View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by The University of North Carolina at Greensboro

Sequence learning in infancy: the independent contributions of conditional probability and pair frequency information

By: Stuart Marcovitch and David J. Lewkowicz

Marcovitch, S., & Lewkowicz, D. J.(2009). Sequence learning in infancy: The independent contributions of conditional probability and pair frequency information. *Developmental Science, 12*, 1020-1025.

Made available courtesy of Wiley-Blackwell: http://www.wiley.com/

***** Note: Figures may be missing from this format of the document**
***** Note: The definitive version is available at www3.interscience.wiley.com**

Abstract
The ability to perceive sequences is fundamental to cognition. Previous studies have shown that infants can learn visual sequences as early as 2 months of age and it has been suggested that this ability is mediated by sensitivity to conditional probability information. Typically, conditional probability information has covaried with frequency information in these studies, raising the possibility that each type of information may have contributed independently to sequence learning. The current study explicitly investigated the independent contribution of each type of information. We habituated 2.5-, 4.5-, and 8.5-month-old infants to a sequence of looming visual shapes whose ordering was defined independently by specific conditional probability relations among pair elements and by the frequency of occurrence of such pairs. During test trials, we tested infants′ sensitivity to each type of information and found that both types of information independently influenced sequence learning by 4.5 months of age.

Article:
Introduction
Our environment consists of events that are often predictable. Predictability can be absolute (e.g. when mommy puts a bottle in my mouth, I will always get food) or it can be probabilistic (e.g. after feeding, mommy usually puts me down for a nap). Abstraction of the underlying structure of an event requires keeping track of adjacent sequential relations based on the redundancy found in the environment. This is a robust ability that appears very early in human development (Gomez & Gerken, 1999; Marcus, Vijayan, Bandi Rao & Vishton, 1999; Saffran, Johnson, Aslin & Newport, 1999) enabling infants to learn artificial grammars (Gomez, 2002; Gomez & Gerken, 1999; Gomez & Maye, 2005), to segment continuous speech utterances into distinct words (Saffran, Aslin & Newport, 1996), to learn visual feature combinations (Fiser & Aslin, 2002), and to perceive the sequential structure of a series of visual (Kirkham, Slemmer & Johnson, 2002), auditory (Saffran et al., 1999), and audiovisual (Lewkowicz, 2004, 2008) elements.

In the seminal study that demonstrated that infants were capable of abstracting sequential relations, Saffran et al. (1996) familiarized 8-month-old infants with a 2 minute speech stream consisting of four different three-syllable words (e.g. bidaku, padoti). The conditional probability of within-word syllables was always 1.0 (e.g. bi was always followed by da), whereas the conditional probability of between-word syllables was 0.33 (e.g. ku could be followed by the first syllable of any of the other three words). Across two studies, infants attended longer to non-words that violated the probability rules, prompting Saffran et al. to raise the possibility that sensitivity to conditional probabilities of sequential elements reflects the operation of an innately biased statistical learning mechanism. Subsequent modifications of the original task (Aslin, Saffran & Newport, 1998; Fiser & Aslin, 2002; Gomez, 2002; Gomez & Gerken, 1999; Gomez & Maye, 2005; Kirkham et al., 2002) have replicated Saffran et al.′s initial findings and, as a whole, these studies have demonstrated that sensitivity to sequential conditional probability relations provides the foundation for an early appearing, domain-general, information-acquisition mechanism.

Kirkham et al. (2002) provided the developmentally earliest example of this type of statistical learning by demonstrating that 2- to 8-month-old infants can learn the conditional probability relations inherent in a sequence of visually presented shapes. Infants were habituated to a series of looming colored shapes that were presented sequentially in randomly presented pairs. The conditional probabilities within pairs were determinate (i.e. $p = 1.0$; the first element always predicted the second element) while the conditional probabilities across pairs were indeterminate (i.e. $p = .33$; the last element of a pair did not predict the first element of the following pair). The question was whether infants could encode the conditional probability relations inherent in these sequences during habituation and then detect violations in these relations during test trials. The test trials indicated that the infants did detect the violations in the conditional probability relations. In a more recent spatiotemporal version of the task, Kirkham, Slemmer, Richardson and Johnson (2007) provided additional evidence of infant sensitivity to conditional probability information. Here, they reported differences in 8-month-olds' saccade latencies when they viewed predictable as opposed to non-predictable locations.

Although the Kirkham et al. (2002) study clearly showed that infants can learn sequential dependencies, it is unclear from that study as well from others like it whether sensitivity to conditional probability information is the sole determinant of successful learning. This is because in a typical statistical learning experiment the frequency of pair occurrence usually covaries with conditional probability information (e.g. if B always follows A, then AB will occur together more often than other pairs such as BC). As a result, infant learning of conditional probability relations may, in fact, reflect learning of frequency information and/or the learning of both types of information.

Recognizing this possibility, Aslin et al. (1998) conducted a study that assessed this issue directly in the auditory domain and provided compelling evidence that at least 8-month-old infants can learn conditional probability information independently of frequency information. In this study, infants were familiarized with four tri-syllabic nonsense words such that two of the words were presented 90 times while the other two words were presented 45 times. After the familiarization phase, the infants were presented with four different test items: two words and two non-words that were created by combining the last syllable of one word with the first two syllables of another word. Importantly, these test items were carefully chosen so that the words and non-words were heard an equal number of times in the familiarization phase. Results showed that infants responded more to non-words than to words, implying that 8-month-old infants can learn conditional probability information even when frequency information is held constant.

In another study, Fiser and Aslin (2002, Experiment 2) investigated 9-month-old infants' response to the spatial relations among distinct objects making up a complex visual scene. Infants were habituated to two base-pair items that were always presented in the same spatial configuration (e.g. circle directly above triangle) and a noise item (e.g. crescent) whose spatial relationship vis-a-vis the base pair varied from trial to trial. Fiser and Aslin manipulated the frequency of the base pairs such that infants saw critical non-base pairs as often as the low frequency base pairs. On test trials, infants looked reliably longer at base pairs over non-base pairs, which the authors interpreted as evidence that infants processed conditional probabilities when learning complex spatial relationships.

Despite the converging evidence that by 8 months of age infants can perceive and learn conditional probability relations when frequency information is held constant, it is still possible that these two types of information may interact in infant sequence learning. The currently available data do not provide independent information regarding the contribution of frequency information to sequence learning because frequency information was either confounded with conditional probability information (e.g. Kirkham et al., 2002; Saffran et al., 1996), or because it was perfectly controlled (Aslin et al., 1998; Aslin, Slemmer, Kirkham & Johnson, 2001; Fiser & Aslin, 2002). As a result, it is still possible that frequency information may interact with conditional probability information in infant sequence learning. Determining whether this is the case is critical because in their everyday life infants are usually exposed to both types of information. As a result, the purpose of the current study was to investigate this possibility directly by manipulating each type of information independently and then assessing its role in infant learning across an age range where previous studies have reported successful statistical learning (2 to 8 months). We used the Kirkham et al. (2002) visual learning task except that we varied the frequency of each pair's occurrence as well as the conditional probabilities within the pairs during the habituation phase. In this way, we provided infants with the opportunity to encode frequency and/or probability information. If infants encoded only conditional probability relations, then during test trials they should exhibit longer looking at pairs that violate those relations. Alternatively, if infants encoded only frequency information, they should exhibit longer looking at relatively infrequent or novel pairs. Finally, if they encoded both types of relations, then they should exhibit longer looking when either type of information changes.

Method

Participants

Ninety-six infants (48 girls) were included in the final sample: 32 2.5-month-olds (M = 11.3 weeks, Range: 9.6–12.9), 33 4.5-month-olds (M = 19.0 weeks, Range: 16.9–21.0), and 31 8.5-month-olds (M = 36.9 weeks, Range: 34.0–39.9). An additional 79 infants were tested but not included in the final sample because of failure to complete the experiment due to fussiness or disinterest (34 2.5-month-olds, 36 4.5-month-olds, nine 8.5-month-olds). Finally, the infants were tested in one of two university laboratories (n = 44 in Lab A and n = 52 in Lab B)[1] whose set up was identical except for the size of the monitor (see below).

Apparatus

Infants were seated in an infant seat 50 cm from a 15 inch (Lab A) or 17 inch (Lab B) computer monitor. Those infants who refused to sit in the infant seat were seated on their parent’s lap. Parents were not informed about the specific hypothesis under test and were instructed to sit as still as possible and not to talk nor interact with their infant.

The stimuli were the same six colored shapes used by Kirkham et al. (2002): yellow circle, pink diamond, blue cross, green triangle, turquoise square, and red octagon. The shapes were presented one at a time and loomed from either 2.7 to 16.5 cm (3.1–18.3', Lab A) or 3.0 to 19.5 cm (3.4–21.3', Lab B) in height within 1 s.

Procedure

We used an infant-controlled habituation/test procedure (i.e. the length of each trial, regardless of whether it was a habituation or a test trial, was controlled by the infant’s looking at the stimulus-display monitor). Whenever the infant looked at the monitor, the appropriate movie (i.e. trial) began and as soon as the infant either looked away from the monitor for more than 2 seconds or accumulated a total of 60 seconds of looking, the trial ended. To attract the infant’s attention back to the monitor following a look-away, we presented a beeping sound through a set of speakers positioned on each side of the monitor. As soon as the infant looked back at the monitor the sound was turned off and the movie showing the visual stimuli began to play. The habituation trials continued until the infant reached a habituation criterion which required that the total duration of looking during the last four habituation trials declined to 50% of the total duration of looking during the first four habituation trials. Once an infant reached the habituation criterion, the next trial constituted the start of the test phase.

Habituation phase

During the habituation phase, infants watched a continuous stream of looming shapes with specific constraints on their sequential ordering. 2 These constraints were designed to pit conditional probability information against frequency information. Table 1 a shows the shape pairs presented during the habituation phase, their frequency, and the conditional probabilities between the pair members. As can be seen in Table 1a, the conditional probability constraints were such that A could be followed by B or D equally often, whereas C was always followed by D, and E was always followed by F. In terms of overall frequency of occurrence, however, pairs A-B and A-D occurred four times as often as pairs C-D and E-F.

Test phase

Each infant received four test trials that were counterbalanced using a Latin Square design such that each test trial appeared equally often in each ordinal position. Each test trial consisted of the repeated presentation of a single pair of items (i.e. the first element followed by the second element) with a 1 s break between each pair presentation. Note that these test trials are analogous to the Saffran et al. (1996) test trials where a word was presented repeatedly, and different from the Kirkham et al. (2002) test trials where all six elements were presented either randomly or according to the probability rules. The test pairs presented here were: Test AB, Test CD, Test FE, and Test AF (see Table 1b).

The Test AB pair was used to determine whether infants attended primarily to frequency information. Because this was a high-frequency pair during the habituation phase, we expected infants to look more at the other pairs than at this pair if they attended primarily to frequency during the habituation phase. The Test CD pair was used to determine whether infants attended primarily to conditional probability information. Because this pair had a determinate conditional probability during the habituation phase, we expected infants to look more at the other pairs than at this pair if they attended primarily to the conditional probability information during the habituation phase. The Test FE pair was a control pair that was almost never shown in the habituation phase and had an indeterminate conditional probability.

Table 1 *(a) Frequency of occurrence and conditional probabilities of the shape pairs in the habituation phase, (b) Test pairs and their relevant characteristics*

(a) Shape pair	Frequency	Conditional probability
AB	40%	.50
AD	40%	.50
CD	10%	1.00
EF	10%	1.00

(b) Test pair	Relevant characteristics
TEST AB	High frequency and indeterminate transitional probability
TEST CD	Low frequency and determinate transitional probability
TEST FE	Almost never shown and indeterminate transitional probability
TEST AF	Never shown, but contains frequent element 'A'

Note: For approximately half of the infants: A-Pink Diamond; B-Green Triangle; C-Turquoise Square; D-Blue Cross; E-Yellow Circle; F-Red Octagon. For the remaining infants: A-Turquoise Square; B-Yellow Circle; C-Green Triangle; D-Red Octagon; E-Pink Diamond; F-Blue Cross.

Finally, the Test AF pair was presented to determine whether infants were responding to individual shapes instead of shape pairs. One potential limitation of the current design is that infants' preferences for the test pairs may depend upon the frequency of presentation of the individual elements as opposed to the frequency of the pair of elements co-occurring or the conditional probabilities within the pairs. To assess whether individual element frequency affected looking times, we compared the looking duration to the Test AF pair with looking to the Test FE pair. Both of these test pairs were presented infrequently, if at all, in the habituation phase and subsequently both test pairs had an indeterminate conditional probability within the pair (i.e. the first element did not necessarily predict the second element). The primary difference between the two pairs is that the Test AF pair contains the element A which occurred very frequently (about 40% of all elements presented) in the habituation phase, while the Test FE pair contains the element E which occurred infrequently (about 5% of all elements presented). Note that both test pairs also contain the infrequently presented element F (also about 5% of all elements presented). Thus, we reasoned that if infants were tracking the individual-element frequency, then they would look less at the Test AF pair than at the Test FE pair. Alternatively, if infants were not tracking individual-element frequency but rather were tracking pair frequency or transitional probability relations, then they should not exhibit differential looking at the Test AF and Test FE pairs.

Results

As is often the case with looking-time data from infants, the data were not normally distributed. As a result, we employed the nonparametric Wilcox on matched pairs test for data analysis. In our first analysis, we asked whether infants were tracking element frequency information. Thus, we compared looking duration to the Test AF pair with looking duration to the Test FE pair. Results indicated that infants did not exhibit differential looking at these two test pairs, at 2.5 months, $z = 1.25$, $p = .21$; 4.5 months, $z = 1.56$, $p = .12$; or 8.5 months, $z = 1.00$, $p = .32$. These results indicate that infants did not track single-element frequency information and, thus, justify examining the data to determine whether infants were tracking pair conditional probability information and/or pair frequency information.

To do so, we created a composite score by computing the mean duration of looking at the FE and AF test pairs (note that both of these test pairs had at a low frequency of occurrence and were characterized by indeterminate conditional probability relations during the habituation phase). To assess whether infants attended to pair frequency information, we compared the AB test pair (high pair frequency) to the composite score (low pair frequency). Longer looking times to the composite score were exhibited by 17 (53%) of the 2.5-month-olds, $z = 0.37$, $p = .71$, 22 (67%) of the 4.5-month-olds, $z = 3.03$, $p = .002$, and 25 (81%) of the 8.5-month-olds, $z = 3.25$, $p = .001$. These findings indicate that the 4.5- and 8.5-month-old infants tracked pair frequency information but that the 2.5-month-old infants did not.

To assess whether infants attended to conditional probability information, we compared the CD test pair (determinate conditional probability) to the composite score (indeterminate conditional probability). Longer looking times to the composite score were exhibited by 17 (53%) of the 2.5-month-olds, $z = 0.09$, $p = .93$, 25 (76%) of the 4.5-month-olds, $z = 2.08$, $p = .04$, and 20 (65%) of the 8.5-month-olds, $z = 2.25$, $p = .02$. These findings indicate that the 4.5- and the 8.5-month-old infants tracked the conditional probability information but that the 2.5 month-old infants did not.

Discussion

Previous statistical learning studies with infants between 2 and 8 months of age (Aslin et al., 1998; Fiser & Aslin, 2002; Kirkham et al., 2002; Saffran et al., 1999) have found that infants can recognize violations in the conditional probability linking the members of a pair of sequentially presented elements. Importantly, however, these studies did not investigate the role of pair frequency, either because pair frequency information was confounded with conditional probability information, or pair frequency information was controlled. Given that in their daily lives infants are exposed to sequences whose temporal structure is defined concurrently by frequency and conditional probability information, it is important to assess the contribution of each type of information to infant sequence learning. Thus, in the present study we investigated sequence learning in 2.5-, 4.5-, and 8.5-month-old infants. Results showed that starting at 4.5 months of age, infants are able to independently track pair frequency and conditional probability information.

These findings replicate and extend Kirkham et al.′s (2002) results from their 5-month-old and 8-month-old age groups. In that study, infants exhibited longer looking during test trials where visual stimuli were presented in a random sequence than in test trials where the visual stimuli bore the same pair conditional probability relations as in the habituation phase. This finding suggests that infants were tracking pair frequency information, conditional probability information, or both. The current study expands on Kirkham et al.′s (2002) findings by showing that infants are, in fact, sensitive to both frequency and conditional probability information and that they are able to track such information during a short-term learning task. Interestingly, and in contrast to Kirkham et al.′s (2002) study, here we did not find evidence of learning of either type of information in 2.5-month-old infants. The reason for this inconsistency is presently unknown, although it is possible that the methodological differences between the two studies contributed to the difference in outcome. Specifically, the test trials in the current study compared infants’ response to pattern differences, whereas Kirkham et al. compared infants’ response to sequential patterns versus random sequences. [3] Thus, the task may have been more challenging in the current study. One interesting possibility is that the reasons that the 2.5-month-old infants in the Kirkham et al. experiment exhibited evidence of successful learning is because frequency and conditional probability information covaried and that they actually benefited from this redundancy of information. Whatever the ultimate reason for the difference in outcome might be, additional research is needed to shed further light on it.

Sensitivity to pair frequency independent of conditional probability is critical in daily life because patterned or sequential sensory input often contains items with low conditional probabilities that need to be learned. For example, infants must learn that both ‘baby’ and ‘basin’ are words despite the fact that the conditional probability within the word is indeterminate (i.e. less than 1). Sensitivity to frequency above and beyond statistics can serve to facilitate learning by allowing the infant to take advantage of multiple sources of information. The present study demonstrates that by 4.5 months of age, infants are sensitive to both frequency and conditional probability information.

Given that our findings are the first to demonstrate that frequency and conditional probability information can play independent roles in infant sequence learning starting at 4.5 months of age, and given the absence of other data on this issue, it is currently only possible to speculate on the development of perceptual sensitivity to each type of information. One possible developmental pattern might be that sensitivity to frequency information emerges first in development and that it then helps infants discover the conditional probability relations that are typically inherent in sequences. Such a developmental pattern is consistent with well-established findings showing that adults exhibit an implicit sensitivity to frequency information and that this sensitivity is invariant with age (Hasher & Zacks, 1984). Thus, one reasonable scenario that is generated by our hypothetical developmental pattern is that infants must first be repeatedly exposed to a sequence of elements and that through such exposure they get to discover the inherent probabilistic structure of such a sequence. Furthermore, when younger infants are first exposed to a repeating pattern, they only respond to the inherent frequency information and ignore the probabilistic information. In contrast, older infants are capable of extracting the conditional probability information, in part, because the differential frequency information selectively directs their attention to those features that are also probabilistically related and perhaps allows them to consider the deeper nature of the information available (see Marcovitch & Zelazo, 2009; Zelazo, 2004, for similar ideas on other infant tasks). Needless to say, we recognize the speculative nature of this hypothesis and that it now needs to be tested.

Although the current data provided evidence that element frequency was not a significant predictor of looking duration in the test trials in this study, it must almost certainly play a role in processing pair frequency and conditional probability information. Here, we used a standard habituation procedure in which habituation trials are administered until looking declines by 50%. It is assumed that this provides infants with sufficient time to encode the relevant pairings. It is possible that other habituation procedures that utilize a more lenient habituation criterion, or familiarization techniques that do not guarantee that the infant has habituated, may be more prone to the influence of the individual elements. Consequently, it is essential that control trials (such as the AF test pair in the current study) be administered to ensure that infants are not responding to the individual elements irrespective of their intended pairings.

Finally, the current findings raise some interesting questions for future studies. First, they raise questions regarding the domain-generality of 4.5-month-olds’ sensitivity to pair frequency and conditional probability information. For example, do infants of this age exhibit the ability to track frequency and conditional probability information in the auditory modality and is this ability affected by multisensory redundancy (i.e. by concurrent visual and auditory sequence specification)? The possibility that multisensory redundancy might affect learning and discrimination of each type of sequential information is supported by findings that infants’ response to serial order information is facilitated by multisensory redundancy (Lewkowicz, 2004). Finally, our findings from the youngest infants are in direct contrast to previous speculations that infants may be born with the ability to compute conditional probabilities and the ability to encode pair frequency. The results of testing these hypotheses will yield important insights into the perceptual foundations of sequence learning, in particular, and pattern detection, in general, and, ultimately into the development of fundamental higher-level cognitive abilities that make the learning and understanding of language and event structure possible.

Notes:

1 There was no effect of laboratory, nor did this variable interact with any of the independent variables and thus will not be discussed further.

2 Due to limitations in the programming software, we were unable to randomize shape pairing and order for every infant. Instead, randomization was achieved in the following manner. A string of 360 consecutive pairs were generated randomly following the frequencies outlined in Table 1 (i.e. 144 each of AB and AD and 36 each of CD and EF). Then the string was sectioned into 12 habituation trials consisting of 30 pairs. Further randomization was achieved by presenting the 12 habituation trials in a different random order for each infant.

3 We thank an anonymous reviewer for suggesting this possibility.

References

Aslin, R.N., Saffran, J.R., & Newport, E.L. (1998). Computation of conditional probability statistics by 8-month-old infants. Psychological Science, 9, 321–324.

Aslin, R.N., Slemmer, J.A., Kirkham, N.Z., & Johnson, S.P. (2001, April). Statistical learning of visual shape sequences. Paper presented at the meeting of the Society for Research in Child Development, Minneapolis, MN.

Fiser, J., & Aslin, R.N. (2002). Statistical learning of new visual feature combinations by infants. Proceedings of the National Academy of Sciences, 99, 15822–15826.

Gomez, R.L. (2002). Variability and detection of invariant structure. Psychological Science, 13, 431–436.

Gomez, R.L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. Cognition, 70, 109–135.

Gomez, R., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. Infancy, 7, 183– 206.

Hasher, L., & Zacks, R.T. (1984). Automatic processing of fundamental information: the case of frequency of occurrence. American Psychologist, 39, 1372–1388.

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.

Kirkham, N.Z., Slemmer, J.A., Richardson, D.C., & Johnson, S.P. (2007). Location, location, location: development of spatiotemporal sequence learning in infancy. Child Development, 78, 1559–1571.

Lewkowicz, D.J. (2004). Perception of serial order in infants. Developmental Science, 7, 175–184.

Lewkowicz, D.J. (2008). Perception of dynamic and static audiovisual sequences in 3- and 4-month-old infants. Child Development, 79, 1538–1554.

Marcovitch, S., & Zelazo, P.D. (2009). A hierarchical competing systems model of the emergence and early development of executive function. Developmental Science, 12, 1–18.

Marcus, G.F., Vijayan, S., Bandi Rao, S., & 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., Johnson, E.K., Aslin, R.N., & Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. Cognition, 70, 27–52.

Zelazo, P.D. (2004). The development of conscious control in childhood. Trends in Cognitive Sciences, 8, 12–17.

Acknowledgements

We thank Marcia Dabbene and Erinn Frances Beck for their assistance. Some of the data for this study were first collected while both authors were at the New York State Institute for Basic Research in Developmental Disabilities, Staten Island, NY, and the remainder were collected while the second author was at Florida Atlantic University. This work was supported by the New York State Office of Mental Retardation and Developmental Disabilities and by NICHD grant R01 HD35849 awarded to DJL.