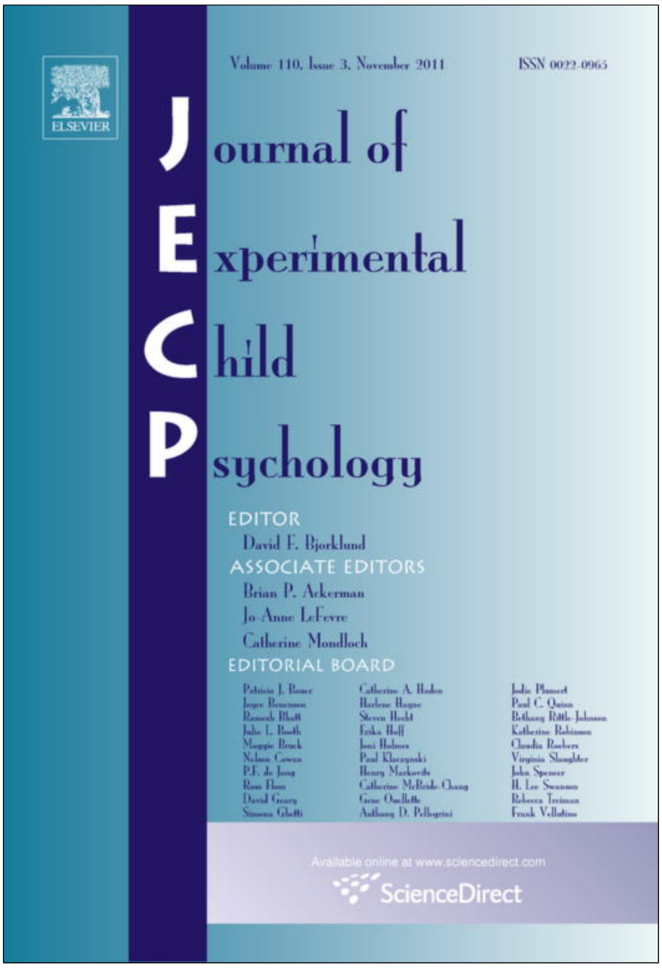
Volume 110, Issue 3, November 2011
ISSN 0022-0965
ELSEVIER
Journal of Experimental Child Psychology
EDITOR
David F. Bjorklund
ASSOCIATE EDITORS
Brian P. Ackerman
Jo-Anne LeFevre
Catherine Mondloch
EDITORIAL BOARD
Available online at www.sciencedirect.com
ScienceDirect

Journal of Experimental Child Psychology 110 (2011) 408–421

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

# Learning to learn: From within-modality to cross-modality transfer during infancy

Julie M. Hupp [a,*], Vladimir M. Sloutsky [b]

[a] Department of Psychology, The Ohio State University at Newark, Newark, OH 43055, USA
[b] Department of Psychology and Center for Cognitive Science, The Ohio State University, Columbus, OH 43210, USA

ARTICLE INFO

*Article history:*
Received 16 January 2011
Revised 6 May 2011
Available online 12 June 2011

*Keywords:*
Generalization
Infancy
Language
Learning
Transfer
Cross-modal processing

ABSTRACT

One critical aspect of learning is the ability to apply learned knowledge to new situations. This ability to transfer is often limited, and its development is not well understood. The current research investigated the development of transfer between 8 and 16 months of age. In Experiment 1, 8- and 16-month-olds (who were established to have a preference to the beginning of a visual sequence) were trained to attend to the end of a sequence. They were then tested on novel visual sequences. Results indicated transfer of learning, with both groups changing baseline preferences as a result of training. In Experiment 2, participants were trained to attend to the end of a visual sequence and were then tested on an auditory sequence. Unlike Experiment 1, only older participants exhibited transfer of learning by changing baseline preferences. These findings suggest that the generalization of learning becomes broader with development, with transfer across modalities developing later than transfer within a modality.


## Introduction

A fundamental aspect of learning is the application of learned knowledge to new situations, or transfer, which hinges on the ability to recognize commonalities between what was previously learned and what is presented in a new situation. There is evidence that this ability appears during infancy, although it is not known how much overlap is required between the two situations for infants to exhibit evidence of transfer. For example, in one study (Marcus, Vijayan, Rao, & Vishton, 1999), infants were familiarized to sequences of syllables representing an ABA pattern (*ga-ti-ga*) and were

* Corresponding author. Fax: +1 740 366 5047.
*E-mail address:* hupp.34@osu.edu (J.M. Hupp).

0022-0965/$ - see front matter 
doi:10.1016/j.jecp.2011.05.002

tested on either new ABA sequences (*wo-fe-wo*) or new ABB sequences (*wo-fe-fe*). Infants looked less when the new ABA sequence was presented than when an ABB sequence was presented, suggesting that something more general than *ga-ti-ga* was learned. Perhaps they learned an abstract rule ABA that was "unbound to the particulars of the training stimuli" (Frank, Slemmer, Marcus, & Johnson, 2009, p. 504), and this more general representation enabled them to detect the commonality between *ga-ti-ga* and *wo-fe-wo*.

However, it remains unclear how abstract an infant's representation is, how bound it is to the particulars of training stimuli, and how the ability to form abstract representations comes about. One argument is that infants have a system in place for learning about the world—a system enabling infants to represent abstract variables in an algebraic manner (Marcus et al., 1999). If this is the case, then when presented with *ga-ti-ti*, they represent *ga-ti-ti* as a set of abstract variables ABB, where A and B can stand for anything. Consequently, participants can recognize ABB even when a given instantiation of the variables is different from the learned instantiation. However, if the mind is truly algebraic, then patterns learned with syllables should be readily recognized under any instantiation. Therefore, a commonality between *ga-ti-ga* and another ABA sequence (rabbit–duck–rabbit) should be readily established because both instantiate ABA. To our knowledge, this has never been demonstrated. Furthermore, given that even later in development transfer is often difficult to achieve (e.g., DeLoache (1991); see also Barr (2010) and Detterman (1993), for reviews), there is likely a limit to how far infants would transfer the syllable pattern of ABA. Then, even under the most optimistic construal, representation is relatively specific to the training set. Although the specificity of the initial representation is unknown, the ability to form increasingly abstract representations is likely to improve with development and learning (for related arguments, see Elman et al. (1996), Goldstone and Landy (2010), Kloos and Sloutsky (2008), Morris and Sloutsky (1998), Sloutsky (2010), and Smith, Colunga, and Yoshida (2010)). For example, Son, Smith, and Goldstone (2008) argued that abstraction and generalization are indicative of differences between immature and mature learners; the latter seem to know the right similarities over which to generalize past experiences (see also Hayne, Barr, and Herbert (2003)).

There is much evidence exemplifying this point with respect to word learning. For example, Woodward, Markman, and Fitzsimmons (1994) demonstrated that 13-month-olds applied newly learned words only to objects that were used during learning, whereas 18-month-olds (who are more experienced word learners) exhibited evidence of generalization; they also applied the learned words to similar objects. Furthermore, Smith (2003) demonstrated that previous experience with category learning is a predictor of whether 20-month-olds recognize an abstract caricature of a concrete object. These studies indicate that breadth of generalization (and transfer of learning) increases in the course of development, suggesting that abstraction itself might be a product of development and learning.

The current study tested this hypothesis outside of word learning. It examined the ability of infants to generalize learning of sequence information within and across modalities. Several studies have demonstrated this ability in adults. For example, Altmann, Dienes, and Goode (1995) found that adults can transfer learning of artificial grammar across modalities, and more recently Hupp, Sloutsky, and Culicover (2009) found that adults can learn to attend to a particular component of a temporal sequence and transfer this learning across modalities.

There are reasons to believe that the ability to transfer learning across modalities is a product of development. First, unless learning and transfer involve an amodal relation that could be expressed in either modality (e.g., rhythm or tempo), transfer across modalities may require some experience in integrating cross-modal information. At the same time, there is evidence that younger participants may miss this experience (Robinson & Sloutsky, 2007a, 2010; Sloutsky & Robinson, 2008). Second, transfer across modalities may require a representation that is relatively independent of a particular modality and, thus, is more abstract (cf. Barnett and Ceci (2002), for related arguments).

If the ability to generalize or transfer information across modalities is a product of development, then older infants (who have greater experience in integrating cross-modal information in the course of word learning than younger infants) should be more likely to form a more abstract representation of sequence information than younger prelinguistic infants. Accordingly, they may generalize learned knowledge more broadly (i.e., across modalities) than younger prelinguistic infants. To examine these possibilities, we conducted four experiments examining transfer of learning within and across modalities. Given that sequence learning is more likely to be achieved in the auditory modality than in the visual modality (Conway & Christiansen, 2005; Robinson & Sloutsky, 2007b), we trained participants on visual sequences and examined their transfer to new visual or auditory sequences. In Experiment 1, we trained 8- and 16-month-olds to attend to a particular part of a visual sequence and then tested their ability to generalize this learning to new visual sequences. In Experiment 2, we repeated training on visual sequences but tested generalization to auditory sequences.

## Experiment 1A: visual baseline

The goal of Experiment 1A was to establish infants' baseline preferences within the visual modality. Are participants more likely to detect changes in the beginning or at the end of a visual sequence? To answer this question, infants were familiarized to a target visual sequence. Then they were tested with modified familiarization sequences, where an element was added to either the beginning or the end of the familiarization sequence.

### *Method*

#### *Participants*

In all reported experiments (1A, 1B, 2A, and 2B), parents' names were collected from local birth announcements and contact information was obtained through local directories. All infants were full-term (i.e., >2500 g birth weight) with no auditory or visual deficits, as reported by parents. The majority of infants were Caucasian. No infant participated in more than one experiment. Participants were 8-month-olds ($n$ = 22, 13 boys and 9 girls, mean age = 245.86 days, $SD$ = 7.32) and 16-month-olds ($n$ = 25, 17 boys and 8 girls, mean age = 490.28 days, $SD$ = 7.95). An additional 16 infants were excluded for failing to complete the experiment due to excessive fussiness ($n$ = 15) or parental interference ($n$ = 1).

#### *Apparatus*

Infants were seated on parents' laps approximately 100 cm from a 152 $\times$ 127-cm projection screen. An NEC GT2150 LCD projector was mounted to the ceiling approximately 30 cm behind infants. Two Boston Acoustics 380 speakers were 76 cm apart and mounted at infants' eye level on either side of the screen. A Dell Dimension 8200 computer with *Presentation* software (Neurobehavioral Systems, 2003) was used to present stimuli and record visual fixations.

Visual fixations were recorded online by trained research assistants in all reported experiments. Two video streams (i.e., stream of stimulus presentation and stream of infants' fixations) were projected onto two Dell flat panel monitors in an adjacent room, and a Sony DCR-PC120 camcorder recorded both video streams. This split-screen recording was used to establish interrater reliability. Then 18.7% ($n$ = 40) of the infants were coded offline by two additional coders. The average reliability between online and offline coders across all reported experiments was $r$ = .93.

#### *Materials*

The stimuli consisted of object sequence videos created using *Macromedia Flash* (Macromedia Studio MX, 2002). Infants were familiarized to a visually presented two-part object sequence (green triangle → cube) that flashed sequentially from left to right for 1 s each. The videos were presented centrally at the bottom of the screen (see Table 1 for stimuli used in all reported experiments). Each video presented the sequence four times and lasted 9 s. The video was repeated 10 times for a total of 40 sequence presentations and 90 s of familiarization. The two test sequences of interest consisted of a modified familiarization sequence; for the "Pre+" sequence, a component (orange musical note) was added to the beginning/left of the familiarization sequence, whereas for the "Post+" sequence, the component was added to the end/right of the familiarization sequence. There was also a novel sequence (red semicircle → light bulb) used to determine novelty preference, which was the third and final test trial. Each of the three test sequences was also presented four times and lasted 9 s. After each video in the experiment, the screen was black for 750 ms.

**Table 1**
Stimuli used across experiments.

| | Training | Familiarization | Pre+ | Post+ | Novel |
|---|---|---|---|---|---|
| Experiment 1A<br>Visual Baseline | NA |  |  |  |  |
| Experiment 1B<br>Visual Train-Visual Test |  | | | | |
| Experiment 2A<br>Auditory Baseline | NA | *Ki-Tu* | *FO-Ki-Tu* | *Ki-Tu-FO* | *Gu-Ti-Hi* |
| Experiment 2B<br>Visual Train-Auditory Test |  | | | | |

*Note.* For interpretation of the references to color in the text regarding objects shown in this table, the reader is referred to the web version of this article. NA, not applicable.

### *Design and procedure*

Infants were familiarized with a spatiotemporal visual sequence (object sequence flashing left to right). At test, two test trials of interest were presented in a random order: Pre+ (object added to beginning/left of the familiarized sequence) and Post+ (object added to end/right of the familiarized sequence). For 8-month-olds, 11 received the Pre+ trial first and 11 received the Post+ trial first. For 16-month-olds, 13 received the Pre+ trial first and 12 received the Post+ trial first. The final test trial was the novel sequence.

### *Results and discussion*

As shown in Fig. 1A, participants reduced looking by 33.90% between the first three and last three familiarization trials, paired-samples $t(46) = 7.74$, one-tailed $p < .001$, $d = 1.13$. An analysis of variance (ANOVA) on average looking across all familiarization trials did not reveal a significant main effect of age. In addition, participants demonstrated novelty preference by exhibiting longer looking to the novel sequence test trial than the average of their own last three familiarization trials, paired-samples $t(46) = 5.80$, one-tailed $p < .001$, $d = 0.85$.

The primary analyses focused on looking to the first test trial of interest (Pre+ or Post+) compared with the mean of each infant's last three familiarization trials. For this and all other experiments reported here, mean looking times to the last three familiarization trials, Pre+ items, and Post+ items are presented in Table 2. Individual difference scores were calculated for each participant by subtracting his or her average looking time to the last three familiarization trials from his or her looking time to the first test trial of interest (Pre+ or Post+). For each of the baseline conditions (Experiments 1A and 2A), the difference scores were compared across age groups and trial types. For each of the transfer conditions (Experiments 1B and 2B), the difference scores were compared with the appropriate baseline condition. For this experiment, these individual difference scores (i.e., recovery scores) were subjected to a 2 (Age: 8 or 16 months) × 2 (Trial Type: Pre+ or Post+) between-participants ANOVA. There

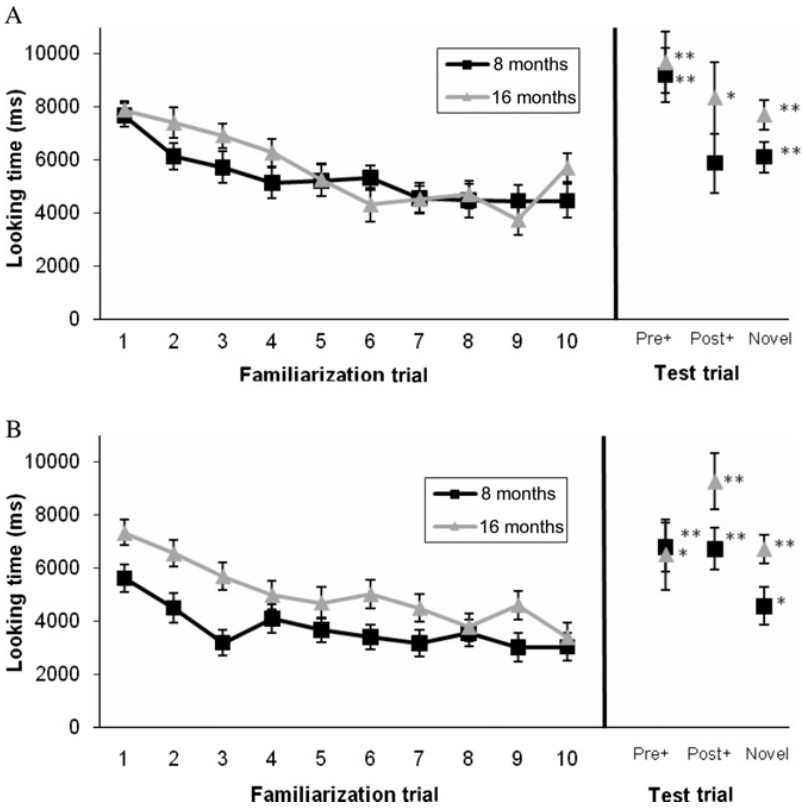


**Fig. 1.** (A) Visual baseline looking times across trials for 8- and 16-month-olds. (B) Looking times in the visual modality after training to attend to the end of a visual sequence for 8- and 16-month-olds. Pre+ and Post+ trials were between participants. Bars represent standard errors. Differences from the last three familiarization trials are indicated with one-tailed *p* values: *$p < .05$; **$p < .01$.

**Table 2**
Mean looking times and standard deviations by age and trial type for each experiment.

| | Trial type | | |
|---|---|---|---|
| | Familiarization | Pre+ | Post+ |
| *Experiment 1A: visual baseline* | | | |
| 8-month-olds | 4467 (2160) | 9205 (3362)** | 5885 (3716) |
| 16-month-olds | 4714 (1926) | 9678 (4118)** | 8347 (4673)* |
| *Experiment 1B: visual training–visual test* | | | |
| 8-month-olds | 3209 (2280) | 6790 (3603)** | 6721 (2958)** |
| 16-month-olds | 3931 (2401) | 6521 (4971)* | 9279 (4148)** |
| *Experiment 2A: auditory baseline* | | | |
| 8-month-olds | 3737 (1458) | 7157 (3001)* | 9405 (1510)** |
| 16-month-olds | 3492 (2229) | 5667 (3994) | 5094 (3754) |
| *Experiment 2B: visual training–auditory test* | | | |
| 8-month-olds | 5127 (2041) | 6687 (2877)* | 6208 (3569) |
| 16-month-olds | 5543 (2130) | 8265 (2947)** | 9516 (2368)** |

*Note.* Looking times are in milliseconds (ms). Standard deviations are in parentheses. Differences from familiarization are indicated with one-tailed *p* values:
* $p \leqslant .05$
** $p \leqslant .01$.

was only a main effect of trial type, with Pre+ sequences having higher difference scores ($M$ = 5061.39 ms, $SD$ = 3647.17) than Post+ sequences ($M$ = 2320.54 ms, $SD$ = 3822.22), $F(1, 43) = 6.77$, $p = .01$, $d = 0.70$. Further examination of the data in Table 2 indicates that although the interaction did not reach significance, there was a numerical difference between the age groups; whereas the younger group exhibited exceedingly strong Pre+ preference, this preference was noticeably weaker in the older group. Based on these data, it was concluded that none of the groups had Post+ preference, whereas there was evidence for Pre+ preference.

The main goal of this experiment was to establish whether infants had a baseline preference for either the beginning or the end of a visual sequence. The results indicate that infants had a preference for the beginning, as evidenced by greater recovery to Pre+ items than to Post+ items; they were more likely to notice when an item was added to the beginning of a visual sequence than when it was added to the end of a visual sequence. Given that participants were unlikely to exhibit Post+ preference (they exhibited Pre+ preference), we trained participants to pay attention to the end of the visual sequence and tested their transfer of learning to novel visual sequences. Experiment 1B examined this within-modality transfer.

## Experiment 1B: visual training–visual test

Infants were trained to attend to the end of a visual sequence and were then presented with the same familiarization and test visual sequences described in Experiment 1A. Note that the visual stimuli used in training differed from those used in familiarization and testing. Transfer of learning would be inferred from participants exhibiting a difference from baseline preferences as a result of training.

### *Method*

#### *Participants*

Participants were 8-month-olds ($n = 29$, 15 boys and 14 girls, mean age = 250.14 days, $SD = 8.58$) and 16-month-olds ($n = 29$, 16 boys and 13 girls, mean age = 494.34 days, $SD = 9.71$). An additional 59 infants[1] were excluded for failure to complete the experiment due to excessive fussiness ($n = 45$), failure to look during test trials ($n = 11$), or failure to look to at least 50% of familiarization trials ($n = 3$).

#### *Materials, design, and procedure*

The experiment had three phases: training, familiarization, and testing. All infants were trained to attend to the end of a visual sequence. Training was immediately followed by a familiarization session and testing phase in the visual modality.

The training video consisted of two blue objects (sun → heart) that flashed sequentially for 1 s each in a left/right configuration centered at the bottom of the screen. During each presentation, the last component was made salient by doing something exciting (e.g., jumping, sliding back and forth, tilting, expanding). The saliency of the final component was used to direct infants' attention to the end of the visual sequence. The video presented this sequence four times and lasted for 9 s. The video was repeated continuously 10 times for a total of 40 sequence presentations and 90 s of training. After each video in the experiment, the screen was black for 750 ms.

Then participants were presented with familiarization followed by testing, which were the same as in Experiment 1A. Infants were familiarized to a visual sequence (green triangle → cube) and then tested with visual sequences that had a component (orange musical note) added to the beginning (Pre+) or the end (Post+) as well as a novel sequence (red semicircle → light bulb) as the final test trial to measure novelty preference. For 8-month-olds, 15 received the Pre+ trial first and 14 received the Post+ trial first. For 16-month-olds, 14 received the Pre+ trial first and 15 received the Post+ trial first. Their recovery scores (looking to Pre+ or Post+ compared with the average of their own last three familiarization trials) were compared with the infants' recovery scores in the no-training visual baseline established in Experiment 1A.

[1] The exclusion rate was especially high in this experiment given the increase in session length once the training phase was added and the fact that the stimuli were very similar across the entire session.

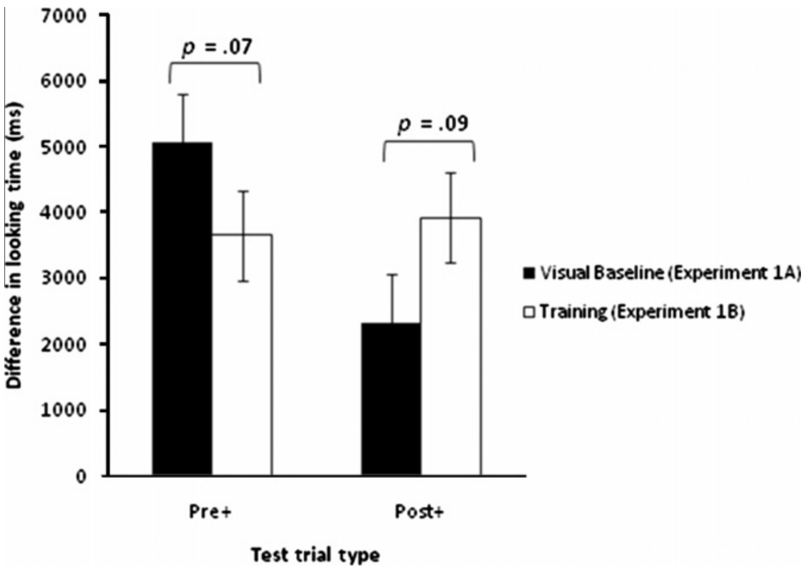


**Fig. 2.** Recovery times (test trial minus infant's average of last three familiarization trials) in the visual modality without training and after training to attend to the end of a visual sequence collapsed across ages. Bars represent standard errors. Differences between baseline recovery scores and recovery scores after training are indicated with one-tailed *p* values.

### *Results and discussion*

As shown in Fig. 1B, participants reduced looking by 33.77% between the first three and last three familiarization trials, paired-samples $t(57) = 7.01$, one-tailed $p < .001$, $d = 0.92$. An ANOVA on average looking across all familiarization trials revealed a significant main effect of age, $F(1, 57) = 7.80$, $p < .01$, partial $\eta^2 = .12$, with 16-month-olds looking on familiarization trials longer than 8-month-olds. Participants demonstrated novelty preference by exhibiting longer looking to the novel sequence than the average of their own last three familiarization trials, paired-samples $t(57) = 4.66$, one-tailed $p < .001$, $d = 0.61$. As shown in Table 2 and Fig. 1B, participants exhibited significant recovery to the Pre+ and Post+ items; that is, their looking to these items exceeded their own average looking on the last three familiarization items.

Because the procedure in Experiment 1B differed from that in Experiment 1A in that the training phase was added to Experiment 1B, we could not directly compare looking to the Pre+ and Post+ items across the experiments given that the absolute values of looking times are likely to differ considerably with the addition of a training phase. Instead, we compared the amount of recovery to the Pre+ and Post+ sequences in relation to their own last set of familiarization sequences (i.e., difference or recovery scores). These recovery scores to Pre+ and Post+ test trials after training (i.e., in Experiment 1B) were compared with those without training (i.e., in Experiment 1A).

A 2 (Age: 8 or 16 months) × 2 (Condition: baseline or training) × 2 (Trial Type: Pre+ or Post+) between-participants ANOVA on the difference scores revealed a significant Condition × Trial Type interaction, $F(1, 97) = 4.36$, $p < .05$, partial $\eta^2 = .04$. To tease this apart, each trial type was analyzed separately collapsed across age groups. As shown in Fig. 2, infants marginally increased the amount of looking to the test trial that had the last component changed (Post+), $F(1, 50) = 2.24$, one-tailed $p = .07$, $d = 0.41$, and decreased looking to the test trial that had the first component changed (Pre+), $F(1, 51) = 1.94$, one-tailed $p = .09$, $d = 0.38$.

Therefore, as a result of training, there was a change in the baseline preference. This is particularly noteworthy because infants were trained on one visual sequence and tested on a different visual sequence. These results suggest that infants of both age groups have the potential to learn to shift their attention to the end of the visual sequence (away from the beginning) and transfer this learning to a novel visual sequence. Although the ability to shift attention as a result of learning has been demonstrated before in older participants (e.g., Sloutsky & Fisher, 2008; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002), here we found evidence of attentional learning in younger infants. In addition, participants were able to transfer their learning from the studied sequence to a novel sequence. To examine how abstract their representation is, we conducted Experiment 2, in which we trained infants on a visual sequence and examined their transfer to an auditory sequence. To be able to determine a change in auditory sequence processing as a result of training, it was necessary to first determine baseline auditory sequence processing (Experiment 2A).

## Experiment 2A: auditory baseline

Infants were familiarized to a target auditory sequence and tested with sequences where components were added to either the beginning or the end of the sequence. The goal was to establish participants' baseline preferences within the auditory modality.

### *Method*

#### *Participants*

Participants were 8-month-olds ($n = 20$, 12 boys and 8 girls, mean age = 250.05 days, $SD = 6.00$) and 16-month-olds ($n = 21$, 12 boys and 9 girls, mean age = 491.43 days, $SD = 6.16$). An additional 19 infants were excluded for failure to complete the experiment due to excessive fussiness ($n = 18$) or parental interference ($n = 1$).

#### *Materials*

Infants were presented with a familiarization video that consisted of a red circle centered at the bottom of the screen that flashed while they heard a two-syllable pseudoword being played on speakers. Pseudowords were chosen to specifically investigate whether learning of sequences consisting of speech-like sounds could be affected by learning of non-speech-like sequences in the transfer experiment (Experiment 2B). The target word was created using *Cool Edit* (Syntrillium Software, 2000) by randomly connecting discrete syllables recorded by a female speaker (*Ki-Tu*) with .06 s between syllables. The video presented this sequence five times and lasted 9 s. The video was repeated 10 times for a total of 50 sequence presentations and 90 s of continuous familiarization. Test items were created by adding a randomly selected syllable to either the beginning of the familiarization item (Pre+: *FO-Ki-Tu*) or the end of the familiarization item (Post+: *Ki-Tu-FO*). There was also a novel sequence used to determine novelty preference (*Gu-Ti-Hi*), which was presented as the third and final test trial. Each test video presented the test item five times and lasted 9 s. After each video in the experiment, the screen was black for 750 ms. During each familiarization and test video, the auditory sequence was presented more times than the corresponding visual sequence from Experiment 1 to control for presentation time (i.e., five times for the auditory sequence vs. four times for the visual sequence). As a result, similar to visual sequences, auditory sequences were presented for a total duration of 9 s.

#### *Design and procedure*

Infants were familiarized with a sequence of auditory syllables appearing as a pseudoword. Then they were randomly presented with two test trials: Pre+ (component added to beginning of familiarized sequence) and Post+ (component added to end of familiarized sequence). For 8-month-olds, nine received the Pre+ trial first and 11 received the Post+ trial first. For 16-month-olds, 10 received the Pre+ trial first and 11 received the Post+ trial first. The final test trial was the novel auditory sequence.

### *Results and discussion*

As shown in Fig. 3A, participants reduced looking by 49.24% between the first three and last three familiarization trials, paired-samples $t(40) = 11.68$, one-tailed $p < .001$, $d = 1.83$. An ANOVA on average looking across all familiarization trials did not reveal a significant main effect of age. Once again, participants demonstrated novelty preference by exhibiting longer looking to the novel test sequence than the average of their own last three familiarization trials, paired-samples $t(40) = 8.07$, one-tailed

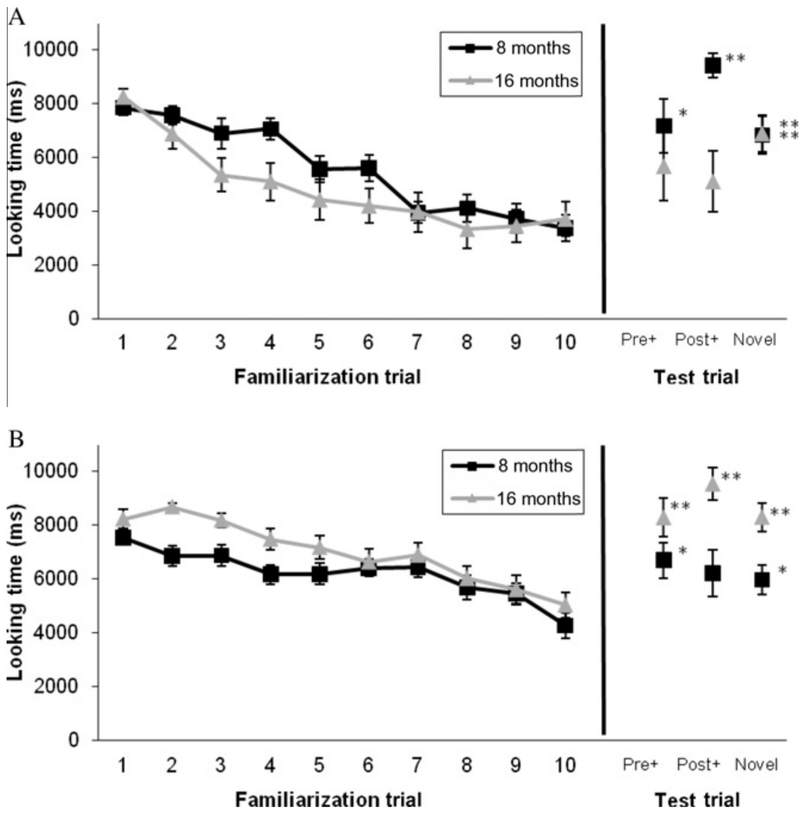


**Fig. 3.** (A) Auditory baseline looking times across trials for 8- and 16-month-olds. (B) Looking times in the auditory modality after training to attend to the end of a visual sequence for 8- and 16-month-olds. Pre+ and Post+ trials were between participants. Bars represent standard errors. Differences from the last three familiarization trials are indicated with one-tailed *p* values: $^{*}p < .05$; $^{**}p < .01$.

$p < .001$, $d = 1.26$. As shown in Table 2 and Fig. 3A, only 8-month-olds exhibited recovery to the Pre+ or Post+ trials, whereas 16-month-olds failed to do so on either test trial type, indicating that 16-month-olds were likely considering the test stimuli as variants of the familiarized sequence.

The primary analyses focused on looking to the first test trial of interest (Pre+ or Post+) compared with the average of their own last three familiarization trials. These recovery scores were subjected to a 2 (Age: 8 or 16 months) × 2 (Trial Type: Pre+ or Post+) between-participants ANOVA. There was only a main effect of age, $F(1, 37) = 4.58$, $p < .05$, $d = 0.67$, with 8-month-olds ($M = 4097.78$ ms, $SD = 2860.36$) having higher recovery scores than 16-month-olds ($M = 1875.47$ ms, $SD = 3418.24$).

It was concluded that participants had no preference in the auditory modality; they were equally likely to notice changes in the beginning and at the end of the auditory sequence. Therefore, to examine transfer of learning across modalities, we conducted Experiment 2B, in which we once again trained infants to pay attention to the end of a visual sequence and examined changes in baseline preferences in the auditory modality as a result of this training.

## Experiment 2B: visual training–auditory test

Participants were trained to attend to the end of a visual sequence, and then they were tested on the same auditory sequences as in Experiment 2A. Training was identical to that in Experiment 1B (i.e., final object of visual sequence was salient), whereas familiarization and testing were similar to those in Experiment 2A (i.e., participants were familiarized to auditory sequence and then tested on Pre+ and Post+ auditory sequences). Note that testing sequences were presented in a different modality than training sequences and that the spatial component present in the visual training sequences (left to right) was absent in the auditory test sequences. Therefore, any transfer of learning from the visual modality to the auditory modality would be substantially broader than that observed in Experiment 1B.

### Method

#### Participants

Participants were 8-month-olds ($n = 34$, 17 boys and 17 girls, mean age = 252.97 days, $SD = 12.33$) and 16-month-olds ($n = 33$, 18 boys and 15 girls, mean age = 490.73 days, $SD = 11.12$). An additional 18 infants were excluded for failure to complete the experiment due to excessive fussiness ($n = 14$), failure to look to at least 50% of familiarization trials ($n = 1$), technical difficulty ($n = 2$), or sibling interference ($n = 1$).

#### Materials, design, and procedure

All infants were trained to attend to the end of a visual sequence (sun → heart) as in Experiment 1B, where the final object in the visual sequence was made salient. This training was immediately followed by familiarization and test phases in the auditory modality. Similar to Experiment 2A, infants were first familiarized to an auditory sequence (*Ki-Tu*) and then presented with sequences that had components added to either the beginning of the sequence (Pre+: *FO-Ki-Tu*) or the end of the sequence (Post+: *Ki-Tu-FO*). For 8-month-olds, 19 received the Pre+ trial first and 15 received the Post+ trial first. For 16-month-olds, 17 received the Pre+ trial first and 16 received the Post+ trial first. The final test trial was the novel auditory sequence (*Gu-Ti-Hi)*, which was used to establish novelty preference.

### Results and discussion

As shown in Fig. 3B, participants reduced looking by 30.55% between the first three and last three familiarization trials, paired-samples $t(66) = 8.88$, one-tailed $p < .001$, $d = 1.09$. An ANOVA on average looking across all familiarization trials revealed a significant main effect of age, $F(1, 65) = 5.00$, $p < .05$, partial $\eta^2 = .07$, with 16-month-olds looking on familiarization trials longer than 8-month-olds. Infants also demonstrated novelty preference by exhibiting longer looking to the novel sequence than the average of their own last three familiarization trials, paired-samples $t(66) = 5.07$, one-tailed $p < .001$, $d = 0.62$. As shown in Table 2 and Fig. 3B, 16-month-olds exhibited significant recovery to both the Pre+ and Post+ items, whereas 8-month-olds exhibited recovery only to the Pre+ items.

A recovery score was calculated for each infant by subtracting looking to the average of his or her last three familiarization trials from looking to his or her first auditory test sequence (Pre+ or Post+). These difference scores were compared with their equivalents in the no-training auditory baseline condition (Experiment 2A).

A 2 (Age: 8 or 16 months) × 2 (Condition: baseline or training) × 2 (Trial Type: Pre+ or Post+) between-participants ANOVA on the difference scores revealed a significant Condition × Age interaction, $F(1, 100) = 14.23$, $p < .001$, partial $\eta^2 = .13$. To tease this apart, the age groups were analyzed separately. As shown in Fig. 4, after training, 8-month-olds failed to increase looking to the auditory Post+ item compared with the baseline. In fact, they decreased their looking to the auditory Post+ item, $F(1, 24) = 15.15$, $p = .001$, $d = 1.23$, as well as to the auditory Pre+ item, $F(1, 26) = 2.08$, one-tailed $p = .08$, $d = 0.57$, after being trained in the visual modality. In contrast, 16-month-olds increased looking to the auditory Post+ item (but not to the auditory Pre+ item) in comparison with the auditory baseline condition, $F(1, 25) = 2.77$, one-tailed $p = .05$, $d = 0.63$. Overall, these findings suggest that only 16-month-olds exhibited evidence of transfer across modalities from the visual modality to the auditory modality.

Taken together, results of the two experiments indicate that for 8-month-olds transfer is limited to the learning modality, whereas 16-month-olds can transfer across modalities. These findings suggest that transfer becomes broader in the course of development.

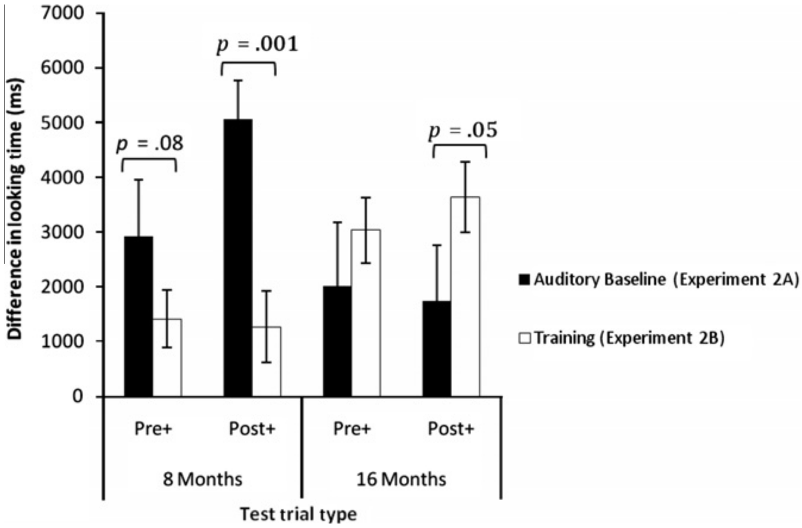


**Fig. 4.** Recovery times (test trial minus infant's average of last three familiarization trials) in the auditory modality without training and after training to attend to the end of a visual sequence for 8- and 16-month-olds. Bars represent standard errors. Differences between baseline recovery scores and recovery scores after training are indicated with one-tailed *p* values.

## General discussion

The goal of this research was to examine the development of the ability to transfer learned knowledge to new situations. To address this issue, we focused on learning to attend to the end of a sequence. The reported experiments present two novel findings pertaining to the issue of generalization of learning. First, both 8- and 16-month-olds can transfer learning to attend to the end of a visual sequence to a novel visual sequence. Second, 16-month-olds, but not 8-month-olds, can transfer this learning across modalities—from visually to aurally presented sequences. These findings present evidence that what is initially learned is fairly specific and is bound to the learned modality, resulting in narrow transfer. At the same time, transfer becomes broader in the course of development.

The problem of transfer (or applying learned knowledge to a new situation) has a long tradition in psychology (see Barnett and Ceci (2002) and Detterman (1993), for reviews). There is a wide range of opinions varying from rather pessimistic assessments that what is learned is bound to specifics of a learning situation, and thus transfer is unlikely (see Detterman (1993) and Lave and Wenger (1991)), to rather optimistic ones arguing that even during infancy what is learned is not bound to the particulars of the training stimuli (e.g., Frank et al., 2009) and, thus, very broad transfer is likely. The current research was based on the idea that transfer is neither impossible nor given but rather is itself a product of learning and development. The reason for this is that the level of abstraction of what is learned comes in degrees, and so does transfer. For example, when taught that 2 + 4 + 6 = 6 + 4 + 2, a participant may learn that commutativity holds only for these three numbers, only for monotonically increasing triples of even numbers, or for all possible addends. Each of these learning outcomes may result in different transfer of learning outside of the learning situation.

We argue that training resulted in participants shifting attention to the end of a visual sequence. However, what they learned as a result of such training is not clear. Specifically, they could have learned to shift attention to the end of (a) the trained visual sequence, (b) any visual sequence, or (c) a sequence. Each of these possibilities would result in different transfer, with more abstract learning resulting in transfer across modalities. Our research suggests that even 8-month-olds exhibit evidence for the second possibility (b), whereas 16-month-olds exhibit evidence for the third possibility (c). These findings are critically important in the context of extant literature on transfer, with some arguing that transfer never happens and others arguing that from very early in development initial representation is not bound to details of specific stimuli and, as a result, exceedingly broad generalizations are possible. Overall, the current research presents evidence that transfer is present early in development but is initially limited by details of the training situation. At the same time, it becomes broader with development, suggesting that development and learning may result in the ability to form more abstract representations. Therefore, results of the current research indicating early onset and broadening of transfer with development may have interesting implications for theories of learning and transfer.

The reported findings present interesting parallels with literature on visual development and on word learning, also presenting evidence that transfer may be broadening with development. For example, there is evidence from visual development (see Bhatt and Quinn (2011), for a review) that when trained on displays that are organized into columns by luminance, 3- and 4-month-olds can recognize organization by luminance in novel displays but cannot recognize columns organized by shape. That is, they exhibit evidence of luminance-to-luminance transfer but not of luminance-to-shape transfer. At the same time, 6- and 7-month-olds exhibit evidence of both luminance-to-luminance and luminance-to-shape transfer. Therefore, transfer in 6- and 7-month-olds appears to be broader than transfer in 3- and 4-month-olds.

Similarly, in the word learning literature, researchers demonstrate that generalization of the learned word becomes progressively broader with age (Smith, 2003; Son et al., 2008; Woodward et al., 1994). For example, Vlach and Sandhofer (2011) demonstrated that 2.5-year-olds could extend an object label only when the learning and testing contexts matched, whereas 4-year-olds could extend labels across discrepant contexts. Here again, transfer in older participants appears to be broader than that in younger participants. One explanation of this is that transfer broadens as children become more experienced word learners. The current findings indicate that these effects are not limited to visual development or word learning; in a more general learning situation, transfer became broader with development. Perhaps as children become more experienced *learners*, they form broader representations of what is being learned and, as a result, exhibit broader transfer. Therefore, it is possible that development results in learning to learn. This learning to learn, in turn, may result in representations that "are not so narrow as to disallow the process of generalization, yet not so broad that it applies to everything" (Aslin, 2011, p. 41).

One particular aspect of the current findings deserves more discussion. As shown in Experiment 2, training with visual sequences affected attention allocated to auditory sequences. Given that the latter were English syllables, it is possible that at least 16-month-olds perceived these speech-like sequences as linguistic (cf. Saffran, Aslin, & Newport, 1996). For example, work by Saffran and colleagues demonstrated that exposure to similar speech-like sequences facilitates subsequent performance on a learning task for 17-month-olds (Graf Estes, Evans, Alibali, & Saffran, 2007) and on a word generalization task for 22-month-olds (Lany & Saffran, 2010). Similar to this idea, Hupp and colleagues (2009, Experiment 3D) found that when adults were trained to attend to the end of a visual sequence, their subsequent label extensions were affected. More specifically, their baseline label extensions (i.e., without training) relied heavily on the beginning of the linguistic sequence; however, after training to attend to the end of a visual sequence, this reliance on the beginning of the linguistic sequence decreased significantly.

If infants perceived the stimuli as linguistic in nature, then perception of linguistic sequences was affected by nonlinguistic visual training. In this case, current findings indicate that language learning could be affected by learning outside of language, casting doubt on the idea of modular impenetrability of language. If language learning could be affected by nonlanguage factors, these findings should be considered as evidence for domain-general and nonmodular (vs. strictly modular) language learning mechanisms (cf. Hupp et al. (2009), for related arguments). Future research should further examine the domain generality of language learning during infancy. For example, if training in the visual modality would affect morphological, syntactical, or semantic aspects of spoken language, such findings would strengthen the domain-general argument. In addition, future research should further examine the development of this mechanism by both modifying the procedure to make it less taxing for 8-month-olds and examining the change in processing between 8 and 16 months of age.

In summary, this research found generalization of learning to broaden with development between 8 and 16 months of age; whereas younger participants exhibited evidence of transfer only within the trained modality, older participants transferred across modalities. Transfer is an essential component of learning because it allows individuals to take learning beyond the particulars of the stimuli. Although the developmental mechanisms of this remain unclear, these findings suggest that representation of structure undergoes development, resulting in the ability to transfer new knowledge more broadly. These findings have important implications for our understanding of learning and transfer during infancy.


## Acknowledgments

This research was supported by National Science Foundation (NSF) Grant BCS-0720135 and by National Institutes of Health (NIH) Grant R01HD056105 to V.M.S. We thank Chris Robinson, Anna Fisher, and Catherine Best for their helpful comments on earlier drafts of this article. A portion of the data was presented at the 2007 Conference of the Society for Research in Child Development in Boston.


## References


Altmann, G. T. M., Dienes, Z., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 899–912.

Aslin, R. N. (2011). Perceptual organization of visual structure requires a flexible learning mechanism. *Infancy, 16*, 39–44.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*, 612–637.

Barr, R. (2010). Transfer of learning between 2D and 3D sources during infancy: Informing theory and practice. *Developmental Review, 30*, 128–154.

Bhatt, R. S., & Quinn, P. C. (2011). How does learning impact development in infancy? The case of perceptual organization. *Infancy, 16*, 2–38.

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 24–39.

DeLoache, J. S. (1991). Symbolic functioning in very young children: Understanding of pictures and models. *Child Development, 62*, 736–752.

Detterman, D. K. (1993). The case for the prosecution: Transfer as an epiphenomenon. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 1–24). Westport, CT: Ablex.

Elman, J. L., Bates, E., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Frank, M. C., Slemmer, J. A., Marcus, G. F., & Johnson, S. P. (2009). Information from multiple modalities helps 5-month-olds learn abstract rules. *Developmental Science, 12*, 504–509.

Goldstone, R. L., & Landy, D. (2010). Domain-creating constraints. *Cognitive Science, 34*, 1357–1377.

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science, 18*, 254–260.

Hayne, H., Barr, R., & Herbert, J. (2003). The effect of prior practice on memory reactivation and generalization. *Child Development, 74*, 1615–1627.

Hupp, J. M., Sloutsky, V. M., & Culicover, P. W. (2009). Evidence for a domain general mechanism underlying the suffixation preference in language. *Language and Cognitive Processes, 24*, 876–909.

Kloos, H., & Sloutsky, V. M. (2008). What's behind different kinds of kinds: Effects of statistical density on learning and representation of categories. *Journal of Experimental Psychology: General, 137*, 52–72.

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

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.

Macromedia Studio MX (2002). *Macromedia flash MX [computer software]*. San Francisco: Macromedia Studio MX.

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283*, 77–80.

Morris, A. K., & Sloutsky, V. M. (1998). Understanding of logical necessity: Developmental antecedents and cognitive consequences. *Child Development, 69*, 721–741.

Neurobehavioral Systems (2003). *Presentation (version .76) [computer software]*. Albany, CA: Neurobehavioral Systems.

Robinson, C. W., & Sloutsky, V. M. (2007a). Visual processing speed: Effects of auditory input on visual processing. *Developmental Science, 10*, 734–740.

Robinson, C. W., & Sloutsky, V. M. (2007b). Visual statistical learning: Getting some help from the auditory modality. In D. McNamara & G. Trafton (Eds.), *Proceedings of the XXIX annual conference of the cognitive science society* (pp. 611–616). Mahwah, NJ: Lawrence Erlbaum.

Robinson, C. W., & Sloutsky, V. M. (2010). Development of cross-modal processing. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*, 135–141.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928.

Sloutsky, V. M. (2010). Mechanisms of cognitive development: Domain-general learning or domain-specific constraints? *Cognitive Science, 34*, 1125–1130.

Sloutsky, V. M., & Fisher, A. V. (2008). Attentional learning and flexible induction: How mundane mechanisms give rise to smart behaviors. *Child Development, 79*, 639–651.

Sloutsky, V. M., & Robinson, C. W. (2008). The role of words and sounds in infants' visual processing: From overshadowing to attentional tuning. *Cognitive Science, 32*, 342–365.

Smith, L. B. (2003). Learning to recognize objects. *Psychological Science, 14*, 244–250.

Smith, L. B., Colunga, E., & Yoshida, H. (2010). Knowledge as process: Contextually cued attention and early word learning. *Cognitive Science, 34*, 1287–1314.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science, 13*, 13–19.

Son, J. Y., Smith, L. B., & Goldstone, R. L. (2008). Simplicity and generalization: Short-cutting abstraction in children's object categorization. *Cognition, 108*, 626–638.

Syntrillium Software (2000). *Cool Edit (2000 version) [computer software]*. Phoenix, AZ: Syntrillium Software.

Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word learning. *Journal of Experimental Child Psychology, 108*, 394–401.

Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology, 30*, 553–566.