



ELSEVIER

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

Visual statistical learning in the newborn infant

Hermann Bulf^a, Scott P. Johnson^b, Eloisa Valenza^{c,d,*}^a Department of Psychology, Piazza Ateneo Nuovo 1, 20126 Milano, Italy^b Department of Psychology, UCLA, 3291 Franz Hall, Los Angeles, CA 90095-1563, United States^c Department of Psicologia dello Sviluppo e della Socializzazione, Università degli Studi di Padova, Via Venezia 8, 35131 Padova, Italy^d Interdepartmental Center for Cognitive Science (CISC), Università degli Studi di Padova, Via Venezia 8, 35131 Padova, Italy

ARTICLE INFO

Article history:

Received 12 November 2010

Revised 10 June 2011

Accepted 19 June 2011

Available online 13 July 2011

Keywords:

Statistical learning

Newborns

ABSTRACT

Statistical learning – implicit learning of statistical regularities within sensory input – is a way of acquiring structure within continuous sensory environments. Statistics computation, initially shown to be involved in word segmentation, has been demonstrated to be a general mechanism that operates across domains, across time and space, and across species. Recently, statistical learning has been reported to be present even at birth when newborns were tested with a speech stream. The aim of the present study was to extend this finding, by investigating whether newborns' ability to extract statistics operates in multiple modalities, as found for older infants and adults. Using the habituation procedure, two experiments were carried out in which visual sequences were presented. Results demonstrate that statistical learning is a general mechanism that extracts statistics across domain since the onset of sensory experience. Intriguingly, present data reveal that newborn learner's limited cognitive resources constrain the functioning of statistical learning, narrowing the range of what can be learned.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Our sensory environments are full of regularities distributed in space and time. The speed and accuracy with which an organism extracts environmental regularities are important for its adaptation. Thus, a central question concerns how infants learn so much in so little time, without explicit instruction.

One learning mechanism available to infants is the detection of statistical relations among elements in the sensory array. Saffran, Aslin, and Newport (1996) reported evidence for statistical learning in 8-month-olds, who appeared to use subtle statistical relations for word segmentation in a continuous stream of artificial speech. During 2 min exposure to a structured sequence, infants learned

that some syllable groupings were more likely to co-occur than others, providing evidence for a statistical computation mechanism that may contribute to early language acquisition by segmenting the speech stream into units.

Statistical learning exists broadly across species and sensory modalities. Animal species learn statistically structured speech streams (Toro & Trobalón, 2005) and infants parse streams of musical tones based on statistical probabilities (Saffran, Johnson, Aslin, & Newport, 1999). Kirkham, Slemmer, and Johnson (2002) documented 2–8-month-olds' ability to detect statistical information in sequences of discrete, looming shapes, with no reliable age difference in performance. By 8 months, infants compute statistics in spatiotemporal visual sequences (Kirkham, Slemmer, Richardson, & Johnson, 2007), and by 9 months, infants encode the underlying spatial statistical structure of multiple-element scenes (Fiser & Aslin, 2002).

Overall, these results are consistent with a domain-general statistical learning device that is available early and operates across modalities, across time and space, and across species, suggesting that statistical learning

* Corresponding author at: Department of Psicologia dello Sviluppo e della Socializzazione, Università degli Studi di Padova, Via Venezia 8, 35131 Padova, Italy. Tel.: +39 049 8276583; fax: +39 049 8276500.

E-mail address: eloisa.valenza@unipd.it (E. Valenza).

URL: <http://dpss.psy.unipd.it/ita/people/valenza.htm> (E. Valenza).

might be a predisposed, general associative mechanism (Kirkham et al., 2002, 2007). This hypothesis was supported by a recent study investigating event-related brain amplitudes when sleeping newborns heard a stream of syllables containing statistical cues to word boundaries, providing the first evidence for statistical learning at birth, at least in the presence of linguistic materials (Teinonen, Fellman, Näätänen, Alku, & Huotilainen, 2009).

The first aim of the present paper was to investigate whether newborns extract statistical regularities in a variety of domains, as do older infants (Kirkham et al., 2002) and adults (Fiser & Aslin, 2001; Saffran et al., 1999), in this case from input with which they have no experience (visual input). Our second goal was to explore the possibility that statistical learning is influenced by the restricted cognitive capacities of newborn infants, or whether its functioning does not differ from statistical learning shown by older infants. In adults and older infants, the interaction between the learner's cognitive characteristics and the complexity of the input to-be-learned appears to be a critical factor in building a unitary representation of a sequence defined only by statistical information (Kirkham et al., 2007; Saffran & Thiessen, 2007). We hypothesized that limited processing ability (e.g., limitations in attentional or working memory capacities) shown by newborns might reduce statistical learning efficiency.

To address these aims, newborns were habituated to a probabilistic visual sequence of discrete geometric forms in a procedure adapted from Kirkham et al. (2002), who examined visual statistical learning in 2–8-month-olds. After habituation, newborns viewed a random sequence free of statistical cues. If the ability to compute visual statistical relations over time is present and available at birth, we expected newborns to discriminate structured from random visual sequences following the learning phase.

To manipulate task complexity, 48 newborns were randomly assigned to one of two conditions in which the number of elements that gave rise to the learning sequence was manipulated: a *high-demand condition* (HDC) or a *low-demand condition* (LDC, Fig. 1).

In the HDC we tested newborns' discrimination of a sequence of six looming discrete shapes structured by transitional probabilities from the same pattern of shapes presented in a random order. The learning sequence was composed of the same number of stimuli previously presented to 2–8-month-olds by Kirkham et al. (2002). In the LDC, a less demanding task was created; the number of stimuli that gave rise to the statistically structured sequence was reduced from 6 to 4. We reasoned that this manipulation may better match the newborn baby's limited cognitive resources, though preserving the essential features of any statistical learning design.

2. Method

2.1. Subjects

Forty-eight 1–3-day-old infants (M age = 38 h, SD = 17 h, range = 16–86 h) were tested using an infant-controlled habituation technique. Nine additional newborns were tested but excluded from the final sample because of position bias (looking more than 80% in one direction, $n = 3$) or change in state ($n = 6$). Infants in both experiments had a normal delivery, birth weight between 2470 and 4050 g, and an Apgar score of 9 or 10 at 5 min. Newborns were tested only if awake and alert, and after the parents had provided informed consent.

2.2. Stimuli

In the HDC, stimuli were six black and white shapes presented one at a time in a continuous stream, with no break or delay between shapes (Fig. 1). Each shape was presented for 2 s and loomed from 2 to 10 cm in height (about 3.8–19°). The habituation sequence consisted of three randomly-ordered shape pairs (pair 1: square followed by X shape; pair 2: circle followed by hexagon; pair 3: triangle followed by octagon). The initial member of a shape pair predicted the next member, and the next

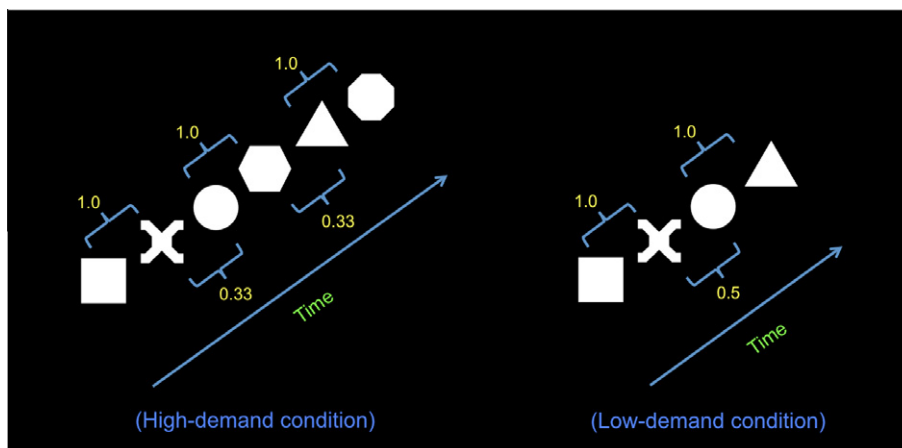


Fig. 1. Schematic representation of the learning sequence presented during habituation in the two experiments, showing transitional probabilities defining pairs (higher within pairs than between pairs).

stimulus after a pair was constrained to be the initial member of one of the three allowable pairs. There were no pauses between pairings, thus only transitional probabilities defined between-stimulus boundaries, with a transitional probability within pairs of 1.0 and a transitional probability between pairs of 0.33. The computer randomized shape pairing for each infant. During the test phase, the habituation sequence was presented with a novel sequence produced by randomly ordering the same six shapes, with the constraint that there were never two identical shapes in a row. In the LDC, the number of shapes was reduced from 6 to 4. Consequently, in this condition the habituation sequence consisted of two randomly-ordered shape pairs (pair 1: square followed by X shape; pair 2: triangle followed by circle), with a transitional probability within pairs of 1.0 and a transitional probability between pairs of 0.5. As in the HDC, during the test phase the familiar (structured) sequence was presented alongside a novel (random) sequence.

2.3. Apparatus and procedure

Stimuli were presented on a 30-in. Apple Cinema monitor 30 cm from the newborn. Plain white curtains were drawn on both sides to prevent distraction. A central flickering attention getter was used to attract the newborns' attention at the start of the habituation and test phases. Above the monitor, a video camera recorded the newborn to monitor looking behavior on-line and for off-line coding.

The newborn sat on an experimenter's lap in front of the monitor. This experimenter was naïve to hypothesis and stimuli, and fixed his/her gaze on a camera on the ceiling throughout the session. Testing began with the onset of the central flickering light. When the newborn's gaze was properly aligned with the attention getter, the sequence of trials was started by a second experimenter who watched the newborn on a monitor and pressed a key on the computer keyboard which turned off the attention get-

ter, activated onset of the stimuli, and recorded looking times. The stimulus was presented simultaneously on both the left and the right side during the habituation phase, and the amount of looking time was recorded irrespective of the side. A look-away criterion of 2 s determined the end of each fixation. To ensure that this criterion was strictly respected, the software automatically compacted two consecutive fixations that were not separated by a time interval of at least 2 s. The habituation phase was terminated when, from the fourth trial, looking times across any three consecutive trials were 50% or less than the total of the first three. At this point the stimuli were automatically turned off and the central flickering attention getter was turned on. When the newborn's gaze was realigned to the center of the screen, the test phase began. Each newborn was given two paired presentations of the familiar and novel test stimuli in which the position of the stimuli was reversed; initial left-right order of presentation was counterbalanced. The attention getter flickered between the first and the second presentation but did not flicker while the test stimuli were shown. A presentation lasted until each stimulus had been fixated on at least once and 20 s looking time accumulated. Looking times during the test phase were subsequently coded by a different observer unaware of the stimuli presented. Mean reliability between coders, calculated on 33% of the test phases, was $r(14) = 0.97$, $p < 0.001$, $N = 16$ (Pearson correlation).

3. Results

All newborns reached the habituation criterion. In the HDC, the average time to habituate was (mean \pm SD) 66.93 ± 33.95 s. In the LDC, the average time to habituate was (mean \pm SD) 64.71 ± 21.28 s. No statistical difference was found in the average habituation time between the two conditions, $t_{46} = 0.64$, $p = 0.53$, independent samples (all t statistics reported here were two-tailed).

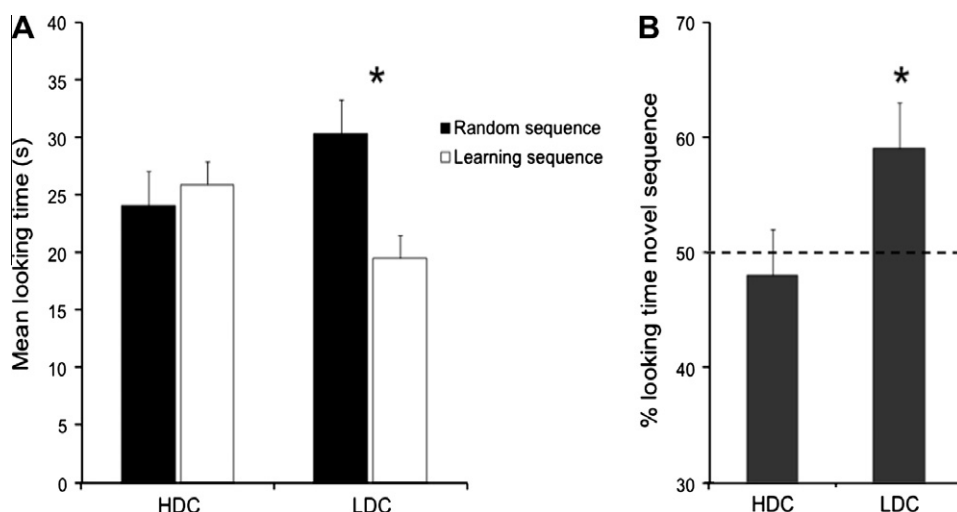


Fig. 2. Results of the two experiments. (A) Mean looking time after habituation. Newborn babies looked longer at the random (novel) sequence than at the learning (familiar) sequence only under conditions of fewer pairs (LDC). In the more demanding condition (HDC), newborns looked equally at the two test sequences. (B) Percentage of time spent looking at the novel sequence during the test phase. Dashed line indicates chance level. * $p < 0.05$.

Newborns' posthabituation visual preference was probed for the familiar vs the novel sequence. Fig. 2A shows that, during the test phase, only newborns habituated to the LDC condition consistently looked longer at the novel sequence than at the familiar one. A comparison of looking times to the novel (random) and to the familiar (learning) sequences was conducted with a two (sequence: novel vs familiar) \times two (condition: HDC vs LDC) mixed analysis of variance (ANOVA). No effects of sequence, $F_{1,46} = 2.21$, $p = 0.14$, or condition, $F_{1,46} = 0.003$, $p = 0.95$, were obtained. There was a significant interaction between sequence and condition, $F_{1,46} = 4.4$, $p = 0.042$. The interaction between sequence and condition was due to significant longer looking at the novel sequence (mean \pm SD = 30.26 \pm 16.31 s) than at the familiar sequence (mean \pm SD = 19.49 \pm 8.56 s) by newborns habituated to the low-demand sequence, $t_{23} = 2.39$, $p = 0.026$, paired samples. In contrast, newborns habituated to high-demand sequence looked equally at the novel sequence (mean \pm SD = 24.05 \pm 11.44 s) and at the familiar sequence (mean \pm SD = 25.88 \pm 9.69 s), $t_{23} = 0.46$, $p = 0.65$, paired samples. There were reliable difference in mean preference scores for both the novel (LDC = 30.26 s vs HDC = 24.05 s, $t_{46} = 2.072$, $p = 0.044$) and the familiar sequences (LDC = 19.49 s vs HDC = 25.88 s, $t_{46} = 2.4$, $p = 0.019$).

To test for recognition of the habituation sequence, a novelty preference score (percentage) was computed for each newborn (Fig. 2B). A t test for independent samples comparing the mean novelty preference in the two HDC and LDC conditions reached statistical significance, $t_{46} = 2.072$, $p = 0.044$. In the HDC the mean preference for the novel (random) sequence (mean \pm SD = 48 \pm 18%) was not different than chance ($t_{23} = 0.67$, $p = 0.51$), but it was significantly higher than chance in the LDC (mean \pm SD = 59 \pm 20%, $t_{23} = 2.165$, $p = 0.041$).

This outcome demonstrates that newborns discriminated between the structured habituation sequence and the random sequence only in the LDC, but not in the HDC, meaning that visual statistical learning appears to be functional even at birth, but heavily constrained by newborns' limited attentional resources.

4. General discussion

In the present study, newborns were shown to learn the statistical structure of a continuous sequence of discrete shapes. After habituation to a structured visual sequence defined solely by transitional probabilities between stimuli, newborns showed a reliable preference for a random novel sequence in which the order of shapes violated the transitional probability that defined grouping in the habituation stream. In tandem with the Teinonen et al. (2009) study on newborns' detection of statistics in linguistic sequences, our results provide the first evidence for the existence of a learning device that computes statistics across sensory domains at birth.

Crucially, we observed visual statistical learning only when a minimal number of stimuli (four) composed the learning sequence (LDC), but when the learning task was presumably more demanding (six shapes), performance

dropped to chance (HDC). This finding provides evidence that visual statistical learning at birth is constrained by the newborn's limited cognitive resources, preventing identification of relevant visual information necessary to detect statistical structure when a task previously used to assess 2-month-olds' learning skills (Kirkham et al., 2002) is presented to a newborn. This is the first demonstration of the effects of resource limitations to neonates' learning statistical information. Visual statistical learning is available at birth but highly constrained; cognitive limitations narrow the range of what can be learned. Notably, selecting and encoding visual information changes dramatically in the first 2 months after birth, as demonstrated in many cognitive abilities such as object unity perception (Johnson & Aslin, 1995; Johnson, Davidow, Hall-Haro, & Frank, 2008; Johnson, Slemmer, & Amso, 2004; Valenza & Bulf, 2007; Valenza, Leo, Gava, & Simion, 2006), and motion detection (Atkinson & Braddick, 1989; Wattam-Bell, 1992, 1996).

Why was newborns' ability to detect statistics observed only when the number of shapes that composed the learning sequence was reduced from 6 to 4? We consider three possibilities. First, a learning sequence composed of a minimal number of shapes (LDC) can be easier to process because the amount of information to be represented and stored in visual short-term memory is reduced. Neural network simulations have shown optimal learning when constrained by severe memory limitations, restricting the range of data they are exposed to during initial learning (Elman et al., 1996). Memory limitations act as a filter on the input, and focus learning on just a subset of the data. Seen in this light, limitations on memory capacity assume a positive character: "the importance of starting small" (Elman, 1993). Second, the difference in performance between the HDC and the LDC may stem from the frequency of occurrence of each shape pair. On this account, since no statistical difference was found in the average habituation time between the HDC and the LDC, newborns may have had more chances to discover the statistical structure of the visual sequence shown during habituation in the LDC. Third, selective attention may have played a role in newborns' performance. On this account, detection of statistical regularities was possible only when a reduced number of elements composed the learning sequence (LDC), facilitating identification of relevant visual information. Although statistical learning may proceed incidentally in some circumstances (Saffran, Newport, Aslin, Tunick, & Barrueco, 1997), there is evidence from adults that statistical learning is compromised when engagement of attention to a different task is required (Baker, Olson, & Behrmann, 2004; Toro, Sinnett, & Soto-Faraco, 2005; Turk-Browne, Jungé, & Scholl, 2005). The more proficient infants become at information processing during development, therefore, the more likely it is they will detect and utilize information relevant to cognitive tasks.

As argued by Saffran and Thiessen (2007), the nature of the input influences learning outcomes, and the same mechanism can yield different results as a function of prior knowledge (Saffran, 2001), input structure (Saffran, Reeck, Niehbur, & Wilson, 2005), input familiarity (Gebhart, Newport, & Aslin, 2009), and the learner's age (Saffran & Griepentrog, 2001) or species (Newport, Hauser, Spaepen,

& Aslin, 2004). For example, infants and adults parse a stream of tone sequences from statistical information provided by pitch cues, but this ability improves between infancy and adulthood (Saffran & Griepentrog, 2001) and is affected by knowledge of tonal structure (Saffran, 2003), providing evidence that auditory statistical learning depends on the developmental and/or experiential state of the listener. In the case of visual statistical learning, as infants develop, the conditions under which they reveal statistical learning likewise increase in complexity (cf. Kirkham et al., 2007).

It is not yet clear what information newborns utilized to segment the visual sequence into discrete units. Two candidate sources of information examined in previous research are the transitional probabilities that defined shape pairs, and frequencies of occurrence of the pairs (Aslin, Saffran, & Newport, 1998). The difference between a frequency balanced sequence (i.e., a transitional probability within pairs of 0.5 and a transitional probability between pairs of 0.25) and the LDC sequence (i.e., a transitional probability within pairs of 1.0 and a transitional probability between pairs of 0.5) may be discriminable by newborns, but it seems more plausible that the greater difference in probabilities characteristic of the latter would support learning more effectively than the former. Still, this remains an empirical question.

Experiments on learning in neonates have a long history and have tended to examine classical and operant conditioning (Blass, Ganchrow, & Steiner, 1984; Kaye & Bower, 1994; Siqueland, 1968), habituation (Slater, Morison, & Rose, 1984), or association learning (Salter, Quinn, Brown, & Hayes, 1999). The contribution of the present experiments comes from demonstrations of statistical learning without prior exposure to the types of materials composing the sequences infants acquired, and limits in this learning. We discussed as well possible explanations of these limits. To our knowledge, there have been no previous attempts to examine processing or resource limits in neonatal learning of this type.

Our experiments explore behavior in a naive system that has no knowledge of prior or posterior probabilities, and this is a challenge to any statistical learning approach that takes structure as given. Our findings therefore provide clear constraints for theories of statistical learning as a principal foundation for cognitive development (Bates & Elman, 1996), including recently popular Bayesian models, for which the principal developmental mechanism consists of queries to a hypothesis space arising in part from the statistics of an individual's experience with relevant events (e.g., Tenenbaum, Kemp, Griffiths, & Goodman, 2011). The Bayesian learner is an ideal (or nearly so) observer, coming to cognitive tasks with this structured hypothesis space and a prior probability distribution (likelihoods) defined over hypotheses. Observations provide a means of evaluating available hypotheses and selecting those that are determined to have the highest probability, thus updating the current state of knowledge, which can take different forms in different domains.

To compute the priors, a system requires the structures over which the distribution will be calculated (Altmann, 2010). Any comprehensive theory of learning must posit

the developmental sources of these structures, and any comprehensive theory of human learning must take into account the abilities and limitations of newborns. Formal learning models hold promise for understanding probabilistic inferences that yield increasingly abstract knowledge structures in experienced observers, yet any such model is necessarily constrained by the precise means of acquisition, storage, and retrieval of information, without which any hypothesis-space based reasoning system could not function. The Bayesian approach has a characteristic problem of circularity with respect to the *origins* of knowledge. Prior probabilities and likelihoods are *products* of abstract, symbolic systems of knowledge that support inductive reasoning, and they also *contribute* to these knowledge systems. But how does it get off the ground in the first place? How does the information get into the system? How are different kinds of information compared? How is behavior influenced by the state of current knowledge representation of the system? Tenenbaum et al. (2011) suggested that the most formidable challenge to Bayesian models may be to understand how structured symbolic knowledge can be represented in neural circuits. We suggest a second challenge: accounting for acquisition of structured symbolic knowledge in a naive system. It is this challenge that our experiments, and others that explore learning and its limits in newborn infants, can help to meet.

Acknowledgements

The authors are deeply indebted to Dr. B. Dalla Barba and the nursing staff at the Pediatric clinic of the University of Padua for their collaboration. Special thanks are due to the babies who took part in the study and to their parents. All experimental procedures have been licensed by the Responsible Office of the Dipartimento di Pediatria, Università degli Studi di Padova. This research project was supported by a grant from Università degli Studi di Padova, Ricerca Scientifica, fondi quota ex 60% (prot 60A17-5487/09), and by NIH Grants R01-HD48733 and R01-HD40432.

References

- Altmann, G. T. M. (2010). Why emergentist accounts of cognition are more theoretically constraining than structured probability accounts: Comment on Griffiths et al. and McClelland et al. *Trends in Cognitive Sciences*, 14, 340.
- 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.
- Atkinson, J., & Braddick, O. (1989). Development of basic visual function. In A. Slater & G. Brenner (Eds.), *Infant development* (pp. 7–41). Hove, UK: Lawrence Erlbaum Associates Ltd.
- Baker, C., Olson, C., & Behrmann, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological Science*, 15, 460–466.
- Bates, E., & Elman, J. (1996). Learning rediscovered. *Science*, 274, 1849–1850.
- Blass, E. M., Ganchrow, J. R., & Steiner, J. E. (1984). Classical conditioning in newborn humans 2–48 hours of age. *Infant Behavior & Development*, 7, 223–235.
- Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, 48, 71–99.
- Elman, J. L., Bates, E. A., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innate: A connectionist perspective on development*. Cambridge, MA: MIT Press.

- Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher order spatial structures from visual scenes. *Psychological Science*, *12*, 499–504.
- Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences, United States of America*, *99*, 15822–15826.
- Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and non-adjacent dependencies among non-linguistic sounds. *Psychonomic Bulletin & Review*, *16*, 486–490.
- Johnson, S. P., & Aslin, R. N. (1995). Perception of object unity in 2-month-old infants. *Developmental Psychology*, *31*, 739–745.
- Johnson, S. P., Davidow, J., Hall-Haro, C., & Frank, M. C. (2008). Development of perceptual completion originates in information acquisition. *Developmental Psychology*, *44*, 1214–1224.
- Johnson, S. P., Slemmer, J. A., & Amso, D. (2004). Where infants look determines how they see: Eye movements and object perception performance in 3-month-olds. *Infancy*, *6*, 185–201.
- Kaye, K. L., & Bower, T. G. R. (1994). Learning and intermodal transfer of information in newborns. *Psychological Science*, *5*, 286–288.
- Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence of 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.
- Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004). Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology*, *49*, 85–117.
- Saffran, J. R. (2001). Words in a sea of sounds: The output of statistical learning. *Cognition*, *81*, 149–169.
- Saffran, J. R. (2003). Absolute pitch in infancy and adulthood: The role of tonal structure. *Developmental Science*, *6*, 37–45.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.
- Saffran, J. R., & Griepentrog, G. J. (2001). Absolute pitch in infant auditory learning: Evidence for developmental reorganization. *Developmental Psychology*, *37*, 74–85.
- 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.
- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, *8*, 101–105.
- Saffran, J. R., Reeck, K., Niehbur, A., & Wilson, D. P. (2005). Changing the tune: Absolute and relative pitch processing by adults and infants. *Developmental Science*, *8*, 1–7.
- Saffran, J. R., & Thiessen, E. D. (2007). Domain-general learning capacities. In E. Hoff & M. Shatz (Eds.), *Handbook of language development* (pp. 68–86). Cambridge: Blackwell.
- Salter, A., Quinn, P. C., Brown, E., & Hayes, R. (1999). Intermodal perception at birth: Intersensory redundancy guides newborn infants' learning of arbitrary auditory–visual pairings. *Developmental Science*, *2*, 333–338.
- Siqueland, E. R. (1968). Reinforcement patterns and extinction in human newborns. *Journal of Experimental Child Psychology*, *6*, 431–442.
- Slater, A., Morison, V., & Rose, D. (1984). Habituation in the newborn. *Infant Behavior & Development*, *7*, 183–200.
- Teinonen, T., Fellman, V., Nääätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience*, *10*, 21.
- Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, *331*, 1279–1285.
- Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, *97*, B25–B34.
- Toro, J. M., & Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception & Psychophysics*, *67*, 867–875.
- Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*, 552–564.
- Valenza, E., & Bulf, H. (2007). The role of kinetic information in newborns' perception of illusory contours. *Developmental Science*, *10*, 492–501.
- Valenza, E., Leo, I., Gava, L., & Simion, F. (2006). Perceptual completion in newborn human infants. *Child Development*, *77*, 1810–1821.
- Wattam-Bell, J. (1992). The development of maximum displacement limits for discrimination of motion direction in infancy. *Vision Research*, *32*, 621–630.
- Wattam-Bell, J. (1996). Motion processing in one-month-old infants: Preferential looking experiments. *Vision Research*, *36*, 1671–1677.