Developmental Science DOI: 10.1111/j.1467-7687.2007.00627.x

REPORT

Visual processing speed: effects of auditory input on visual processing

Christopher W. Robinson and Vladimir M. Sloutsky

Center for Cognitive Science, Ohio State University, USA

Abstract

The ability to process simultaneously presented auditory and visual information is a necessary component underlying many cognitive tasks. While this ability is often taken for granted, there is evidence that under many conditions auditory input attenuates processing of corresponding visual input. The current study investigated infants' processing of visual input under unimodal and cross-modal conditions. Results of the three reported experiments indicate that different auditory input had different effects on infants' processing of visual information. In particular, unfamiliar auditory input slowed down visual processing, whereas more familiar auditory input did not. These results elucidate mechanisms underlying auditory overshadowing in the course of cross-modal processing and have implications on a variety of cognitive tasks that depend on cross-modal processing.

Introduction

Infants live in a dynamic world where they constantly encounter information presented to different sensory modalities. Under many conditions, related information is presented to auditory and visual modalities (e.g. words and the objects that these words denote or objects and the sounds that these objects produce) and infants have to process and integrate this cross-modal information.

Under some conditions (e.g. when the same amodal relation such as rhythm or rate) is presented cross-modally, cross-modal presentation is likely to facilitate processing of the amodal relation (see Lewkowicz, 2000; Lickliter & Bahrick, 2000, for reviews). At the same time, under other conditions (e.g. when auditory–visual pairings are arbitrary), cross-modal presentation is likely to hinder processing of arbitrary auditory–visual stimuli. These arbitrary pairings reflect situations when a person has to process the appearance of a novel item (e.g. an animal or a machine) and the sound it produces. Another important example of arbitrary pairings is a situation when a person has to simultaneously process a novel word and a novel referent of this word. In both of these situations, young participants are more likely to encode a visual stimulus when presented unimodally than when paired with an auditory stimulus (e.g. Napolitano & Sloutsky, 2004; Robinson, Howard & Sloutsky, 2005; Robinson & Sloutsky, 2004a, 2004b, 2007; Sloutsky & Napolitano, 2003). Therefore, it was concluded that auditory input overshadows or attenuates processing of arbitrarily paired visual input.

The study of auditory overshadowing reveals several important regularities. First, there is a developmental pattern, with infants exhibiting mostly auditory overshadowing (Robinson & Sloutsky, 2004a), young children exhibiting evidence of both auditory and visual overshadowing (Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2004a; Sloutsky & Napolitano, 2003), and adults exhibiting mostly visual overshadowing (Colavita, 1974; Colavita & Weisberg, 1979). In addition, early in development, auditory overshadowing is mediated by stimulus familiarity and processing time: under relatively brief processing conditions, familiar auditory input overshadows unfamiliar visual input (Napolitano & Sloutsky, 2004), whereas under protracted processing conditions, familiar auditory input is less likely to interfere with processing of visual input (Robinson & Sloutsky, 2004b).

However, the mechanism underlying auditory overshadowing effects remains unknown. One possibility is that unfamiliar auditory input engages attention and dominates initial processing, thus attenuating processing

Address for correspondence: Vladimir M. Sloutsky or Christopher W. Robinson, Center for Cognitive Science, Ohio Stadium East, 1961 Tuttle Park Place, Ohio State University, Columbus, OH 43210, USA; e-mail: sloutsky.1@osu.edu or robinson.777@osu.edu

of a corresponding visual input. At the same time, familiar auditory input may not only quickly engage attention (Christie & Klein, 1995), but it may also quickly release attention, thus making familiar auditory input less likely to interfere with processing of visual input. If these considerations are correct, then different types of auditory input may result in a different time course of processing corresponding visual input. In particular, unfamiliar auditory stimuli may be more likely to slow down visual processing compared to more familiar auditory stimuli.

The current study tests this idea by employing a 'continuous familiarization' procedure (see Fantz, 1964; Mather & Schafer, 2005; Roder, Bushnell & Sasseville, 2000; Rose, Feldman & Jankowski, 2002, for similar procedures). As in previous research examining visual processing speed, infants were familiarized to two simultaneously presented visual stimuli: one visual stimulus remained unchanged across familiarization (hereafter, *familiar*), whereas the other visual stimulus changed on every trial (hereafter, *novel*). Given that infants prefer novel stimuli to familiar stimuli, infants should reliably prefer the novel stimuli when the repeated stimulus becomes familiar. Visual processing speed can then be inferred by the amount of familiarization needed for infants to shift attention to the *novel* stimulus.

Visual stimuli were either presented in isolation (unimodal presentation) or paired with an auditory stimulus (cross-modal presentation). Experiment 1 compared processing of the same visual stimuli when presented in isolation or when paired with a word or unfamiliar sound. Experiments 2 and 3 examined the effect of auditory familiarity on visual processing.

Experiment 1

Method

Participants

Fifty-eight 14-month-olds (37 boys and 21 girls, $M = 437$ days, $SD = 65$ days) participated in this experiment. Parents' names were collected from local birth announcements, and contact information was obtained through local directories. All children were full-term (i.e. > 2500 g birth weight) with no auditory or visual deficits, as reported by parents. A majority of infants were Caucasian. Infants had to contribute a mean to each of the six familiarization blocks (i.e. each block is four familiarization trials) to be included in the current experiment. An additional 20 infants were tested but not included in the current experiment.

Materials and design

The experiment had three between-subjects conditions: (1) no-auditory baseline, (2) unfamiliar sound, and (3) word. Across the conditions, participants were presented with the same visual stimuli, which were realistic pictures of animals and commonplace objects (e.g. ball, dog, car, etc.). Each visual stimulus was approximately 36 cm × 36 cm and images were presented in pairs with approximately 50 cm separating the two images. The auditory stimuli consisted of a linguistic label or a laser sound. The label was produced by a female experimenter in infant-directed speech (i.e. 'Look at the dax'). Both auditory stimuli were presented by a computer for 1200 ms at the onset of each trial at 65–68 dB. Previous research has demonstrated that the laser sound was unfamiliar to 4-year-olds (Robinson & Sloutsky, 2004a), and thus it was assumed that this sound would also be unfamiliar to infants in the current study. Furthermore, in contrast to previous research examining auditory overshadowing effects (Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2004a, 2004b; Sloutsky & Napolitano, 2003), the visual stimuli in the current experiment and in all subsequent experiments were presented for longer durations than the auditory stimulus: while auditory stimuli were presented for only 1200 ms, the visual stimuli were presented for 8000 ms.

Apparatus

Infants were seated on parents' laps approximately 100 cm away from a 152 cm × 127 cm projection screen. A NEC GT2150 LCD projector presented images to the infants and was mounted on the ceiling approximately 30 cm behind the infant (130 cm away from the projection screen). Two Boston Acoustics 380 speakers presented auditory stimuli to infants. These speakers were 76 cm apart from each other and mounted in the wall at the infant's eye level. The projector and speakers received visual and auditory signals from a Dell Dimension 8200 computer, which was controlled by *Presentation* software. This computer was also used to record visual fixations. Fixations were recorded online by pressing a button on a 10-button USB gamepad when infants were looking at the stimulus and releasing the button when infants looked away from the stimulus.

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. A random sample of 33% of the infants was coded offline. Offline coders concealed the half of the

	Unimodal presentation			Cross-modal presentation		
		Modality			Modality	
		Visual	Auditory		Visual	Auditory
Time	Trial 1:	V1 V2	None	Trial 1:	V1 V2	A1
	Trial 2:	V1 V3	None	Trial 2:	V1 V3	A1
	Trial 3:	V1 V4	None	Trial 3:	V1 V4	A1

Figure 1 *Overview of the procedure.*

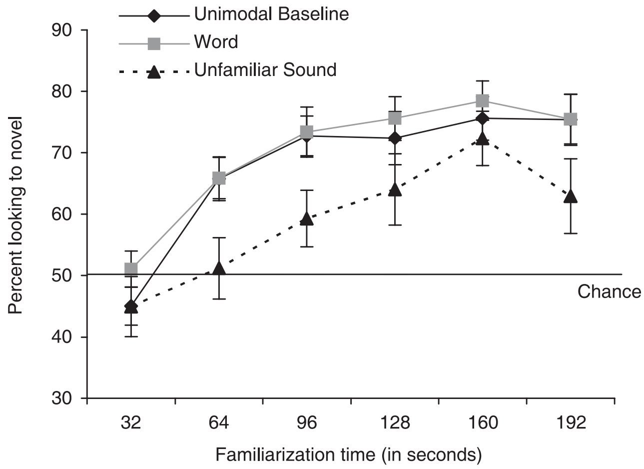

Figure 2 *Visual processing speed by Stimulus Condition and Time in Experiment 1. Error bars represent standard errors of the mean.*

split-screen associated with the stimulus presentation, thus blinding themselves to the auditory and visual information presented to infants. Offline coders then coded infants' visual fixations at a resolution of 30 frames per second. Reliabilities for online and offline coders were calculated for each infant and averaged across all reported experiments, average $r = .94$.

Procedure

The procedure consisted of 24 familiarization trials. On each trial, infants were simultaneously presented with two visual stimuli (e.g. VIS_1 and VIS_2), which were presented for 8000 ms. Each successive trial consisted of presenting a new visual stimulus and the old visual stimulus (see Figure 1 for an overview of the procedure). Twenty infants heard no auditory input (unimodal baseline condition), 18 infants heard the laser sound at the onset of each trial (unfamiliar sound condition) and 20 infants heard a word at the onset of each trial (word condition). The left–right location of the novel and familiar stimuli was held constant across familiarization for each infant and counter-balanced between subjects.

Results and discussion

Analyses focused on infants' attention to the visual stimuli and speed of visual processing. Infants' attention to the visual stimuli was measured by summing fixation durations to both visual stimuli across the 24 familiarization trials. Infants could accumulate 192 s of looking during the experiment (i.e. 24 trials with 8 s duration for each trial). To determine how quickly infants processed the visual stimuli, a novelty preference score was calculated on each familiarization trial and averaged across four familiarization trials (i.e. every 32 s). The amount of familiarization needed before infants demonstrated a reliable novelty preference served as a measure of visual processing speed.

An ANOVA with stimulus condition as a between-subjects factor revealed that looking to the visual stimuli differed across the stimulus conditions, $F(2, 55) = 6.38$, $p < .005$. Independent-samples t-tests confirmed that infants accumulated more looking in the word condition ($M = 128$ s, $SE = 4$ s) and sound condition ($M = 135$ s, $SE = 6$ s) than in the unimodal baseline ($M = 107$ s, $SE = 7$ s), ps $< .014$. No differences were found between the word and sound conditions, $t(36) = 0.98$, $p = .33$, which suggests that the word and sound were equally salient.

Although infants accumulated more looking when visual stimuli were accompanied by an auditory stimulus, this increase in attention did not speed up visual processing (see Figure 2). Rather, infants in the unfamiliar sound condition required 128 s of familiarization before they reliably looked to the novel stimulus, looking to *novel* in Blocks 4–6 was above 50%, one-sample ts > 2.43, ps $> .027$. In contrast, infants in the unimodal condition and in the word condition demonstrated a reliable novelty preference after 64 s of familiarization, looking to *novel* in Blocks 2–6 was above 50%, one-sample ts > 4.44, ps $< .001$.

Additional analyses focused on the percentage of infants that demonstrated a reliable novelty preference, as indicated by the presence of a *critical run* (see Roder *et al.*, 2000, for a similar procedure). As in Roder *et al.*, a novelty preference score was averaged across blocks of three consecutive trials (e.g. Block 1 = Trials 1–3, Block 2 = Trials 2–4, etc.), and a critical run was identified as five consecutive blocks where looking to the novel stimuli exceeded 60%. The percentage of infants demonstrating a critical run and the average onset of the critical run for these infants (i.e. first trial of the critical run) are presented in Table 1. As can be seen in the table, even when excluding infants who did not demonstrate a critical run, the onset of visual discrimination occurred

Table 1 *Average onset of visual discrimination (i.e. critical run) and percentage of infants demonstrating a critical run by stimulus condition*

Stimulus condition	First trial of critical run (averaged across infants who demonstrated a critical run)	Percentage of infants demonstrating a critical run
Unimodal Baseline (Experiment 1)	6.0	85
Unfamiliar Sound (Experiment 1)	8.5	72
Linguistic Label (Experiment 1)	4.6	80
Referential Sound (Experiment 2)	5.6	77
Pre-familiarized Sound (Experiment 3)	5.1	95

earlier in the label condition than in the unfamiliar sound condition, $t(27) = 2.43$, $p < .05$.

In summary, the results of Experiment 1 indicate that unfamiliar sounds slowed down or delayed processing of corresponding visual input, whereas words did not interfere with visual processing (compared to the unimodal baseline). Why did sounds but not words interfere with visual processing? One possibility is that words are processed more efficiently than sounds. A second possibility is that infants may have assumptions about the referential nature of words: words but not sounds refer to objects and categories (e.g. Waxman, 2003; Xu, 2002). A third possibility is that differences may reflect the presence of a referential context. While a referential context may include a host of linguistic and nonlinguistic cues (e.g. pointing, eye gaze, etc.), previous research has demonstrated that simply embedding a linguistic or nonlinguistic cue in a familiar naming frame (e.g. 'Look at the X') is enough to facilitate cross-modal binding (Campbell & Namy, 2003; Namy & Waxman, 2000; Woodward & Hoyne, 1999). Recall that the words in Experiment 1 were presented in a referential context, whereas sounds were presented in isolation. If the latter possibility is the case, then unfamiliar sounds presented in a referential context should have similar effects as words, thus making the unfamiliar sounds less likely to attenuate visual processing. This possibility was examined in Experiment 2.

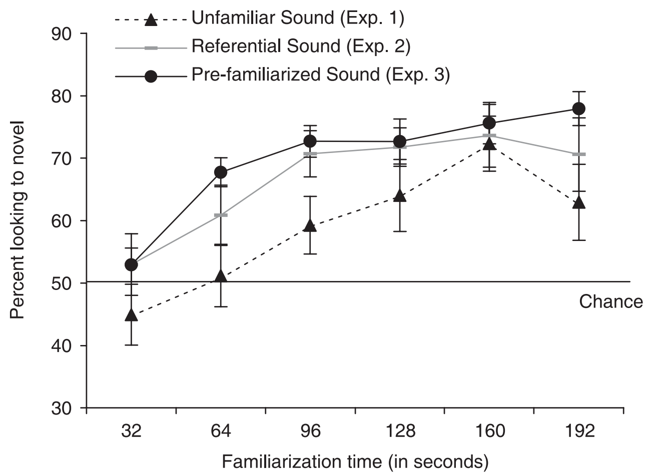

Figure 3 *Visual processing speed by Stimulus Condition and Time across Experiments 1–3. Error bars represent standard errors of the mean.*

Experiment 2

Method

Participants, stimuli and procedure

Thirteen 14-month-olds (eight boys and five girls, $M = 422$ days, $SD = 64$ days) participated in the current experiment. Recruitment procedures and demographics were identical to previous experiments. Two infants were tested but not included in the current experiment because they did not contribute a mean to each of the six familiarization blocks. The procedure was identical to the unfamiliar sound condition of Experiment 1 except that the laser sound was embedded in a referential context (i.e. 'Look at the' <laser sound>).

Results and discussion

As can be seen in Figure 3, as in the word and unimodal conditions of Experiment 1, infants presented with unfamiliar sounds embedded in a referential context processed the visual stimuli after 64 s of familiarization, looking to *novel* in Blocks 2–6 was above 50%, one-sample $ts > 2.24$, $ps < .05$ (see Table 1 for the percentage of infants demonstrating a critical run and the average onset of the critical run). Therefore, unlike Experiment 1 where unfamiliar sounds presented in isolation interfered with visual processing, in Experiment 2, cross-modal interference disappeared when the same sound was presented in a referential context. These findings indicate that effects of words on visual processing found in Experiment 1 stem at least in part from words being presented in a referential context.

Why do effects of sounds presented in a referential context differ from effects of sounds presented in isolation? There are at least two possibilities. First, it is possible that a referential context is more conducive for processing because it engages a listener in social interaction by conveying the intentions of the speaker (e.g. the speaker intended to refer to the object with the nonlinguistic cue). However, it is also possible that a referential context provides a familiar context, which increases the overall familiarity of the auditory stimulus. The goal of Experiment 3 was to test this possibility by examining if stimulus familiarity can account for the effect of a referential context on visual processing.

Experiment 3

Method

Participants, stimuli and procedure

Twenty 14-month-olds (five boys and 15 girls, $M = 469$ days, $SD = 50$ days) participated in the current experiment. Recruitment procedures and demographics were identical to previous experiments. Four infants were tested but not included in the current experiment because they did not contribute a mean to each of the six familiarization blocks.

The stimuli and the experiment proper were identical to the unfamiliar sound condition of Experiment 1. In contrast to Experiment 1, however, infants were presented with a pre-familiarized sound rather than with an unfamiliar sound. Rather than using a sound that is typically found in the infant's environment and making the assumption that it is familiar, the current experiment increased the familiarity of the auditory stimulus by pre-familiarizing infants to the nonlinguistic sound prior to pairing it with the visual stimuli. This manipulation not only ensures that all infants have some familiarity with the auditory stimulus but it also allows for direct comparisons to be made with the other sound conditions (e.g. the same sound that was unfamiliar in Experiment 1 is somewhat familiar in Experiment 3).

Infants in the pre-familiarization phase sat on the parent's lap and heard the laser sound 20 times (in isolation). The auditory stimulus was not paired with any visual stimulus during pre-familiarization, and, as in the experiment proper, the auditory stimulus was 1200 ms in duration and was presented at 65–68 dB. After the pre-familiarization phase, infants were given a short 2–5 minute break and then they participated in the experiment proper, which was identical to Experiment 1 where the sounds were presented in isolation.

Results and discussion

As can be seen in Figure 3, infants in the pre-familiarized sound condition demonstrated a novelty preference after 64 s of familiarization, looking to *novel* in Blocks 2–6 was above 50%, one-sample $ts > 7.64$, $p < .001$, which was similar to the referential context condition (see Table 1 for the percentage of infants demonstrating a critical run and the average onset of the critical run). These findings demonstrate that giving infants an opportunity to become familiar with an auditory stimulus prior to pairing it with a visual stimulus or embedding an unfamiliar auditory stimulus within a familiar context attenuated auditory overshadowing effects. Furthermore, it is important to note that increasing the familiarity of the auditory stimulus attenuated the difference between sounds and words: while words and unfamiliar sounds presented in isolation had different effects on visual processing (Experiment 1), words and pre-familiarized sounds had comparable effects.

General discussion

In many situations infants have to process cross-modal information. While cross-modal presentation facilitates processing of some information such as amodal relations (e.g. Bahrick & Lickliter, 2000), research examining the processing of arbitrary auditory–visual pairings often shows that cross-modal presentation hinders processing of the modality-specific components. Current findings contribute to this latter line of research by demonstrating that unfamiliar auditory input slows down processing of corresponding static visual input (these findings may or may not hold for dynamic visual input), whereas more familiar auditory input did not interfere with processing of corresponding visual input.

These findings highlight mechanisms underlying auditory overshadowing effects, expand research examining the effects of a referential context on binding of arbitrary auditory–visual pairings, and have implications on a variety of cognitive tasks that hinge on cross-modal processing.

First, the reported results indicate that, at least in part, auditory overshadowing effects stem from auditory input slowing down or delaying the onset of visual processing. This is an important finding indicating that auditory input is more likely to affect processing of visual input earlier in the course of processing than later in processing. Recall that unfamiliar sounds and words had comparable effects after 128 s of familiarization (see Figure 1), whereas differences were more pronounced earlier in processing. Second, unlike unfamiliar sounds, words presented in a referential context are less likely to interfere with processing of corresponding visual input. Furthermore, this attenuated interference is not unique to either linguistic input or to a referential context but it rather reflects familiarity of auditory input. In other words, as shown by Christie and Klein (1995), familiar auditory input is fast to engage attention and as suggested by current results, it is also fast to release attention. Therefore, familiar auditory input may have different effects on the encoding of a visual stimulus at different points of processing. Familiar auditory input is likely to result in overshadowing effects early in processing (something that cannot be captured by the current paradigm, but which has been demonstrated with a different paradigm and a different population, see Napolitano & Sloutsky, 2004). At the same time, as shown in the current experiments, familiar auditory input is less likely to interfere with processing of visual input later in the course of processing.

The current findings also have implications for understanding the mechanisms underlying the effects of a referential context on binding of arbitrary auditory–visual pairings (Campbell & Namy, 2003; Namy & Waxman, 2000; Woodward & Hoyne, 1999). In particular, while it is often assumed that a referential context facilitates binding of arbitrary relations because it conveys the intentions of the speaker, the current study suggests that it is possible that some of these effects stem from more efficient processing. This effect can occur in two ways. First, it is possible that a cue presented in a referential context is processed faster than the same cue presented in isolation (cf. Fernald & Hurtado, 2006). Alternatively, it is possible that referential context is a familiar cue which adds to the overall familiarity of auditory input, which in turn attenuates cross-modal interference. While future research is needed to distinguish between these alternative explanations, the current study suggests that some of the effects of a referential context on the binding of arbitrary relations may stem from a referential context attenuating cross-modal interference.

Finally, the current findings reveal important aspects of cross-modal processing and generate many interesting predications pertaining to the effects of words on higher-level tasks. For example, consistent with the current study, we have evidence suggesting that some of the initial differences between words and sounds on category learning stem from unfamiliar sounds (presented in isolation) hindering categorization, rather than from labels facilitating categorization (Robinson & Sloutsky, 2007). However, the current study also suggests that effects of words should change in the course of processing. For example, although words and sounds may have different effects early in the course of processing, this differential effect should attenuate across time, with infants simply requiring more exposure to arbitrary auditory–visual pairings when the auditory stimulus is unfamiliar. Auditory overshadowing may also give words a leg-up on cross-modal binding tasks such as a task of word learning. Assuming that speech is a familiar class of auditory stimuli, infants should be faster at mapping words onto objects than they are at mapping unfamiliar sounds onto objects, especially when these unfamiliar sounds are presented in isolation. While this effect could stem from infants understanding that words denote objects and categories, it could also stem from unfamiliar auditory stimuli attenuating visual processing more than words. If this is the case, then giving infants an opportunity to process an unfamiliar auditory stimulus prior to pairing it with a referent should facilitate auditory–visual binding. Testing these hypotheses is important for understanding how exposure to linguistic input influences cross-modal processing and cross-modal binding which underlie a variety of cognitive tasks, including word learning, concept acquisition, and lexical extension.

Acknowledgements

This research was supported by grants from the National Science Foundation (REC#0208103) and from the Institute for Educational Sciences, US Department of Education (R305H050125).

References

Bahrick, L.E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology*, **36**, 190–201.

Campbell, A.L., & Namy, L.L. (2003). The role of social-referential context in verbal and nonverbal symbol learning. *Child Development*, **74**, 549–563.

Christie, J., & Klein, R. (1995). Familiarity and attention: does what we know affect what we notice? *Memory and Cognition*, **23** (5), 547–550.

Colavita, F.B. (1974). Human sensory dominance. *Perception and Psychophysics*, **16**, 409–412.

Colavita, F.B., & Weisberg, D. (1979). A further investigation of visual dominance. *Perception and Psychophysics*, **25**, 345–347.

Fantz, R.L. (1964). Visual experience in infants: decreased attention to familiar patterns relative to novel ones. *Science*, **46**, 668–670.

Fernald, A., & Hurtado, N. (2006). Names in frames: infants interpret words in sentence frames faster than words in isolation. *Developmental Science*, **9**, F33–F40.

Lewkowicz, D.J. (2000). The development of intersensory temporal perception: an epigenetic systems/limitations view. *Psychological Bulletin*, **126**, 281–308.
Lickliter, R., & Bahrick, L.E. (2000). The development of infant intersensory perception: advantages of a comparative convergent-operations approach. *Psychological Bulletin*, **126**, 260–280.
Mather, E., & Schafer, G. (2005, April). Familiarity and novelty preferences for visual stimuli in the second year of life: initial findings and new directions. Poster presented at the biennial meeting of the Society for Research in Child Development, Atlanta, GA.
Namy, L.L., & Waxman, S.R. (2000). Naming and exclaiming: infants' sensitivity to naming contexts. *Journal of Cognition and Development*, **1**, 405–428.
Napolitano, A.C., & Sloutsky, V.M. (2004). Is a picture worth a thousand words? The flexible nature of modality dominance in young children. *Child Development*, **75**, 1850–1870.
Robinson, C.W., Howard, E.M., & Sloutsky, V.M. (2005). Mechanisms underlying the effects of labels on cognitive development. In B.G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the XXVII Annual Conference of the Cognitive Science Society* (pp. 1878–1882). Mahwah, NJ: Erlbaum.
Robinson, C.W., & Sloutsky, V.M. (2004a). Auditory dominance and its change in the course of development. *Child Development*, **75**, 1387–1401.
Robinson, C.W., & Sloutsky, V.M. (2004b). The effect of stimulus familiarity on modality dominance. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Conference of the Cognitive Science Society* (pp. 1167–1172). Mahwah, NJ: Erlbaum.
Robinson, C.W., & Sloutsky, V.M. (2007). Linguistic labels and categorization in infancy: do labels facilitate or hinder? *Infancy*, **11**, 233–253.
Roder, B.J., Bushnell, E.W., & Sasseville, A.M. (2000). Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy*, **1** (4), 491–507.
Rose, S.A., Feldman, J.F., & Jankowski, J.J. (2002). Processing speed in the 1st year of life: a longitudinal study of preterm and full-term infants. *Developmental Psychology*, **38**, 895–902.
Sloutsky, V.M., & Napolitano, A. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development*, **74**, 822–833.
Waxman, S.R. (2003). Links between object categorization and naming: origins and emergence in human infants. In D.H. Rakison & L.M. Oakes (Eds.), *Early category and concept development: Making sense of the blooming, buzzing confusion* (pp. 213–241). London: Oxford University Press.
Woodward, A.L., & Hoyne, K.L. (1999). Infants' learning about words and sounds in relation to objects. *Child Development*, **70**, 65–77.
Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition*, **85**, 223–250.

Received: 18 August 2006
Accepted: 21 October 2006