PAPER

Effects of auditory input in individuation tasks

Christopher W. Robinson and Vladimir M. Sloutsky

Center for Cognitive Science, The Ohio State University, USA

Abstract

Under many conditions auditory input interferes with visual processing, especially early in development. These interference effects are often more pronounced when the auditory input is unfamiliar than when the auditory input is familiar (e.g. human speech, pre-familiarized sounds, etc.). The current study extends this research by examining how auditory input affects 8- and 14-month-olds' performance on individuation tasks. The results of the current study indicate that both unfamiliar sounds and words interfered with infants' performance on an individuation task, with cross-modal interference effects being numerically stronger for unfamiliar sounds. The effects of auditory input on a variety of lexical tasks are discussed.

Introduction

Even in the earliest stages of word learning, language seems to affect behavior on a variety of cognitive tasks. For example, it is well documented that infants who hear the same word associated with different objects are more likely to group these objects together compared to infants who hear nonlinguistic input associated with the same objects (e.g. Balaban & Waxman, 1997; Fulkerson & Haaf, 2003; Fulkerson & Waxman, 2007; Robinson & Sloutsky, 2007a; but see Campbell & Namy, 2003; Roberts, 1995; and Roberts & Jacob, 1991). Linguistic input also affects performance on an individuation task: 9-month-olds who hear two words referring to two occluded objects are more likely to assume that two objects are concealed compared to infants who hear two nonlinguistic sounds (Xu, 2002; see also Xu, Cote & Baker, 2005, for similar findings). One way of explaining these reported effects is by positing that infants in the early stages of word learning already have some general understanding that words and categories are linked (Fulkerson & Waxman, 2007; Waxman, 2003; Xu, 2002).

However, some of the reported effects of words may also stem from more general factors, such as the dynamics of attention in cross-modal processing. In particular, in each of the above-mentioned tasks, infants have to attend to and encode simultaneously presented words and the objects accompanied by these words. If effects of words stem from the dynamics of cross-modal processing, then factors that facilitate and inhibit cross-modal processing should also affect performance on higher-level tasks such as categorization and individuation. The primary goal of the current study is to examine the ability of more general cross-modal processing factors to account for the differential effects of words and sounds on an individuation task.

Studies examining cross-modal processing have revealed several important findings. One set of findings is the cross-modal facilitation effect: amodal information (e.g. rate, tempo, etc.), where the same information is expressed in multiple modalities, is often easier to acquire when the information is presented cross-modally than when the same information is presented unimodally (e.g. Bahrick & Lickliter, 2000, see also Lewkowicz, 2000; Lickliter & Bahrick, 2000, for reviews). A second set of findings is the cross-modal interference effect: processing of arbitrary, auditory–visual stimuli (e.g. word–object pairings) often results in attenuated processing of one modality compared to the unimodal baseline (e.g. Casasola & Cohen, 2000; Napolitano & Sloutsky, 2004; Robinson, Howard & Sloutsky, 2005; Robinson & Sloutsky, 2004, 2007a, 2007b; Sloutsky & Napolitano, 2003; Sloutsky & Robinson, 2008; Stager & Werker, 1997). For example, in many of the studies reported above, young participants were more likely to process a visual stimulus when it was presented unimodally than when the same visual stimulus was paired with a sound or word.

More recent studies examining the processing of arbitrary, auditory–visual pairings have demonstrated that different auditory stimuli have different effects on visual processing. When young participants are given ample time to process arbitrary, auditory–visual pairings, familiar auditory stimuli, such as human speech and pre-familiarized sounds, are less likely to interfere with visual processing than unfamiliar auditory stimuli (Robinson *et al.*, 2005; Robinson & Sloutsky, 2007b;

Address for correspondence: Vladimir M. Sloutsky, Center for Cognitive Science, 208C Ohio Stadium East, 1961 Tuttle Park Place, Ohio State University, Columbus, OH 43210, USA; e-mail: sloutsky.1@osu.edu

Sloutsky & Robinson, 2008). For example, 8-month-olds in Robinson et al. (2005) and 14-month-olds in Robinson and Sloutsky (2007b) were presented with a continuous familiarization procedure. On each familiarization trial, infants were shown two simultaneously presented visual stimuli: one visual stimulus remained unchanged across 24 familiarization trials (familiar stimulus), whereas the other visual stimulus changed on every trial (novel stimulus). Visual stimuli were presented either in isolation or paired with one of the following auditory stimuli: an unfamiliar sound, a pre-familiarized sound, or a word. Visual processing speed was inferred by the amount of familiarization needed for infants to reliably demonstrate a novelty (or familiarity) preference. It was shown that infants who heard unfamiliar sounds were slower at processing the visual stimuli compared to infants who heard pre-familiarized sounds, words or no auditory input. At the same time, words did not facilitate visual processing above the silent condition. These findings suggest that the differential effects of words and unfamiliar sounds stem from unfamiliar sounds slowing down visual processing rather than from words facilitating visual processing.

Cross-modal processing and higher-order tasks

As shown above, research examining the processing of arbitrary, auditory–visual pairings demonstrates that young infants often have difficulty encoding simultaneously presented auditory and visual input, with unfamiliar sounds often attenuating visual processing more than words. Can these findings account for the differential effects of words and sounds on higher-order tasks such as categorization and individuation tasks? From a general processing account, it could be argued that cross-modal processing is a necessary component underlying many higher-order tasks. Therefore, understanding the dynamics of cross-modal processing may contribute to the understanding of infants' performance on some higher-order tasks. The current approach attempts to ground infants' performance in these tasks in the dynamics of cross-modal processing.

This issue has recently been investigated by Robinson and Sloutsky (2007a) in a study that examined how words and sounds influenced categorization compared to a silent condition. Eight- and 12-month-olds were familiarized to different exemplars from within the same category, and each member of the category was associated either with the same word, the same nonlinguistic sound, or no auditory input was provided (silent condition). After familiarization, infants were presented with novel stimuli from the *familiarized* category and with novel stimuli from a *novel* category. Categorization was inferred from differential looking to the novel category items compared to the familiarized category items. At both 8 and 12 months of age, infants were more likely to form categories in the silent condition than in the word and sound conditions. In addition, while 12-month-olds were more likely to categorize when items were accompanied by words than nonlinguistic sounds, this effect was driven by nonlinguistic sounds hindering categorization rather than from words facilitating categorization. These findings are consistent with previous research examining infants' processing of cross-modal stimuli more generally (Robinson et al., 2005; Robinson & Sloutsky, 2007b; Sloutsky & Robinson, 2008).

The goal of the current study was to examine how words and sounds affect performance on an individuation task. Previous research has demonstrated that individuation is facilitated when two objects are accompanied by unique words, but not when they are accompanied by unique sounds (Xu, 2002). However, without making comparisons to a silent condition, it is unclear whether this effect stems from words facilitating individuation or from sounds interfering with individuation. The current study addressed this issue by examining infants' performance on an individuation task across word, sound, and silent conditions.

Although we are not aware of a published study that has directly compared the effects of words on an individuation task to a silent condition, there is indirect evidence suggesting that words may facilitate individuation above a silent condition. In particular, in some experiments, 10-month-olds failed to individuate objects when the objects were presented in silence (Xu & Carey, 1996), whereas in other experiments, younger infants succeeded on a similar task when the objects were paired with two different words (Xu, 2002). However, procedural differences complicate making direct comparisons between these two studies. In particular, Xu and Carey's (1996) series of experiments (i.e. where older infants failed to individuate in the silent condition) differed from Xu's (2002) procedure in the following ways: infants were given less exposure to the to-be-individuated objects (Experiments 2–3 and 5), infants were presented with different visual stimuli (Experiments 2–5), infants participated in introductory trials prior to participating in the experiment proper (Experiments 2–5), initial preferences were assessed within-subjects (Experiments 4–5), and additional test trials were introduced (Experiment 4). Thus, different patterns of responding in Xu and Carey (1996) and Xu (2002) could stem either from exposure to words or from the other methodological differences between these studies.

Using a procedure similar to the one reported in Xu (2002), the present study directly examined how sounds and words affect performance on an individuation task compared to a silent condition. In Experiment 1, 8-month-olds were familiarized to a duck and a ball that appeared and disappeared from behind an occluder. After familiarization, the occluder dropped, revealing either two objects (expected event) or one object (unexpected event). The duck and the ball were associated either with two different words, two different sounds, or no auditory input was provided (silent condition). Experiment 2 employed a similar methodology with

8- and 14-month-olds to determine how effects of words and sounds change with age and in the course of cross-modal processing. To examine the latter issue, we reduced the amount of familiarization given to infants.

Two outcomes can be predicted based on previous research. First, it is possible that words do facilitate performance on cognitive tasks such as categorization and individuation (Balaban & Waxman, 1997; Xu, 2002). From this perspective, effects of words should exceed both the silent and sound conditions. If words do facilitate individuation, then subsequent experiments should be conducted to determine if these facilitative effects stem from exposure to speech more generally or from infants detecting contingencies between words and objects. According to the former possibility, any linguistic stimulus should facilitate individuation. According to the latter possibility, infants who hear two words paired with two objects (i.e. two-word condition) should form two distinct representations, while infants who hear one word paired with two objects (i.e. one-word condition) should be less likely to individuate the objects. There is evidence suggesting that infants do respond differently in one-word and two-word conditions (Plunkett, Hu & Cohen, 2008; Xu, 2002).

A second possible outcome stems from research that has directly compared infants' categorization performance (as well as their ability to process visual input) in the word and sound conditions to a silent condition (Robinson et al., 2005; Robinson & Sloutsky, 2007a, 2007b; Sloutsky & Robinson, 2008). In these reported studies, unfamiliar sounds interfere with processing of visual input more than words, with words failing to exert any facilitative effects compared to a silent baseline. Based on these findings, it was hypothesized that in an individuation task, words would have no facilitative effect above the silent condition, whereas nonlinguisitc sounds may elicit stronger interference than words.

Experiment 1

Method

Participants and design

Seventy-nine 8-month-olds (36 boys and 43 girls, $M = 253$ days, $SD = 10$ 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. An additional nine infants were tested but were not included in the final sample: seven due to fussiness, one due to parental interference and one due to experimenter error.

There were three between-subjects conditions: word, sound, and silent. Twenty infants heard two different words associated with the occluded objects (word condition), 20 infants heard two different sounds associated with the occluded objects (sound condition) and 24 infants received no auditory input (silent condition). Infants' performance in these conditions was compared to baseline preference, which was assessed in a separate group of infants. Fifteen infants participated in the baseline preference condition (these participants were presented only with the test trials, while not participating in the familiarization phase). Thus, infants in the baseline preference condition had no way of knowing how many objects were hidden behind the occluder.

Apparatus

Infants were seated on a parent's lap 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. Sixty percent of the data was coded offline. Offline coders concealed the half of the 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. Reliability between online and offline coders was assessed for each infant and averaged across Experiments 1 and 2, average $r = .90$.

Stimuli

Two visual images were presented in the current experiment: a duck and a ball (see Figure 1). Each stimulus was approximately 25 cm × 25 cm and appeared and disappeared from behind a 62 cm × 35 cm occluder. In the word condition, the two visual stimuli were associated with two different words (i.e. 'a duck' and 'a ball'). The words were produced by a female experimenter in infant-directed speech and were recorded as 44.1 kHz wav files. In the sound condition, the two images were associated with a laser sound and a static sound. The
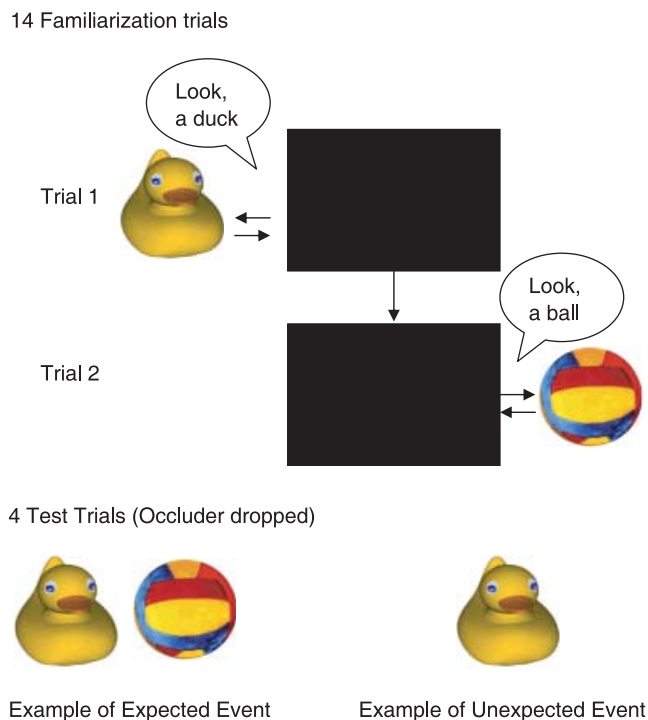
14 Familiarization trials

Figure 1   *Overview of Experiment 1.*

nonlinguistic sounds were unfamiliar to 4-year-olds (Robinson & Sloutsky, 2004), and thus it was assumed that these sounds were also unfamiliar to infants. Previous research has demonstrated that 8-month-olds ably discriminate these auditory stimuli (Robinson & Sloutsky, 2004), thus poor performance in the sound condition cannot stem from poor discrimination of the auditory stimuli. The duration of the linguistic and nonlinguistic stimuli ranged from 0.70 s to 1 s and these stimuli were presented from a Dell 8200 computer at 65–68 dB.

The auditory and visual stimuli were integrated into short animated movies using Macromedia Flash MX. Two types of movies were created: familiarization movies and testing movies. An example of a familiarization sequence is presented in the top portion of Figure 1. Throughout the familiarization phase, the same occluder was presented in the center of a projection screen. Each familiarization trial began with a beeping sound to engage the infant's attention. When the infant looked at the occluder, the beeping sound stopped and an object appeared to the left or to the right of the occluder (either a duck or a ball). The object moved up and down twice, resembling a jumping motion, and then disappeared behind the occluder. Infants in the word and sound conditions heard 'Look' when the object first appeared from behind the occluder, and they heard a word or a sound after the object jumped twice (approximately 2 s after the object first appeared from behind the occluder). Infants in the silent condition received no auditory information. The duration of familiarization movies ranged from 5 s to 16 s (see the Procedure section).

There were two different testing movies. As in familiarization movies, each trial began with a beeping sound to engage the infant's attention. When the infant looked at the occluder, the occluder dropped, revealing either a duck and a ball (expected event) or a duck (unexpected event). Each testing movie was 16 s in duration. Sounds and words were not presented in the testing movies.

Procedure

The overall procedure is presented in Figure 1. As in Xu (2002), the procedure consisted of two phases: familiarization and test. During familiarization, infants were presented with 14 familiarization trials (seven duck trials and seven ball trials). Four of the familiarization trials (trials 3, 4, 7 and 8) were extended to give infants ample time to encode the images. These extended familiarization trials were 16 s in duration, whereas the remaining familiarization trials were only 5 s in duration, with the total of 114 s of familiarization given to each infant. Infants in the word and sound conditions heard auditory input on all 14 familiarization trials, and infants in the silent condition heard no auditory information during the familiarization trials. Familiarization trials were presented in blocks of two trials (one duck trial and one ball trial) and trials were randomized within each block. As in Xu (2002), the two objects were never presented simultaneously, and thus infants could not rely on spatiotemporal information to individuate the objects. The experimenter recorded infants' fixations throughout the familiarization phase by pressing a button on a USB gamepad when infants looked at the stimulus and by releasing the button when infants looked away from the stimulus. *Presentation* recorded a time stamp at the onset of a button press (look to stimulus) and recorded a time stamp when the button was released (look away from stimulus). Fixation durations were calculated for each look (i.e. button release–button press) and total looking was calculated on each trial by summing fixation durations within a trial.

The testing phase consisted of four test trials. Two of the test trials were expected events (i.e. occluder dropped, revealing two objects), and two of the test trials were unexpected events (i.e. occluder dropped, revealing only one object). Expected and unexpected test trials alternated for each infant (e.g. expected, unexpected, expected, unexpected), with approximately half of the infants seeing an expected event as the first test trial and the remaining infants seeing an unexpected event as the first test trial. As in Xu (2002), during the testing phase, participants were shown two short re-familiarization trials prior to each test trial to remind them of the occluded objects. Each test trial and accompanying short familiarization trials had a unique occluder color (i.e. black, orange, pink, and red) to delineate the testing trials. Thus, while the occluder disappeared and changed color after each test trial, the occluder did not disappear or change color between re-familiarization and test. As

in the familiarization phase, accumulated looking was calculated on each test trial by pressing a button on the USB gamepad when infants were looking at the stimuli.

Results and discussion

Preliminary analyses focused on the mean looking times on the four extended familiarization trials and on habituation rates across the different stimulus conditions. Two means were calculated for each infant. The first mean was calculated by averaging across the first two extended familiarization trials and the second mean was calculated by averaging across the last two extended familiarization trials. These means were then submitted to a 2 (Time: First two trials vs. Last two trials) × 3 (Stimulus Condition: Silent, Word, Sound) ANOVA with Time as a repeated measure. There was a tendency for infants to look longer in the word condition ($M = 10.32$ s, $SE = 0.84$ s) and sound condition ($M = 10.54$ s, $SE = 0.79$ s) than in the silent condition ($M = 8.43$ s, $SE = 0.61$ s); however, the effect of stimulus condition did not reach significance, $F(2, 61) = 2.12$, $p = .13$. The effect of time and the time × stimulus condition interaction did not approach significance, $ps > .34$.

Primary analyses focused on infants' relative looking to the expected and unexpected events on test trials. A difference score was calculated for each infant by subtracting the mean looking time on unexpected trials from the mean looking time on expected trials. Values greater than zero reflect more accumulated looking on expected than unexpected trials and values less than zero reflect more accumulated looking on unexpected than expected trials. Three infants were considered outliers (i.e. ±2 SD from the mean) and were excluded from further analyses. Infants' individual difference scores broken up by stimulus condition are presented in Figure 2a, and mean looking times to the expected and unexpected events are presented in Figure 2b.

The difference scores were submitted to a 4 (Stimulus Condition: Silent, Word, Sound, Baseline Preference) × 2 (Test Trial Order: Expected first vs. Unexpected first) ANOVA with stimulus condition and test order manipulated between subjects. The analyses revealed a main effect of stimulus condition, $F(3, 68) = 6.60$, $p < .001$. Difference scores in each condition were compared to infants' baseline preference by using a Dunnett test, which adjusted for multiple comparisons. The results indicated that infants' relative looking in the silent condition ($M = 2.01$ s, $SE = 0.36$ s) differed from baseline preference ($M = -0.37$ ms, $SE = 0.55$ s), $p = .01$, $\eta_p^2 = 0.29$. In contrast, infants' looking in the word condition ($M = 1.16$ s, $SE = 0.41$ s), $p = .15$, $\eta_p^2 = 0.14$, and sound condition ($M = 0.30$ s, $SE = 0.80$ s), $p = .73$, $\eta_p^2 = 0.01$, did not differ from baseline preference. While the word condition probably would have reached significance with additional power, it is clear that the effect in the word condition did not exceed the silent condition. In fact, the effect for the word condition ($\eta_p^2 = 0.14$) was numerically smaller than that
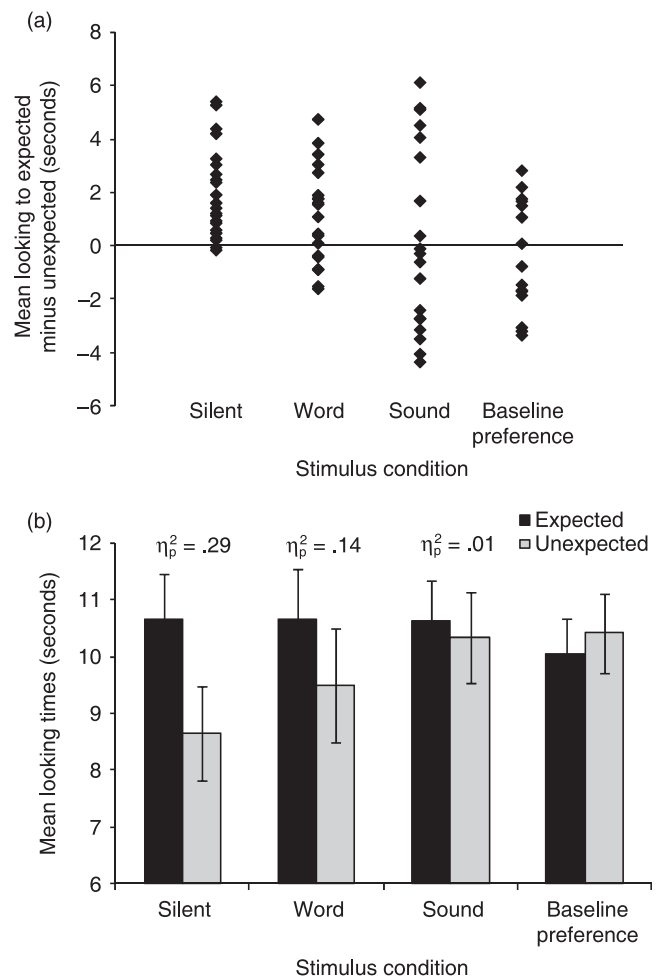

Figure 2 *Individual difference scores broken up by stimulus condition are presented in Figure 2a. Looking times to expected and unexpected events across the different stimulus conditions are presented in Figure 2b. Error bars in Figure 2b represent standard errors of the mean, and values denote effect sizes compared to baseline preference.*

for the silent condition ($\eta_p^2 = 0.29$). At the same time, the effect of the word condition was numerically larger than that of the sound condition.

Planned comparisons were also conducted comparing the silent, word and sound conditions. One-tailed *t*-tests revealed that the silent condition exceeded the sound condition, $t(39) = 2.03$, $p = .025$, and was marginally above the word condition, $t(40) = 1.55$, $p = .064$.

The above analysis also revealed an effect of test order, $F(1, 68) = 17.45$, $p < .001$, and a significant stimulus condition × test order interaction, $F(3, 68) = 3.93$, $p = .012$. These findings suggest that infants' pattern of looking differed as a function of what test item they saw first (i.e. expected trial or unexpected trial). Despite these effects, it is important to note that there was no evidence that effects of words exceeded the silent condition (see Figure 3). In particular, for those infants who saw an expected trial first, infants' looking did not differ across the stimulus conditions, one-way ANOVA, $F(3, 36) =$
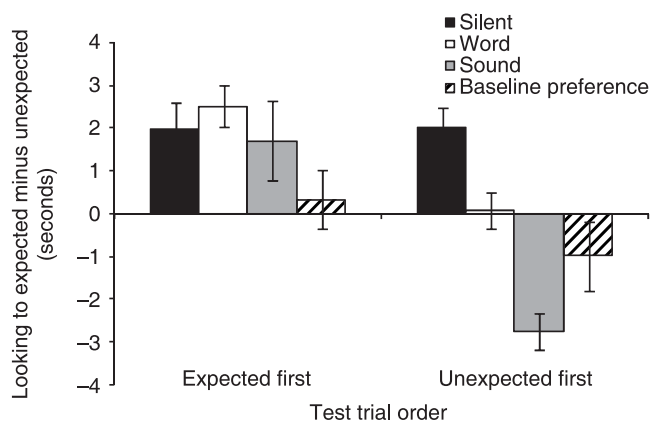
Figure 3 *Mean difference scores broken up by stimulus condition and test order. Error bars represent standard errors of the mean.*

1.15, $p = .34$. In contrast, differences between the stimulus conditions were more pronounced for those infants who saw an unexpected trial first, one-way ANOVA, $F(3, 32) = 12.80$, $p < .001$. Independent-sample t-tests with a Bonferroni adjustment indicated that the silent, word and sound conditions all differed from each other, adjusted ps $< .024$. However, only the silent condition differed from baseline preference, $t(17) = 3.52$, $p = .018$, whereas the word and sound conditions did not differ from baseline preference, ps $> .64$.

The current findings are consistent with previous research examining the effects of words and sounds on visual discrimination, visual processing speed, and categorization tasks (Robinson *et al.*, 2005; Robinson & Sloutsky, 2007a, 2007b; Sloutsky & Robinson, 2008). In particular, differences between words and sounds, when found, are likely to stem from sounds interfering with visual processing more than words, with words having no facilitative effect above the silent condition.

While the current findings are consistent with previous research examining cross-modal processing, it is important to note that infants' pattern of looking differed from the findings reported by Xu (2002). In particular, infants in the current experiment showed a preference to look at the expected event, whereas, infants in Xu (2002) showed a preference to look at the unexpected event (see General discussion for possible explanations). Given this discrepancy, we deemed it necessary to conduct Experiment 2 to examine the replicability of the observed pattern.

Experiment 2

The current experiment had four goals. As mentioned above, one goal was to ensure that the overall pattern of results in Experiment 1 was replicable. To achieve this goal we presented infants in the current experiment with the duck–ball trials that were used in Experiment 1. A

second goal was to assess the generalizability of the results of Experiment 1 by examining whether these results are restricted to events that consist of familiar objects and familiar words, such as duck and ball. To achieve this goal, we introduced new events that consisted of novel entities and novel words. A third goal of Experiment 2 was to examine how effects of auditory input on individuation tasks change with age. To achieve this goal, the current experiment examined how 8- and 14-month-olds responded to the same events.

The final goal was to examine how effects of auditory input change in the course of processing. To account for some of the effects of words on a variety of tasks, we have proposed a set of factors underlying the dynamics of attention in cross-modal processing (Robinson & Sloutsky, 2007a, 2007b; Sloutsky & Robinson, 2008). These hypothetical factors are: (a) auditory input is often faster to engage attention than visual input, especially early in development and (b) processing of the details of a visual stimulus may not start until the auditory stimulus is fully processed (i.e. cross-modal processing is serial in nature). These ideas generate novel and interesting predictions pertaining to how words and sounds affect performance on a variety of tasks and how these effects change in the course of processing. For example, assuming (a) and (b), one hypothesis that can be tested is that words and sounds should elicit stronger cross-modal interference early in the course of processing, especially if there is not enough time for the auditory modality to release attention. This issue was addressed in Experiment 2 by shortening the amount of familiarization.

Method

Participants, design and apparatus

Eleven 8-month-olds (five boys and six girls, $M = 263$ days, $SD = 26$ days) and 10 14-month-olds (five boys and five girls, $M = 432$ days, $SD = 63$ days) participated in this experiment. Subject recruitment and demographics were identical to Experiment 1. An additional four infants were tested but were not included in the final sample due to fussiness. As mentioned above, the amount of familiarization was reduced in the current experiment. This manipulation was not only necessary to examine how infants respond to visual stimuli in the early stages of cross-modal processing but was also implemented to make the task more engaging for the older infants: recall that infants in Experiment 1 were given 114 s of familiarization to two objects before reaching the testing phase. The reduction of familiarization enabled us to present participants with different types of events (i.e. duck–ball trials as well as with novel images and words), thus assessing the generalizability of the findings in Experiment 1. In the current experiment, stimulus condition (silent, word, sound, baseline preference), test trial (expected vs. unexpected) and event type (duck–ball vs. mecki–nobi) were manipulated within subjects, and

Table 1 *Overview of Experiment 2. Stimulus condition (silent, word, sound, baseline preference), Trial type (expected and unexpected) and Event type (duck–ball vs. mecki–nobi) were manipulated within subjects*

Stimulus condition	16 Testing trials			
	8 duck–ball events		8 mecki–nobi events	
Silent	1 Expected Trial	1 Unexpected Trial	1 Expected Trial	1 Unexpected Trial
Word	1 Expected Trial	1 Unexpected Trial	1 Expected Trial	1 Unexpected Trial
Sound	1 Expected Trial	1 Unexpected Trial	1 Expected Trial	1 Unexpected Trial
Baseline preference	1 Expected Trial	1 Unexpected Trial	1 Expected Trial	1 Unexpected Trial

age was a between-subjects factor (see Table 1 for overview of Experiment 2). The apparatus was identical to Experiment 1.
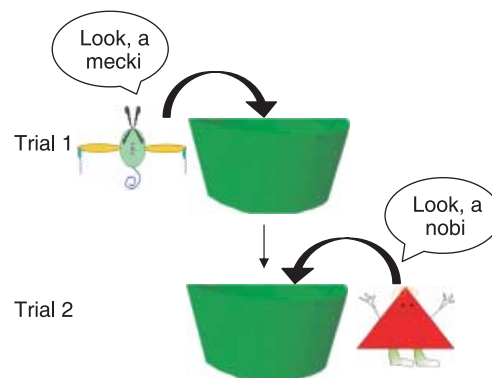
Stimuli and procedure

To ensure that the findings in Experiment 1 are replicable, the duck–ball trials from Experiment 1 were presented to infants in Experiment 2. The duck–ball familiarization trials were identical to Experiment 1 (see Figure 1), except that the duration of each familiarization trial was reduced to 3 s. Recall that none of the familiarization trials in Experiment 1 was shorter than 5 s. The words that infants heard were identical to Experiment 1 (i.e. 'a duck' and 'a ball'). To ensure that poor performance in the sound condition of Experiment 1 was not stimulus specific, two new nonlinguistic sounds were created by using preset functions in Cool Edit 2000 (chord and out of control). The sounds and words were all dynamic in that they changed in pitch and amplitude across time. As in Experiment 1, auditory stimuli ranged from 0.70 s to 1 s and were presented at 65–68 dB.

There were also two new types of events, which consisted of novel objects and novel words (see Figure 4 for stimuli and overview of these events). On each familiarization trial, a visual stimulus appeared on one side of the screen, the novel animal-like creature bounced twice, and then jumped into the green container. The second familiarization trial was similar except that a different animal appeared on the opposite side of the screen, bounced twice, and then jumped into the same green container. On word trials, each object was associated with a different word ('a mecki' and 'a nobi'). The words were spoken by a female experimenter in infant-directed speech and saved as 44.1 kHz wav files. The nonlinguistic sounds were created by using preset functions in Cool Edit 2000 (DTMF signal and out of control). The sounds and words were dynamic in that they changed in pitch and amplitude across time. As in the duck–ball trials, each mecki–nobi familiarization trial was 3 s.

In the current experiment, familiarization and testing trials were blocked (two familiarization trials → one test trial) and each infant received 16 of these familiarization → testing blocks. Thus, in contrast to Experiment 1 where infants were given four test trials, infants in the current experiment received 16 test trials (see Table 1 for

Example of Familiarization Trials

Trial 1

"Look, a mecki"

Trial 2

"Look, a nobi"

Example of Expected Test Trial   Example of Unexpected Test Trial

Figure 4 *Examples of mecki–nobi trials presented in Experiment 2.*

test trials). The two familiarization trials preceding each test trial were combined into a single familiarization movie (hereafter, familiarization phase). Each familiarization phase began with a red curtain that concealed the entire screen; the red curtain was animated using Macromedia Flash MX and it took 1 s for the curtain to fully open. The opening of the curtain revealed either a black occluder (on duck–ball trials) or a green container (on mecki–nobi trials). Infants were then presented with two familiarization trials (see Figures 1 and 4 for examples). After the second familiarization trial, the occluder began to drop (on duck–ball trials) or the green container began to rotate 180° (on mecki–nobi trials). Looking times were recorded during the entire familiarization phase (approximately 12 s in duration) by pressing a button on a USB response pad when infants were looking at the stimuli and by releasing the button when infants looked away. As in Experiment 1, *Presentation* recorded a time stamp at the onset of a button press (look to stimulus)

and recorded a time stamp when the button was released (look away from stimulus). Fixation durations were calculated for each look (i.e. button release–button press) and total looking was calculated on each trial by summing fixation durations within a trial.

After each pair of familiarization trials, infants were given one test trial. On duck–ball trials, infants saw either one visual stimulus (unexpected test trial) or two visual stimuli (expected test trial). On mecki–nobi trials, either one animal fell out of the container (unexpected test trial) or two animals fell out of the container (expected test trial). Each test trial was 10 s in duration and looking times were recorded by pressing a button on a USB response pad when infants were looking at the stimuli. As in Experiment 1, no auditory input was provided at test. After the 10 s test trial, the red curtain slowly closed, which marked the end of the trial. The next familiarization → testing block began when the red curtain began to open. The order of the familiarization → testing blocks was randomized for each infant. Infants' looking times in the silent, sound, and word conditions were compared to their own baseline preference. Baseline preference trials were identical to the test trials in the other conditions except that infants did not see the familiarization phase prior to test. Baseline preference trials were randomly intermixed with the silent, sound, and word trials in the experiment proper.

Results and discussion

Preliminary analyses focused on the accumulated looking during familiarization. Recall that each familiarization phase was approximately 12 s; thus, habituation analyses were not conducted in the current experiment. Fixation durations were summed within each familiarization phase and means were calculated for each stimulus condition and for each event type. A 3 (Stimulus Condition: Silent, Word, Sound) × 2 (Event Type: Duck–ball vs. Mecki–nobi) × 2 (Age: 8 months vs. 14 months) mixed ANOVA revealed a significant effect of stimulus condition, $F(2, 38) = 6.55$, $p = .004$. Infants accumulated more looking in the word condition ($M = 10.66$ s, $SE = 0.24$ s) and sound condition ($M = 10.64$ s, $SE = 0.21$ s) than in the silent condition ($M = 10.10$ s, $SE = 0.23$ s), paired-sample ts > 2.89, ps < .01. No difference was found between the word and sound conditions, $t(20) = 0.16$, $p = .87$. The above analysis also revealed an effect of event type, $F(1, 19) = 29.60$, $p < .001$, with infants accumulating more looking on mecki–nobi events ($M = 10.91$ s, $SE = 0.22$ s) than on duck–ball events ($M = 10.03$ s, $SE = 0.22$ s).

Primary analyses focused on infants' relative looking to the expected and unexpected events at test across the four stimulus conditions (silent, word, sound, and baseline preference) and across the two event types (duck–ball vs. mecki–nobi). Eight difference scores (mean looking to expected – mean looking to unexpected) were calculated for each infant. Six responses (4% of the total) were ±2 SD from the mean and were considered outliers. Rather
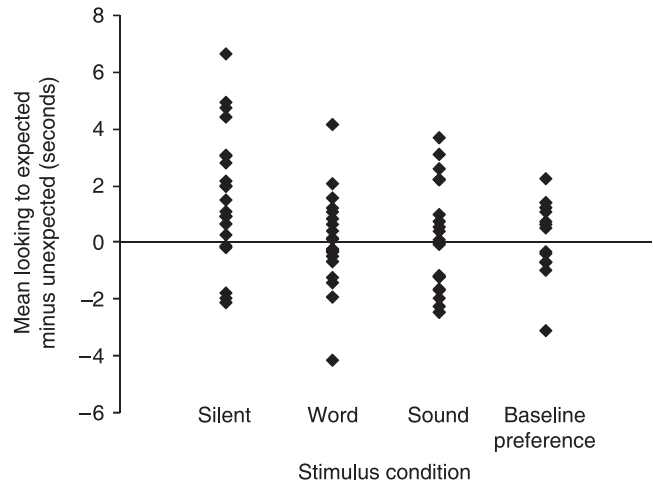

Figure 5 *Individual difference scores broken up by stimulus condition in Experiment 2.*

than exclude the infant from the entire sample, each outlier was replaced by the mean from the corresponding condition. The difference scores were submitted to a 4 (Stimulus Condition: Silent, Word, Sound, Baseline Preference) × 2 (Event Type: Duck–ball vs. Mecki–nobi) × 2 (Age: 8 months vs. 14 months) mixed ANOVA, with stimulus condition and event type being within-subjects variables. Given that there was no significant effect of age, nor did age interact with other variables (all ps > .22), the two age groups were collapsed.

The analyses revealed a main effect of stimulus condition, $F(3, 57) = 3.99$, $p = .012$ (see Figure 5 for individual difference scores broken up by stimulus condition). Difference scores in each condition were then compared to infants' baseline preference by conducting paired-sample t-tests with a Bonferroni adjustment. The results indicated that infants' looking in the silent condition ($M = 1.67$ s, $SE = 0.51$ s) differed from baseline preference ($M = -0.26$ s, $SE = 0.29$ s), $t(20) = 3.08$, $p = .018$, $\eta_p^2 = .32$. The word condition ($M = 0.05$ s, $SE = 0.36$ s), $\eta_p^2 = .03$, and sound condition ($M = 0.14$ s, $SE = 0.40$ s), $\eta_p^2 = .02$, did not differ from baseline preference, ps > .90. Planned comparisons were also conducted comparing the silent, word and sound conditions. The silent condition exceeded the word and sound conditions, ts > 2.40, ps < .05, and the sound and word conditions did not differ from each other, $t < 1$.

The above analyses also revealed an effect of event type, $F(1, 19) = 4.15$, $p = .056$. Infants accumulated more looking to expected trials on mecki–nobi events ($M = 0.74$ s, $SE = 0.23$ s) than on duck–ball events ($M = 0.06$ s, $SE = 0.22$ s). Despite this overall difference, however, the stimulus condition × event type interaction did not approach significance, $p = .79$. As can be seen in Figure 6a (duck–ball events) and Figure 6b (mecki–nobi events), the same overall pattern emerged across the two event types, with effects in the silent condition always being larger than the word and sound conditions.
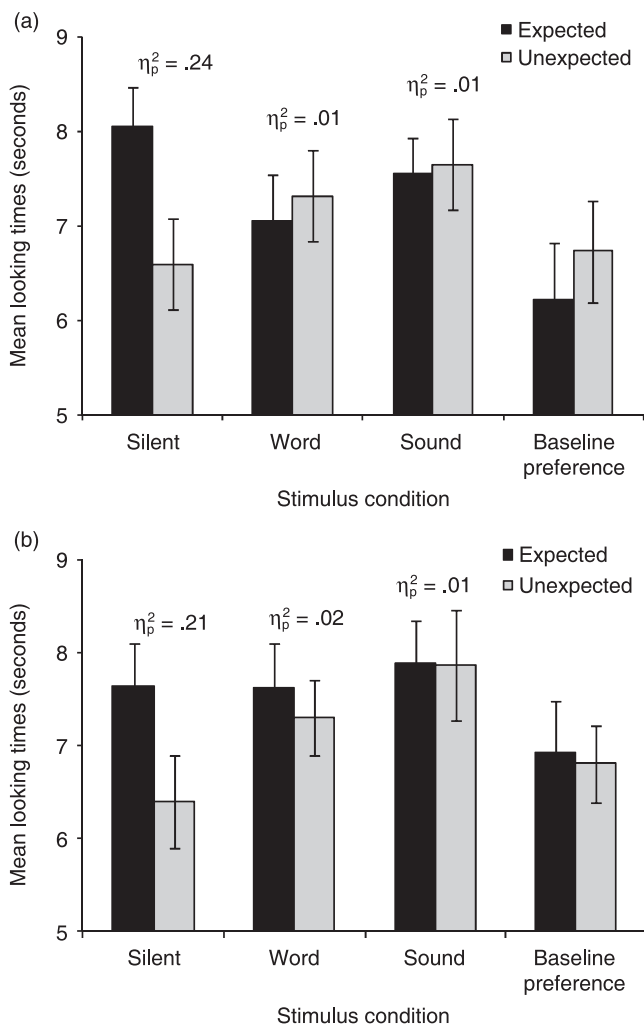
Figure 6 *Looking times to expected and unexpected events on duck–ball events are presented in Figure 6a. Looking times to expected and unexpected events on mecki–nobi events are presented in Figure 6b. Error bars in Figures 6a and 6b represent standard errors of the mean and values denote effect sizes compared to baseline preference.*

To determine how effects of words and sounds changed in the course of processing, effect sizes on duck–ball events were compared in Experiments 1 and 2 (Figures 2b and 6a, respectively). Early in the course of processing when infants were given only 12 s of familiarization to the duck–ball events (Experiment 2), both unfamiliar sounds ($\eta_p^2 = .01$) and words ($\eta_p^2 = .01$) had small and comparable effects relative to infants' baseline preferences. Furthermore, these effect sizes were considerably smaller compared to the silent condition ($\eta_p^2 = .24$). In Experiment 1 infants were given 114 s of familiarization to the duck–ball events before they reached the testing phase. Under these testing conditions, differences between words ($\eta_p^2 = .14$) and sounds ($\eta_p^2 = .01$) became more pronounced. However, neither condition exceeded the silent condition ($\eta_p^2 = .29$).

In summary, despite the procedural differences between the two reported experiments, Experiment 2 replicated

the overall pattern of results found in Experiment 1. Infants who saw the images presented without an auditory stimulus changed their pattern of looking at test compared to baseline preference, and infants who heard sounds did not change their pattern of looking compared to initial preference. While words had stronger effects in Experiment 1 than in Experiment 2, the effect did not exceed the silent condition in either experiment.

General discussion

The current study examined the effects of words and sounds on individuation tasks in 8- and 14-month-old infants. In both reported experiments, familiarizing infants to two images in silence changed infants' looking behaviors compared to baseline preference, whereas familiarizing infants to the same two images paired with unfamiliar sounds did not appear to have an effect on infants' looking behaviors. Furthermore, effects in the silent condition always exceeded the sound condition, which suggests that the presence of unfamiliar sounds hindered infants' performance in the current tasks. These effects were robust, present at various ages and replicated across varying stimulus conditions and methodologies. At the same time, effects of words appeared to vary across the conditions. Under short familiarization procedures, infants' looking in the word condition did not differ from baseline preference and effects of words were comparable to the unfamiliar sound condition and were significantly below the silent condition (Experiment 2). When infants were given more time to process the two images, effects of words increased (Experiment 1). However, even with extended familiarization, there was no evidence that effects of words exceeded the silent condition.

The current findings are consistent with previous research examining cross-modal processing and effects of auditory input on categorization (Robinson *et al.*, 2005; Robinson & Sloutsky, 2007a, 2007b; Sloutsky & Robinson, 2008). Furthermore, the current findings in conjunction with previous research on cross-modal processing are likely to elucidate performance on a wide range of tasks that hinge on the processing of arbitrary, auditory–visual pairings.

There are many tasks that hinge on the processing of arbitrary, auditory–visual pairings and these tasks require infants to attend, encode, store, and retrieve simultaneously presented auditory and visual information. This process is often difficult for infants (and even preschoolers), especially when the information consists of arbitrarily presented auditory–visual pairings such as word–object and sound–object pairings (e.g. Casasola & Cohen, 2000; Napolitano & Sloutsky, 2004; Robinson *et al.*, 2005; Robinson & Sloutsky, 2004, 2007a, 2007b; Sloutsky & Napolitano, 2003; Sloutsky & Robinson, 2008; Stager & Werker, 1997). The ability to process cross-modal stimuli changes across development (Robinson

& Sloutsky, 2004; Sloutsky & Napolitano, 2003) as well as under different stimulus conditions (Robinson *et al.*, 2005; Robinson & Sloutsky, 2007a, 2007b; Sloutsky & Robinson, 2008).

When infants are given ample time to process cross-modal stimuli, familiar auditory stimuli (e.g. pre-familiarized sounds and human speech) are less likely to attenuate visual processing than unfamiliar auditory stimuli (Robinson *et al.*, 2005; Robinson & Sloutsky, 2007a, 2007b; Sloutsky & Robinson, 2008). While additional research is needed, these published findings and the results of Experiment 1 are consistent with the idea that processing of the details of a visual stimulus may not start until the auditory stimulus is fully processed. Assuming that familiar auditory stimuli are processed more efficiently than unfamiliar auditory stimuli, it is possible that familiar auditory stimuli are simply faster to release attention, thus giving infants more time to process a corresponding visual stimulus.

The current study also suggests that the duration of familiarization may play an important role in cross-modal processing. Under short familiarization procedures such as the ones employed in Experiment 2, infants may not have time to fully disengage from the auditory stimulus. Under these conditions, both familiar and unfamiliar auditory stimuli should attenuate visual processing compared to a silent condition. In contrast, when infants are given more time to process cross-modal stimuli (e.g. Experiment 1), auditory stimuli that are faster to release attention should elicit less cross-modal interference than auditory stimuli that take longer to process. This theoretical approach accounts for the current findings, as well as for some earlier reported findings. At the same time, the current approach does not assume that infants understand the conceptual importance of words or assume that infants have the ability to control their attention and deliberately focus on these conceptually important features. Despite remaining challenges (these are discussed in the Unresolved issues section), the proposed theoretical approach appears promising for understanding the role of language in conceptual development.

Language and conceptual development

What is the role of language in higher-level tasks such as categorization and individuation? It is well established that non-human species such as rhesus monkeys and pigeons appear to have little difficulty on categorization and individuation tasks (Bhatt, Wasserman, Reynolds & Knauss, 1988; Phillips & Santos, 2007; Santos, Sulkowski, Spaepen & Hauser, 2002; Uller, Xu, Carey & Hauser, 1997). Furthermore, pre-linguistic infants as young as 3–5 months of age ably form perceptual categories and can individuate objects when the objects are presented without linguistic input (e.g. Quinn, Eimas & Rosenkrantz, 1993; Wilcox & Baillargeon, 1998a, 1998b). These findings suggest that language is not necessary to categorize and individuate objects in the environment.

Although language may not be necessary, it could be argued that language facilitates performance on these tasks. However, results from studies directly comparing effects of words to a silent condition (Robinson *et al.*, 2005; Robinson & Sloutsky, 2007a, 2007b) indicate that this is not the case: there is no evidence in these above-mentioned studies indicating that linguistic input facilitates visual processing, visual discrimination, or categorization in the early stages of word learning (but see Roberts, 1995). These results highlight a novel theoretical account pertaining to the mechanisms underlying the effects of words on conceptual tasks and change in these mechanisms across development.

Recall that according to one theoretical account, even at the onset of word learning infants have assumptions that words and categories are linked, with words (but not sounds) directing infants' attention to the category-relevant information (Waxman, 2003). Support for this claim comes from the finding that words and sounds often have different effects on categorization and individuation tasks (Balaban & Waxman, 1997; Fulkerson & Waxman, 2007; Xu, 2002). Another possibility (tested in the current research) is that cross-modal processing is a necessary component underlying many cognitive tasks and some of the effects of auditory input on these tasks should be predicted by knowing how infants process cross-modal stimuli more generally. According to this account, words and sounds hinder visual discrimination by slowing down visual processing early in development (Robinson & Sloutsky, 2004, 2007b; Sloutsky & Robinson, 2008). Based on these ideas, it was hypothesized that words and sounds may also attenuate performance on categorization and individuation tasks. This hypothesis has received support in the Robinson and Sloutsky (2007a) study (which focused on a categorization task) and in the current study (which focused on an individuation task).

Unresolved issues

While the current approach is promising, there are several unresolved issues that need to be addressed in future research. First, it is unclear why infants in both reported experiments showed a preference to look to the expected event, whereas infants in Xu (2002) exhibited a preference to look to the unexpected event. We believe that the dynamics of processing may account for these differences. In particular, it is well known that infants often show a familiarity preference early in the course of processing, with younger infants (who are slower at processing) being more likely to show a familiarity preference than older infants (Hunter & Ames, 1988; Hunter, Ames & Koopman, 1983; Roder, Bushnell & Sasseville, 2000; Rose, Gottfried, Melloy-Carminar & Bridger, 1982). Recall that infants in the current study were somewhat younger than those reported in Xu (2002) and were also given less time to process the information (Experiment 2). Therefore, it is likely that infants in the current study

were in the early stages of processing (a lack of significant decline in looking during familiarization of Experiment 1 supports this possibility). Thus, it is possible that infants' preference for the expected events stemmed from a familiarity preference – assuming that expected events are more familiar than unexpected events.

A second issue concerns whether individuation tasks such as the ones reported here can provide definitive answers about the process of individuation *per se*. In individuation tasks, and violation-of-expectation tasks more generally, it is often unclear what these tasks are actually measuring (see Bogartz, Shinskey & Schilling, 2000; Cashon & Cohen, 2000; Schilling, 2000, for similar claims). We briefly review different ways that infants could be processing the current tasks and then discuss how words and sounds may have influenced these processes.

First, it is possible that infants in the current study learned the identity of one of the occluded objects. It is also possible that infants learned that two different individuals were hidden behind the occluder (i.e. individuation proper). Finally, it is possible that infants enumerated the objects – they encoded the number of occluded objects independent of the objects' identities. In all these situations, infants should have been surprised on unexpected trials either because the encoded object disappeared, one of the two individuals disappeared, or because the number of objects changed, respectively. If infants in the current study were novelty responders, then they should have increased looking when presented with these unexpected events. However, if infants were familiarity responders, then they should have accumulated more looking on expected events (i.e. expected events were more similar to familiarization). Infants in the silent condition showed a reliable preference for the expected events, and this preference differed from baseline preferences. In contrast, infants in the word and sound conditions did not show a reliable preference and their looking at test did not reliably differ from baseline preferences. This finding suggests that the presence of words and sounds had no facilitative effect on encoding of visual input (the process that is common for object recognition, individuation and enumeration).

Another interesting issue that will require future research is that infants who hear a single word associated with multiple exemplars (i.e. one-word condition) often respond differently than those infants who hear two words associated with the same exemplars (i.e. two-word condition). For example, in Xu (2002), 9-month-olds who heard two unique words (e.g. 'a duck' and 'a ball') associated with the two occluded objects performed differently from infants who heard the same word associated with the two objects (i.e. 'a toy'). More recently Plunkett *et al.* (2008) presented 10-month-olds with a category learning task, such that a broad to-be-learned category consisted of two narrow subcategories. When members of the to-be-learned category were accompanied by two different words (i.e. each subcategory was accompanied by a separate word), infants learned two categories (Experiment 3). In contrast, when members of both subcategories were accompanied by the same word, infants learned the broader category (Experiment 5).

Why is there a difference between one- and two-word conditions? A tentative answer was offered by Plunkett *et al.* (2008). They argued that when words correlate with category-important features or relations, these correlations affect the extraction of these features or relations. If this is the case, the difference between one-word and two-word conditions is a function of forming word–object associations, which is a function of cross-modal processing. Therefore, it could be predicted that given sufficient familiarization time, the difference between two-words and one-word conditions should be isomorphic to that between two-sounds and one-sound conditions. The current study did not find evidence that the two-word condition facilitated individuation; thus, a one-word condition was not included in the design. Therefore, the hypothesis that the difference between two-word and one-word conditions is a function of cross-modal processing has to be tested in future research.

Despite these unresolved issues, the current study has several important implications. First, auditory input is likely to interfere with visual processing, thus attenuating performance on cognitive tasks compared to a silent condition. Second, sounds exert stronger interference than words: recall that sounds had very small effects in both reported experiments, whereas effects of words became more pronounced under longer familiarization procedures (Experiment 1). This latter finding suggests that effects of words change in the course of processing and that testing at a single point in the course of processing will only partially reveal how words affect performance on higher-order tasks.

Concluding comments

The reported research provides several definitive answers. The current study in conjunction with Robinson and Sloutsky (2007a) indicates that under many conditions, linguistic input does not facilitate performance on cognitive tasks above a silent baseline. Furthermore, effects of auditory input appear to change in the course of processing, with nonlinguistic sounds hindering performance under extended study-phase conditions, and linguistic and nonlinguistic input hindering performance under short study-phase conditions. While the proposed approach can readily account for these findings, these findings pose some problems for language-specific approaches. For example, it is unclear how infants' appreciation for the link between words and concepts can result in words hindering performance on categorization and individuation tasks. It is also unclear why these effects change in the course of processing. Examining the conditions that give rise to cross-modal facilitation and interference will be important for understanding the mechanism(s) underlying effects of words

on a variety of tasks, and it is likely that some of the initial effects of words will be accounted for by understanding how infants allocate their attention to cross-modal stimuli more generally.

In summary, even in the earliest stages of word learning, linguistic and nonlinguistic input can have different effects on a variety of cognitive tasks. These effects have been explained by assuming that infants understand the conceptual importance of words. For these assumptions to affect performance, infants also need to have the ability to focus their attention on the conceptually important features, while ignoring conceptually non-important features. While some of these assumptions may turn out to be true, the current study takes a different approach by attempting to ground these complex behaviors in the component processes underlying these tasks: processing of arbitrary, auditory–visual pairings. This latter approach can account for the differential effects of words and sounds on categorization (Robinson & Sloutsky, 2007a) and can explain some of the effects of auditory input on individuation tasks (current study).

Acknowledgements

References

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

Balaban, M.T., & Waxman, S.R. (1997). Do words facilitate object categorization in 9-month old infants? *Journal of Experimental Child Psychology*, **64**, 3–26.

Bhatt, R.S., Wasserman, E.A., Reynolds, W.F., Jr., & Knauss, K.S. (1988). Conceptual behavior in pigeons: categorization of both familiar and novel examples from four classes of natural and artificial stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, **14**, 219–234.

Bogartz, R.S., Shinskey, J.L., & Schilling, T.H. (2000). Object permanence in five-and-a-half-month-old infants? *Infancy*, **1**, 403–428.

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

Casasola, M., & Cohen, L.B. (2000). Infants' association of linguistic labels with causal actions. *Developmental Psychology*, **36**, 155–168.

Cashon, C.H., & Cohen, L.B. (2000). Eight-month-old infants' perception of possible and impossible events. *Infancy*, **1**, 429–446.

Fulkerson, A.L., & Haaf, R.A. (2003). The influence of labels, non-labeling sounds, and source of auditory input on 9- and 15-month-olds' object categorization. *Infancy*, **4**, 349–369.

Fulkerson, A.L., & Waxman, S.R. (2007). Words (but not tones) facilitate object categorization: evidence from 6- and 12-month-olds. *Cognition*, **105**, 218–228.

Hunter, M.A., & Ames, E.W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In L.P. Lipsitt (Ed.), *Advances in child development and behavior* (pp. 69–95). New York: Academic Press.

Hunter, M.A., Ames, E.W., & Koopman, R. (1983). Effects of stimulus complexity and familiarization time on infant preferences for novel and familiar stimuli. *Developmental Psychology*, **19**, 338–352.

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.

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.

Phillips, W., & Santos, L.R. (2007). Evidence for kind representations in the absence of language: experiments with rhesus monkeys (*Macaca mulatta*). *Cognition*, **102**, 455–463.

Plunkett, K., Hu, J.F., & Cohen, L.B. (2008). Labels can override perceptual categories in early infancy. *Cognition*, **106**, 665–681.

Quinn, P.C., Eimas, P.D., & Rosenkrantz, S.L. (1993). Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. *Perception*, **22**, 463–475.

Roberts, K. (1995). Categorical responding in 15-month-olds: influence of the noun-category bias and the covariation between visual fixation and auditory input. *Cognitive Development*, **10** (1), 21–41.

Roberts, K., & Jacob, M. (1991). Linguistic versus attentional influences on nonlinguistic categorization in 15-month-old infants. *Cognitive Development*, **6**, 355–375.

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. (2004). Auditory dominance and its change in the course of development. *Child Development*, **75**, 1387–1401.

Robinson, C.W., & Sloutsky, V.M. (2007a). Linguistic labels and categorization in infancy: do labels facilitate or hinder? *Infancy*, **11**, 233–253.

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

Roder, B.J., Bushnell, E.W., & Sasseville, A.M. (2000). Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy*, **1**, 491–507.

Rose, S.A., Gottfried, A.W., Melloy-Carminar, P., & Bridger, W.H. (1982). Familiarity and novelty preferences in infant recognition memory: implications for information processing. *Developmental Psychology*, **18**, 704–713.

Santos, L.R., Sulkowski, G.M., Spaepen, G.M., & Hauser, M.D. (2002). Object individuation using property/kind information in rhesus macaques (*Macaca mulatta*). *Cognition*, **83**, 241–264.

Schilling, T.A. (2000). Infants' looking at possible and impossible screen rotations: the role of familiarization. *Infancy*, **1**, 389–402.

Sloutsky, V.M., & Napolitano, A.C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development*, **74**, 822–833.

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

Stager, C.L., & Werker, J.F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, **388**, 381–382.

Uller, C., Xu, F., Carey, S., & Hauser, M. (1997). Is language needed for constructing sortal concepts? A study with nonhuman primates. *Proceedings of the 21st annual Boston University Conference on Language Development* (vol. 21, pp. 665–677). Somerville, MA: Cascadilla Press.

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.

Wilcox, T., & Baillargeon, R. (1998a). Object individuation in infancy: the use of featural information in reasoning about occlusion events. *Cognitive Psychology*, **37**, 97–155.

Wilcox, T., & Baillargeon, R. (1998b). Object individuation in young infants: further evidence with an event-monitoring paradigm. *Developmental Science*, **1**, 127–142.

Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition*, **85**, 223–250.

Xu, F., & Carey, S. (1996). Infants' metaphysics: the case of numerical identity. *Cognitive Psychology*, **30**, 111–153.

Xu, F., Cote, M., & Baker, A. (2005). Labeling guides object individuation in 12-month-old infants. *Psychological Science*, **16**, 372–377.