

The roles of item repetition and position in infant sequence learning

Christina Schonberg (ccs185@g.ucla.edu)

Department of Psychology, 502 Portola Plaza
Los Angeles, CA 90095 USA

Gary F. Marcus (gary.marcus@nyu.edu)

Department of Psychology, 6 Washington Place
New York, NY 10003 USA

Scott P. Johnson (scott.johnson@ucla.edu)

Department of Psychology, 502 Portola Plaza
Los Angeles, CA 90095 USA

Abstract

We examined mechanisms underlying infants' ability to detect, extract, and generalize sequential patterns, focusing on how saliency and consistency of distributional information guide infant learning of the most "likely" pattern in audiovisual sequences. In Experiment 1, we asked if 11- and 14-month-old infants could learn a "repetition anywhere" rule (e.g., ABBC, AABC, ABCC). In Experiment 2 we asked if 11- and 14-month-olds could generalize a "medial repetition" rule when its position is consistent in sequence, and in Experiment 3 we asked if 11-month-olds could identify a nonadjacent dependency occurring at edge positions. Infants were first habituated to 4-item sequences (shapes + syllables) containing repetition- and/or position-based structure, and were then tested with "familiar" structure instantiated across new items or combinations of items vs. "novel" (random) sequences. We found that 11-month-olds failed to learn the repetition rule both when the structure appeared in initial, medial, or final position (Experiment 1) and when it was restricted to the medial position (Experiment 2). Fourteen-month-olds learned repetition rules under both conditions. Finally, in Experiment 3 11-month-olds succeeded in learning a nonadjacent dependency in sequences identical to those used to test repetition learning in Experiment 2. Our results suggest that infants at 11 months, like adults, are relatively insensitive to patterns in the middle of sequences.

Keywords: infant learning; rule learning; sequence learning

Introduction

In the present paper, we examine mechanisms underlying infants' ability to detect, extract, and generalize sequential patterns. Sequence learning is essential for processes ranging from the acquisition of language to everyday activities such as preparing for bed, learning to count, learning to read, and getting ready for school. Insights into development of sequence learning in infancy, therefore, are vital for theories of developmental and cognitive function across a variety of domains.

What kinds of learning mechanisms are available to infants, and what are the limits of these mechanisms? Our particular focus is on two means of knowledge acquisition, "statistical learning" and "rule learning": the extent to which infants can use transitional probability information among items to extract units from an unbroken stream of stimuli

(e.g., Saffran, Aslin, & Newport, 1996) or the extent to which infants can distinguish simple reduplicative patterns from one another (e.g., Gerken, 2006; Marcus, Vijayan, Rao, & Vishton, 1999), respectively. Sequence learning is guided by multiple mechanisms (Arciuli, 2017; Krogh, Vlach, & Johnson, 2013; Thiessen, Kronstein, & Hufnagle, 2013), and its development in infancy can be better understood by investigations of the salience and consistency of statistical and rule-governed structures (Aslin & Newport, 2012, 2014). Some structures, such as identity relations or positions of items in order, might serve as "perceptual primitives," processed by specialized mechanisms to detect and remember specific features in patterned sequences (Endress, Nespor, & Mehler, 2009).

The Saffran et al. (1996) transitional probability task, a well-known example of *statistical* learning, presented 8-month-olds an auditory stimulus consisting of four unique strings (e.g., tupiro, golabu, bidaku, and padoti) presented in random order as a continuous, unsegmented stream for 2 minutes. Infants then heard isolated strings in repetition (e.g., tupiro, tupiro, tupiro...) alternating with "part-word" strings composed of parts of two of the familiar words (e.g., rogola, rogola...) from a speaker located either to the left or right. Infants exhibited a postfamiliarization novelty preference for the part-words relative to the words, implying that they detected the differences in transitional probability across word boundaries in the input sequence.

Rule learning in sequential patterns was assessed by Marcus et al. (1999), who exposed 7-month-olds to strings that followed either an "ABA" pattern (e.g., gah tee gah) or an "ABB" pattern (e.g., gah tee tee). After 2 minutes of exposure, the infants heard the same (familiar) pattern instantiated by different phonemes (e.g., woh fei woh, dee koh dee), and a second (novel) pattern on alternating trials and showed a preference for the novel, a result that extended to a test of ABB vs. AAB. Because transitional probabilities between test-string syllables were zero, performance could not have been based on statistical learning.

Studies of infant rule learning have produced mixed results with respect to the learnability of a simple repetition rule (adjacent, as in AAB or ABB, or nonadjacent, as in ABA). Overall, the findings of these experiments seem to

differ based on how familiar the infants may be with the stimuli: 7-month-olds successfully learn ABA vs. ABB, ABB vs. AAB, and AAB vs. ABB patterns when the stimuli are auditory (Marcus et al., 1999, 2007) or familiar visual stimuli (e.g., faces and animals; Bulf et al., 2015; Saffran et al., 2007). However, when stimuli consist of sequences of colored shapes, learning seems to be more difficult, perhaps because learning visual sequential input is constrained by limits in visual working memory (Johnson et al., 2009). In this case, 8-month-olds learned a late repetition rule (adjacent repetition in the final edge position) when tested vs. nonadjacent repetition (ABB vs. ABA), but failed to learn late vs. early repetition (ABB vs. AAB), early vs. nonadjacent repetition (AAB vs. ABA), and nonadjacent vs. late repetition (ABA vs. ABB). Eleven-month-olds learned all these rules except nonadjacent vs. late repetition.

Adults' learning of repetition-based structure also appears to be constrained by position (Endress, Scholl, & Mehler, 2005): Adults discriminated seven-syllable sequences from sequences of new items based on differences in internal vs. edge repetitions (e.g., ABCDDEF vs. ABCDEFF), but could only generalize when given edge repetitions. In summary, repetition structures in edge positions appear to be reliably learned by both infants and adults, but learnability of internal repetitions remains unknown for infants because the structures tested in previous structural learning studies only involved repetitions located at either the initial or final edge of the sequence (AAB, ABA, and ABB).

In Experiment 1, we asked if 11- and 14-month-olds can detect, extract, and generalize a "repetition anywhere" rule (i.e., ABBC, AABC, ABCC). If infants detect repetition of items during a learning phase, they may subsequently recognize repetitions of new items (in new sequences), which we take as evidence for generalization. However, it may be that consistent position information is a key part of repetition learning at this age, as it appears to be for adults; in this case, variability in the position of the repetition might pose difficulty in its identification and recall.

In Experiment 2, we used a sequence with two possible underlying patterns to examine how consistency and salience contribute to sequence learning. In this experiment, we asked whether 11 and 14-month-olds could generalize a "medial repetition" rule when its position is consistent in sequence, but not at an edge. In Experiment 3, using the same sequence types as Experiment 2, we asked whether 11-month-olds could identify a nonadjacent dependency occurring at initial and final edge positions that may be more salient than the medial position of the repetition.

In all experiments we used an intermodal presentation method in which looming shapes were accompanied by spoken syllables, a method known to facilitate rule learning, relative to visual or auditory only, in 5- and 7-month-olds (Frank, Slemmer, Marcus, & Johnson, 2009; Thiessen, 2012). Because infants as young as newborns look longer at randomly-ordered shape sequences vs. sequences with statistical structure (Addyman & Mareschal, 2013; Bulf, Johnson, & Valenza, 2011; Kirkham, Slemmer, & Johnson,

2002), we reasoned that longer looking at novel vs. familiar sequences (i.e., random vs. structured, respectively) in the current studies would reflect learning and/or generalization of structural and/or statistical structure during habituation.

Experiment 1

Method

Participants Twenty 11-month-olds ($M_{\text{age}} = 11.25$ months; $SD = .297$; 8 girls) and 20 14-month-olds ($M_{\text{age}} = 14.20$ months; $SD = .313$; 9 girls) participated. An additional ten 11-month-olds were tested but excluded for failure to habituate (7), fussiness (2), or preterm birth (1). An additional twelve 14-month-olds were tested but excluded for failure to habituate (8) or fussiness (4).

Materials and Apparatus Visual stimuli consisted of 18 colored shapes (see Figure 1). Auditory stimuli consisted of an inventory of 18 spoken syllables produced with a speech synthesizer and identical to those used in Marcus et al. (1999) (e.g., bah, dee, doo, gei, jai, jah, kei, poh).



Figure 1: Schematic depiction of example habituation and test sequences for Experiment 1.

Shape-syllable pairings were determined randomly (see Figure 1). Sequences were assembled from a randomly chosen set of nine (out of the total 18) shape-syllable combinations (hereafter called "items" for simplicity), so that three items composed each four-item sequence by repeating one of the three items, either the first, second, or third (determined randomly) to yield a repetition in the initial, medial, or final position. Items always appeared in the same order within each habituation sequence.

Procedure Stimuli were presented using Macromedia Director on a Macintosh computer and a 53 cm color screen. In a separate room, the experimenter used closed-circuit video to view the infant and record his or her looking times during the experiment; the experimenter was blind to what was being presented on the screen. Before each habituation trial, a visual attention-getter appeared in the center of the screen to draw the infant's attention. Each shape was presented on a black background and increased in size from 4 cm to 24 cm high (2.4-14.6° visual angle) over a period of 667 ms; onset of syllables (M duration 338 ms) was coincident with the appearance of each shape. Thus each sequence was 2 s long, and sequences were separated by a 667 ms black screen. The 4-item sequences were randomly displayed one after another with no immediate repetition of any specific sequence. When the mean looking time over

four consecutive trials fell to less than 50% compared to the mean looking time for the first four habituation trials (i.e., habituation) infants viewed the test sequences.

At test, infants viewed “familiar” and “novel” four-item sequences drawn from the remaining nine in the total inventory that were not shown during habituation. Familiar test sequences followed the same constraints as those described previously for the habituation sequences. Novel test sequences were composed of the same nine items, presented in sequences of four, and random ordering of items with no constraints except no repeated items in any single sequence. Infants viewed six alternating familiar and novel trials presented in pairs (i.e., three test trial blocks), and viewing order was counterbalanced such that half the infants viewed a familiar trial first (followed by a novel trial) and half the infants viewed a novel trial first (followed by a familiar trial). Preliminary analyses examining sex differences in performance revealed no reliable effects in any of the experiments in this report (all $ps > .05$).

Results and Discussion

A 2 (age group) x 2 (trial type – novel or familiar) x 2 (order – novel or familiar first) x 3 (test trial block) mixed ANOVA on posthabituation looking times revealed a main effect of test trial block, $F(2, 72) = 4.79, p = .011, \eta^2_p = .12$, the result of a decline in looking across trials, and an age group x trial type interaction, $F(1, 36) = 8.03, p = .008, \eta^2_p = .182$. There were no other significant effects. Follow-up t -tests indicated that 11-month-olds did not look differently to novel and familiar test stimuli, $t(19) = -.916, p = .371, ns$, but 14-month-olds looked longer at novel vs. familiar sequences, $t(19) = 3.06, p = .006$. Thus 11-month-olds provided no evidence for learning a “repetition anywhere” rule, whereas 14-month-olds appeared to do so.

Experiment 2

Because 11-month-olds showed no evidence of learning repetitions in variable locations in Experiment 1, in Experiment 2 we examined the role of positional consistency in sequence learning. Here, we tested 11 and 14-month-olds’ learning of a *medial* repetition rule. Habituation sequences comprised two different patterns that could be extracted: a medial repetition rule (a changing identity in the AxxC pattern), and a nonadjacent dependency (between A and C in the AxxC pattern). Experiment 2 specifically tested rule learning, and as in Experiment 1, we reasoned that this learning would be reflected in longer looking during novel vs. familiar test trials.

Method

Participants Twenty 11-month-olds ($M_{age} = 11.15$ months, $SD = .34$; 14 girls) and twenty 14-month-olds ($M_{age} = 14.14$ months, $SD = .39$; 9 girls) participated in Experiment 2. Six additional 11-month-olds were tested but excluded for failure to habituate (4) or fussiness (2). Eight additional 14-month-olds were tested but excluded due to failure to habituate (2), technical error (2), or fussiness (4).

Materials and Apparatus The item stimuli and presentation apparatus were the same as in Experiment 1. Habituation sequences contained both a medial repetition and a nonadjacent dependency between the first and fourth items of the sequence (see Figure 2). Sequences were assembled from a randomly chosen set of ten from the inventory of 18. Again, three items composed each four-item sequence, but the second item was always repeated, instantiating a medial repetition rule. Four items were selected (from the ten) for first and fourth positions in two unique sequences, and three items were selected for the medial positions in each sequence (e.g., ABBC, DEEF, AGGC, DHHF, etc.). The two sequences were presented in alternation during habituation.

The test sequences were constructed such that familiar sequences tested the generalization of the medial repetition rule with new exemplars that did not use consistent shapes/syllables in the first and fourth positions across sequences. Familiar sequences were composed of items drawn from the entire shape inventory, with the constraints that the second item always repeated, and the first and fourth items in sequence could not be one of the four items that occupied those positions in the habituation sequences. The novel sequences followed the same constraints described in Experiment 1.

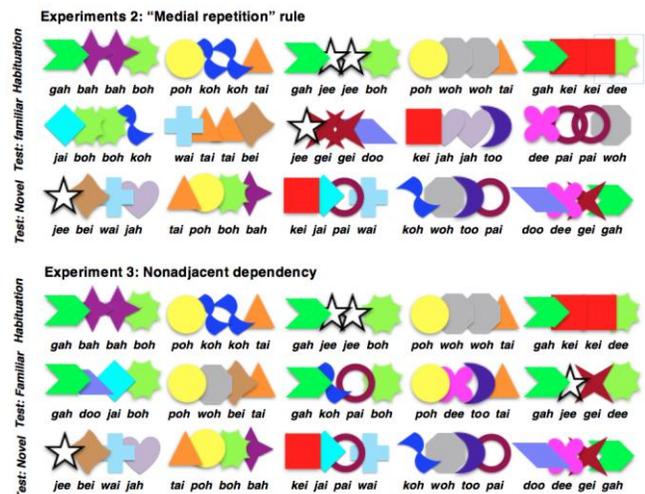


Figure 2: Schematic depiction of example habituation and test sequences for Experiments 2 and 3.

Procedure The procedure was the same as in Experiment 1. Infants were habituated to sequences that contained a medial repetition and a nonadjacent dependency (described above), and an infant-controlled habituation paradigm was used. At test, infants saw six trials that alternated between familiar (i.e., contained a medial repetition) and novel, viewing three of each trial type in total. Infants were randomly assigned to see either a familiar trial or a novel trial first.

Results and Discussion

A 2 (age group) x 2 (trial type) x 3 (test trial block) mixed ANOVA revealed a significant 3-way interaction, $F(2, 76) =$

5.71, $p = .022$, $\eta^2_p = .13$, which stemmed from a relatively precipitous decline in looking across novel trials by older infants but not younger infants. More importantly, there was an age \times trial type interaction, $F(1, 38) = 7.42$, $p = .010$, $\eta^2_p = .16$, due to differences in looking at novel and familiar test sequences. Although this analysis yielded a significant main effect of trial type, $F(1, 38) = 4.36$, $p = .044$, $\eta^2_p = .10$, this effect was driven by 14-month-olds' longer looking to novel test trials, as 11-month-olds looked equally to novel and familiar test trials. There was also a significant main effect of test trial block, $F(2, 76) = 25.11$, $p < .001$, $\eta^2_p = .40$, due to a decline in looking times across trials.

14 Month Olds A 2 (trial type) \times 2 (order) \times 3 (test trial block) mixed ANOVA revealed a significant main effect of trial type, $F(1, 18) = 8.20$, $p = .010$, $\eta^2_p = .313$, due to longer looking overall at the novel test sequence. There was also a reliable main effect of test trial block, $F(2, 36) = 22.37$, $p < .001$, $\eta^2_p = .55$, due to a decline in looking across trials. These main effects were qualified by a significant trial type \times test trial block interaction, $F(2, 36) = 10.70$, $p = .004$, $\eta^2_p = .37$ and a significant trial type \times order \times test trial block interaction, $F(2, 36) = 9.25$, $p = .007$, $\eta^2_p = .34$. There were no other significant effects. The two higher-order interactions were a result of longer looking during the first trial block toward the novel sequence by infants in the novel-first order, relative to infants in the familiar-first order, $t(18) = 2.29$, $p = .034$; comparisons across the second and third trial blocks were *ns*, $ps > .16$.

11 Month Olds A 2 (trial type) \times 2 (order) \times 3 (test trial block) mixed ANOVA yielded a main effect of test trial block, $F(2, 36) = 5.50$, $p = .031$, $\eta^2_p = .23$, the result of a decline in looking across trials, and a significant trial type \times order interaction, $F(1, 18) = 6.44$, $p = .021$, $\eta^2_p = .23$, due to a (nonsignificant) tendency for infants in both order conditions to look longer at the trial type that was presented first. There were no other significant effects; the trial type effect was *ns* at $p = .364$. These results support the conclusion that the 14-month-olds generalized the medial repetition rule, whereas the 11-month-olds did not.

Experiment 3

Because 11-month-olds failed to learn the “medial repetition” rule in Experiment 2, Experiment 3 addressed the possibility that learning structures in 4-item sequences is too difficult for 11-month-olds, perhaps due to limits in visual working memory. We used the same habituation sequences as Experiment 2 but instead tested for statistical learning (specifically, the nonadjacent dependency between A and C in the AxC pattern). As in previous experiments, we reasoned that this learning would be reflected in longer looking during novel vs. familiar test trials.

Method

Participants Twenty 11-month-olds ($M_{age} = 11.16$ months, $SD = .32$; 6 girls) participated in Experiment 3. An additional nine infants were tested but excluded due to failure to habituate (8) or fussiness (1).

Materials and Apparatus The item stimuli and presentation apparatus were the same as in Experiments 1 and 2.

Procedure The procedure was the same as in Experiment 2, except for the structure of the test sequences. The familiar test trials maintained the relation between the first and fourth items across two sequences, using the exact same first and fourth items from habituation, but had no repetitions (i.e., each familiar sequence was composed of four unique items; see Figure 2). The novel test trials followed constraints described previously.

Results and Discussion

A 2 (trial type) \times 2 (order) \times 3 (test trial block) mixed ANOVA yielded a reliable main effect of trial type, $F(1, 18) = 7.86$, $p = .012$, $\eta^2_p = .30$, due to longer looking overall at novel vs. familiar test sequences (see Figure 4). There was also a main effect of trial block, $F(2, 36) = 7.82$, $p = .012$, $\eta^2_p = .30$, due to a decline in looking across trials, and a significant interaction between trial block and trial type, $F(2, 36) = 9.63$, $p = .006$, $\eta^2_p = .35$. Infants looked more toward the novel sequence than the familiar in the first block, $t(19) = 2.75$, $p = .013$, and in the second block, $t(19) = 2.80$, $p = .011$, but not in the third block, $t(19) = .22$, *ns*. The overall presence of a novelty preference suggests that infants abstracted the nonadjacent dependency pattern during habituation.

We compared performance of the 11-month-olds in Experiments 2 and 3 with a 2 (experiment) \times 2 (trial type) \times 3 (test trial block) mixed ANOVA. This analysis revealed a significant main effect of test trial block, $F(2, 76) = 11.44$, $p < .001$, $\eta^2_p = .23$, due to a decline in looking times across trials. More importantly, there was an experiment \times trial type interaction, $F(1, 38) = 6.83$, $p = .013$, $\eta^2_p = .15$, due to differences in looking at novel and familiar test sequences, as noted earlier. Taken together, therefore, the results of Experiments 2 and 3 provide evidence that when both statistical and rule-bound information was available in habituation sequences, 11-month-olds detected the presence or absence of nonadjacent dependencies (ordinal positions of initial and final shapes in sequence; i.e., statistical information) across habituation and test, but not a medial repetition (i.e., rule-bound information).

General Discussion

In a departure from past studies showing that 11-month-olds learn an adjacent repetition rule when the repetition appears in the initial or final positions in sequence (Johnson et al., 2009), we discovered that 11-month-olds failed to learn this rule when the repetition appeared in any position (initial, medial, or final, Experiment 1), or when it was restricted to the medial position (Experiment 2). Fourteen-month-olds, however, appeared to learn repetition rules under both conditions. Finally, in Experiment 3, 11-month-olds succeeded in learning a nonadjacent dependency in sequences identical to those used in Experiment 2. We conclude that 11-month-old infants do not seem to

recognize repetitions when they appear in multiple positions in sequence, or in a consistent middle position, although items at edge positions in sequence appear to be distinctly salient. Similar findings for adults were reported by Endress et al. (2005, 2010).

Repetition and Position as Perceptual Primitives

Infant sequence learning is constrained by saliency and consistency of information well as general limits in attention and memory (Aslin & Newport, 2012, 2014). Some of these constraints are specific to modality (e.g., speech cues; Johnson & Jusczyk, 2001; Thiessen & Saffran, 2003) or the experimental setting (e.g., gaze or action cues; Baldwin, Andersson, Saffran, & Meyer, 2008), but others, such as repetition and position, are domain-general and may operate similarly across contexts. Evidence from the 11-month-olds in Experiment 2 is consistent with findings from adults, who generalized a repetition to new items when it appeared in final position, but not a medial position, in 7-item syllable sequences (Endress et al., 2005). These studies suggest that item position, most notably final position, is more salient than item repetition, relatively speaking; these findings are consistent with the well-documented serial position curve (Ebbinghaus, 1885) and the recency effect in memory (Baddeley & Hitch, 1974)

Because any particular set of items in a group potentially supports an infinite number of possible structures and generalizations thereof, a learner must determine the most likely pattern given a limited amount of experience with it. One way in which this problem may be constrained is by a “gradient of generalization:” the most consistent information across a distribution produces the best learning (Aslin & Newport, 2014; Gerken, 2006). In Experiments 2 and 3, information for medial repetition and nonadjacent dependency was available, yet 11-month-olds learned only the statistical information. Notably, 14-month-olds appeared to learn a repetition rule both when it was restricted to the medial position (Experiment 2) and when it was free to appear in initial, medial, or final position (Experiment 1), implying that important developments in structural learning consist of the “separation” of perceptual primitives such that they become less interdependent and perhaps more salient on their own.

Infant Sequence Learning in Context

Rule learning and generalization for shorter sequences can be observed in infants as young as 4 months (Dawson & Gerken, 2009). Rule learning in 5-month-olds from 3-item shape-syllable sequences was also reported, using a similar design (Frank et al., 2009). Studies that tested for statistical learning reported that 3-month-olds appeared to recognize violations of serial order in 3-item shape-sound sequences (Lewkowicz, 2008), and 5-month-olds segmented shape sequences from differences in transitional probability (Marcovitch & Lewkowicz, 2009; Slone & Johnson, 2015). (To our knowledge, there is no published evidence for rule or statistical learning in auditory or visual sequences prior to

4 months.) By 8 months, infants seem to use a “chunking” mechanism to segment shape sequences when tested for learning of “illusory” sequences or “embedded” units in streams of looming shapes (Slone & Johnson, 2016; cf. Endress & Mehler, 2009; Giroux & Rey, 2009).

Finally, consider the findings (from Experiments 2 and 3) that 11-month-olds extracted statistical patterns, but not rules, from identical sequences. In a previous test of multiple pattern learning, adults listened to speech streams that could be interpreted in terms of rules or statistical relations (Endress & Bonatti, 2007). With briefer listening times, participants learned the rules, but did not identify the statistical structure without substantially longer exposure durations. This result led to the claim that there is a fast-working mechanism for extracting rule-bound patterns, and a second slower mechanism that requires additional time to learn associations among items. Yet the infants we observed appeared to learn statistical relations, but not rules, during a relatively brief period of habituation. The reasons for this effect are unclear. Recently, 8-month-old infants were found to learn different statistical structures (transitional probabilities and “chunks” of items) as a function of exposure time (Slone & Johnson, 2016), and it may be that 11-month-olds would learn rules in the current stimulus set if they accumulated more looking times than allowed for by the infant-controlled habituation method. Nor is it clear from the current studies or the larger literature whether, in general, rule learning systems might come “on line” earlier during development than statistical learning systems, or vice-versa. These questions await future study.

Conclusions

Perceptual primitives may be best thought of as helping to support learning by attracting learners’ attention and memory resources to likely structures in the environment. Yet they do not seem to attract attention automatically, as do some sensory primitives such as motion or high contrast, nor are they automatically committed to memory, for either infants or adults. Rather, evidence to date suggests that infants at birth can discriminate certain rules and statistical patterns when compared to unstructured input, but learning and generalization of rules develop across the first year after birth and beyond. On this account, perceptual primitives such as repetition and position serve as building blocks upon which more complex structures can be built.

Acknowledgments

This research was supported by NIH Grant R01-HD73535. We thank the research staff at the NYU Infant Perception Lab and the UCLA Baby Lab, and our infant participants and their families.

References

Addyman, C., & Mareschal, D. (2013). Local redundancy governs infants’ spontaneous orienting to visual-temporal sequences. *Child Development, 84*, 1137-1144.

- Arciuli, J. (2017). The multi-component nature of statistical learning. *Philosophical Transactions of the Royal Society B*, 372, 20160058.
- Aslin, R. N., & Newport, E. L. (2012). Statistical learning: from acquiring specific items to forming general rules. *Current Directions in Psychological Science*, 21, 170-176.
- Aslin, R. N. & Newport, E. L. (2014). Distributional language learning: mechanisms and models of category formation. *Language Learning*, 64: Cognitive Neuroscience Supplement 2, 86–105.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, 8, 47-89.
- Baldwin, D., Andersson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action via statistical structure. *Cognition*, 106, 1382–1407.
- 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.
- Bulf, H., Johnson, S. P., and Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, 121, 127–132.
- Dawson, C., & Gerken, L. (2009). From domain-general to domain-sensitivity: 4-month-olds learn an abstract repetition rule in music that 7-month-olds do not. *Cognition*, 111, 378–382.
- Ebbinghaus, H. (1885). *Über das Gedächtnis (On memory)*. Leipzig: Duncker & Humblot.
- Endress, A., & Bonatti, L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition*, 105, 247-299.
- Endress, A. & Mehler, J. (2009). The surprising power of statistical learning: When fragment knowledge leads to false memories of unheard words. *Journal of Memory and Language*, 60, 351-367.
- Endress, A. D., Nespore, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences*, 13, 348–353.
- Endress, A. D., Scholl, B. J., & Mehler, J. (2005). The role of saliency in the extraction of algebraic rules. *Journal of Experimental Psychology: General*, 134, 406-419.
- Frank, M. C., Slemmer, J. A., Marcus, G. F., & Johnson, S. P. (2009). Information from multiple modalities helps five-month-olds learn abstract rules. *Developmental Science*, 12, 504–509.
- Gerken, L. (2006). Decisions, decisions: Infant language learning when multiple generalizations are possible. *Cognition*, 98, B67-B74.
- Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences (USA)*, 105, 14222-14227.
- Gervain, J., Berent, I., & Werker, J. F. (2012). Binding at birth: the newborn brain detects identity relations and sequential position in speech. *Journal of Cognitive Neuroscience*, 24, 564-574.
- Giroux, I. & Rey, A. (2009). Lexical and sublexical units in speech perception. *Cognitive Science*, 33, 260-272.
- Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, 44, 548–567.
- 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, 2–18.
- Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition*, 83, B35-B42.
- Krogh, L., Vlach, H. A., & Johnson, S. P. (2012). Statistical learning across development: Flexible yet constrained. *Frontiers in Psychology*, 3, 598.
- Marcovitch, S., & Lewkowicz, D. L. (2009). Sequence learning in infancy: The independent contributions of conditional probability and pair frequency information. *Developmental Science*, 12, 1020-1025.
- Marcus, G. F. (2000). *Pabiku* and *ga ti ga*: Two mechanisms children could use to learn about language and the world. *Current Directions in Psychological Science*, 9, 145-147.
- Marcus, G., Fernandes, K., & Johnson, S. (2007). Infant rule learning facilitated by speech. *Psychological Science*, 18, 387–391.
- Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283, 77–80.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926-1928.
- Saffran, J. R., Pollack, 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, 669–680.
- Slone, L. K. and Johnson, S. P. (2015). Infants' statistical learning: 2-and 5-month-olds' segmentation of continuous visual sequences. *Journal of Experimental Child Psychology*, 133, 47-56.
- Slone, L. K., & Johnson, S. P. (2016). *When learning goes beyond statistics: infants represent visual sequences in terms of chunks*. Manuscript submitted for publication.
- Thiessen, E. D. (2012). Effects of inter- and intra-modal redundancy on infants' rule learning. *Language Learning and Development*, 8, 197–214.
- Thiessen, E. D., Kronstein, A. T., and Hufnagle, D. G. (2013). The extraction and integration framework: a two-process account of statistical learning. *Psychological Bulletin*, 139, 792-814.
- Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology*, 39, 706–716