning when predictive dependencies characterized nonlinguistic phrase structure. Infants who heard materials containing predictive dependencies organizing the sounds successfully discriminated familiar test sentences from novel, grammar-inconsistent test sentences. Infants who heard materials that did not contain predictive dependencies failed to discriminate the same set of test items.

One limitation of this research is that the test stimuli did not allow to test generalization or category-based learning – the familiar test items were identical to the exposure items for infants in both conditions. In order to learn linguistic phrase structure, learners must be able to identify multiple members of the same lexical category, and discover grammatical patterns over categories rather than individual exemplars, as in the current study. Even without testing generalization, though, these results are suggestive, and consistent with prior studies. It is particularly striking that infants exposed to the less predictable grammar failed to distinguish the test items, despite the fact that they had heard all the consistent test items during familiarization (and none of the inconsistent items). This is exactly the same pattern of results observed by Saffran et al. (2008, Exp. 1) with linguistic stimuli. Moreover, it is important noting that discrimination in the P-Grammar condition did not necessarily imply hierarchical learning. Although P-Grammar comprised a hierarchical pattern (i.e., C-Phrase was embedded in B-Phrase), infants could have distinguished consistent vs. inconsistent test strings without forming hierarchical representations of the grammatical structure.

In line with previous research on predictable events (e.g., Bar, 2007; Benitez & Smith, 2012; Benitez & Saffran, 2018; see also Aslin, 2014), this evidence points to predictability as an important constraint on learning. Our findings suggest that even for nonlinguistic auditory structures, learning is facilitated when particular types of predictable patterns signal the structure of the input. Moreover, along with findings observed in linguistic tasks (Gomez & Gerken, 1999; Saffran et al., 2008), and in some nonhuman species (see Santolin & Saffran, 2017; Milne, Wilson, & Christiansen, 2018 for recent reviews), this evidence supports the existence of domain-general learning processes serving the beginning of syntax acquisition. Our next step is to extend such results to see whether infants can generalize nonlinguistic “syntactic” patterns to include new exemplars, or whether, alternatively, this remarkable ability is constrained by the cognitive domain in which it operates.

Low-level statistical patterns are unlikely to explain infants’ pattern of success and failure in our task. Saffran et al. (2008) addressed this possibility and found that n-gram properties of the stimuli (e.g., average item frequencies, bigram and trigram frequencies and probabilities) could not explain why infants showed discrimination after exposure to the P-Grammar but not the NP-Grammar. Importantly, since the materials used in the current research are structurally identical to those of the prior study, including the selection of test items, we are confident that the different learning outcomes reflect the grammatical structure of the two sets of exposure materials. Similarly, it is unlikely that element-based learning could entirely account for our results. If this was the case, to succeed at test, it would have been sufficient for the infants to track the patterns shared across both grammars (e.g., strings begin with A-Phrase).

Modern theories of human language learning propose that the learning mechanisms underlying language acquisition have shaped the structure of natural languages (e.g., Chater & Christiansen, 2010; Christiansen & Chater, 2008, 2016; Hawkins, 2004; Saffran, 2001, 2002). The organization of natural languages reflects the cognitive and learning abilities that are used to acquire such structures. The present research is aligned with this view, and supports the hypothesis that human learning capacities have shaped the type of structures – linguistic and nonlinguistic – that can be learned. Sensory inputs with predictive dependencies are easier to track than others lacking these cues. It is likely

the case that constraints on learning do not affect just the acquisition of linguistic structures but also patterns displaying similar complex structures such as music or visual scenes. Future studies comparing learning across modalities, as well as across ages and species, will be highly informative in addressing these fascinating issues.

Acknowledgments

We are grateful to Christine Potter, Viridiana Benitez and Nuria Sebastian Galles for helpful comments and feedbacks on this work. We are also grateful to participants and families, and to the Infant Learning Lab at University of Wisconsin-Madison for help and technical support during data collection. This research was supported by a postdoctoral fellowship from the Foundation Marica De Vincenzi ONLUS in association with the Department of Psychology and Cognitive Science of the University of Trento granted to C.S, and by grants from the National Institute of Health to J.R.S. (R37HD037466) and to the Waisman Center (U54 HD090256).

Disclosure statement

The authors declare no conflict of interest.

Funding

This work was supported by the National Institute of Health [R37HD037466].

References

- Aslin, R. (2014). Infant learning: Historical, conceptual, and methodological challenges. *Infancy*, 19(1), 2–27. doi:10.1111/inf.12036
- Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11(7), 280–289. doi:10.1016/j.tics.2007.05.005
- Benitez, V., & Saffran, J. (2018). Predictable events enhance word learning in toddlers. *Current Biology*, 28(17), 2787–2793. doi:10.1016/j.cub.2018.06.017
- Benitez, V., & Smith, L. (2012). Predictable locations aid early object name learning. *Cognition*, 125(3), 339–352. doi:10.1016/j.cognition.2012.08.006
- Bulf, H., Brenna, V., Valenza, E., Johnson, S. P., & Turati, C. (2015). Many faces, one rule: The role of perceptual expertise in infants' sequential rule learning. *Frontiers in Psychology*, 6, 1595. doi:10.3389/fpsyg.2015.01595
- Chater, N., & Christiansen, M. H. (2010). Language acquisition meets language evolution. *Cognitive Science*, 34(7), 1131–1157. doi:10.1111/j.1551-6709.2009.01049.x
- Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, 31(5), 489–509. doi:10.1017/S0140525X08004998
- Christiansen, M. H., & Chater, N. (2016). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, 39. doi:10.1017/S0140525X15001053
- Ferguson, B., Franconeri, S. L., & Waxman, S. R. (2018). Very young infants learn abstract rules in the visual modality. *PloS One*, 13(1), e0190185. doi:10.1371/journal.pone.0190185
- Ferguson, B., & Lew-Williams, C. (2016). Communicative signals support abstract rule learning by 7-month-old infants. *Scientific Reports*, 6, 25434.
- Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70(2), 109–135.
- Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford, UK: Oxford University Press.

- Johnson, S. P., Fernandes, K. J., Frank, M. C., Kirkham, N., Marcus, G., Rabagliati, H., & Slemmer, J. A. (2009). Abstract rule learning for visual sequences in 8- and 11-month-olds. *Infancy, 14*(1), 2–18. doi:[10.1080/15250000802569611](https://doi.org/10.1080/15250000802569611)
- Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283*(5398), 77–80.
- Milne, A. E., Wilson, B., & Christiansen, M. H. (2018). Structured sequence learning across sensory modalities in humans and nonhuman primates. *Current Opinion in Behavioral Sciences, 21*, 39–48. doi:[10.1016/j.cobeha.2017.11.016](https://doi.org/10.1016/j.cobeha.2017.11.016)
- O'Grady, W. (2005). *Syntactic carpentry: An emergentist approach to syntax*. Mahwah, NJ: Lawrence Erlbaum.
- Saffran, J. R. (2001). The use of predictive dependencies in language learning. *Journal of Memory and Language, 44*(4), 493–515. doi:[10.1006/jmla.2000.2759](https://doi.org/10.1006/jmla.2000.2759)
- Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language, 47*(1), 172–196. doi:[10.1006/jmla.2001.2839](https://doi.org/10.1006/jmla.2001.2839)
- Saffran, J. R., Hauser, M., Seibel, R., Kapfhamer, J., Tsao, F., & Cushman, F. (2008). Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition, 107*(2), 479–500. doi:[10.1016/j.cognition.2007.10.010](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. doi:[10.1146/annurev-psych-122216-011805](https://doi.org/10.1146/annurev-psych-122216-011805)
- Saffran, J. R., Pollak, S. D., Seibel, R. L., & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition, 105*(3), 669–680. doi:[10.1016/j.cognition.2006.11.004](https://doi.org/10.1016/j.cognition.2006.11.004)
- Santolin, C., & Saffran, J. R. (2017). Constraints on statistical learning across species. *Trends in Cognitive Sciences*. doi:[10.1016/J.TICS.2017.10.003](https://doi.org/10.1016/J.TICS.2017.10.003)