Categorization in Infancy: When Sounds and Labels Hinder Category Learning

Christopher W. Robinson (robinson.777@osu.edu)
Center for Cognitive Science
The Ohio State University
208F Ohio Stadium East, 1961 Tuttle Park Place
Columbus, OH 43210, USA

Vladimir M. Sloutsky (sloutsky.1@osu.edu)
Center for Cognitive Science
The Ohio State University
208C Ohio Stadium East, 1961 Tuttle Park Place
Columbus, OH 43210, USA

Abstract

Why do linguistic labels facilitate category learning more than other types of auditory input? From one perspective, it has been argued that effects of labels stem from young infants having broad assumptions that labels and categories are linked. From a different perspective, it is possible that both labels and sounds deplete attentional resources and hinder category learning, with sounds attenuating visual processing more than linguistic labels. The current study investigated effects of auditory input on categorization in 12-month-old infants, and the findings are consistent with the latter view: compared to a silent condition, both linguistic labels and nonlinguistic sounds hindered category learning in young infants, and differential effects of labels and sounds stemmed from nonlinguistic sounds interfering with category learning more than labels. The role of auditory input in a variety of cognitive tasks is discussed.

Keywords: Cognitive Development, Attention, Language Acquisition, Psychology, Human Experimentation.

Introduction

The ability to form categories by treating discriminable entities as members of an equivalence class is an important component of human cognition (see Murphy, 2002 for a review). This ability appears very early in development (e.g., Oakes, Madole, & Cohen, 1991; Quinn, Eimas, & Rosenkrantz, 1993; Younger & Furrer, 2003). Furthermore, exposure to linguistic input seems to facilitate category learning above and beyond exposure to non-linguistic auditory input (see Waxman, 2003 for a review).

Why do labels facilitate categorization more than non-linguistic auditory input? According to the language-specific view, even prelinguistic infants are "equipped with a broad, universally shared expectation, linking words to commonalities among objects," (Waxman, 2003, p. 220). Thus, at the onset of word learning, linguistic labels and nonlinguistic sounds have different effects on cognitive tasks because infants have some general understanding of the conceptual importance of labels (e.g., Balaban & Waxman, 1997; Xu, 2002; see also Waxman, 2003 for a review). For example, in Balaban & Waxman's (1997) study, 9-month-old infants were familiarized to a category and the to-be-categorized entities were associated with the same label, the same content-filtered speech, or the same nonlinguistic sound. After familiarization, infants were presented with two novel entities: one from the familiar category and one from a novel category. If infants formed a category during familiarization, then they should perceive a novel entity from a familiar category as old and show a reliable preference to look to the novel category. The findings demonstrate that infants who heard speech (label or content-filtered) were more likely to look to the novel category than infants who heard sounds.

However, there are recent findings suggesting that effects of labels may stem from more general (rather than language-specific) factors. For example, there is evidence that both labels and sounds often overshadow (i.e., attenuate processing of) visual input (Napolitano & Sloutsky, 2004; Robinson, Howard, & Sloutsky, 2005; Robinson & Sloutsky, 2004a; 2004b; Sloutsky & Napolitano, 2003). Therefore, it is possible that both labels and sounds result in less efficient category learning than a no-sound condition.

If this is the case, how could the overshadowing hypothesis explain the earlier found effects of labels on categorization (e.g., Balaban & Waxman, 1997)? One possibility is that the familiarity of the auditory input can account for the degree and likelihood of overshadowing effects. It has been shown that unfamiliar auditory input exerts greater overshadowing effects than more familiar auditory input (Robinson, Howard, & Sloutsky, 2005; Robinson & Sloutsky, 2004b). Given that speech sounds of a native language are more familiar than completely novel sounds, it seems plausible that after multiple exposures labels would exert weaker overshadowing effects on visual input than unfamiliar sounds.

The goal of the current study was to further examine overshadowing effects by examining how auditory input affects category learning in young infants. Infants in the current task were familiarized to different members of a

category (i.e., cat), and during familiarization each individual cat was associated with the same auditory stimulus (either sound or word) or each individual cat was presented without an auditory stimulus (i.e., silent condition). After familiarization, infants were simultaneously presented with two novel stimuli (i.e., novel cat and novel bear). If infants formed a category during familiarization, then they should perceive a novel cat as old and look longer to the novel category (i.e., compared to infants' initial preference). Furthermore, if labels facilitate categorization then infants in the label condition should accumulate more relative looking to the novel category at test compared to infants in the silent and sound conditions. However, if auditory input interferes with category learning, then nonlinguistic sounds and linguistic labels should hinder categorization compared to the silent condition. It was hypothesized that both words and sounds would hinder categorization and effects of labels over sounds would stem from sounds hindering category learning more than words.

Method

Participants Seventy-one 12-month-olds (45 boys and 26 girls, M = 373 days, SD = 13 days) participated in this experiment. A majority of infants were Caucasian and none of the infants had auditory or visual deficits, as reported by parents. Eighteen infants heard no auditory input during familiarization (i.e., silent condition), 21 infants heard the same nonlinguistic sound throughout familiarization (i.e., sound condition), and 18 infants heard the same linguistic label throughout familiarization (i.e., label condition). An additional 14 infants were presented with the same testing stimuli, however, these infants were not familiarized to a category prior to testing (i.e., initial preference condition).

Apparatus Infants sat on parents' laps 100 cm away from a 152 cm x 127 cm projection screen, which was located approximately 5 cm above the infant's eye level. A Sony DCR-TRV40 camcorder was used to capture infants' fixations and was projected to a Dell flat panel monitor in the observation room. An NEC GT2150 LCD projector was mounted on the ceiling approximately 30 cm behind the infant. Two Boston Acoustics 380 speakers were 76 cm apart from each other and mounted in the wall. The speakers and camcorder were concealed by black felt and located directly below the projection screen. Two small lights were located behind the infant to ensure that the room was dimly lit throughout the entire procedure. In an adjacent room, a Dell Dimension 8200 computer with *Presentation* software was used to present stimuli to the infants, as well as to record the onset and offset of infants' visual fixations. Fixations to the visual stimuli were recorded online by pressing one of two buttons on a 10-button USB game pad when infants were looking at the stimuli and releasing the buttons when infants looked away from the stimuli.

Stimuli The auditory stimuli consisted of a non-speech sound (i.e., laser sound) and an infant-directed linguistic label (i.e., "a cat"), which was produced by a female experimenter. Both auditory stimuli were dynamic (i.e., changing in pitch and amplitude), and they were each presented at 65-68 dB for 1000 ms. The visual stimuli consisted of 10 familiarization stimuli (i.e., 10 different cats) and 4 test stimuli (i.e., 2 novel cats and 2 bears). All visual stimuli were realistic representations of cats and bears and were presented at approximately 36 cm x 36 cm in size (see Figure 1 for examples of visual stimuli).

Figure 1. Examples of Visual Stimuli

Familiarization Stimuli Testing Stimuli

Procedure The procedure consisted of two phases: a familiarization phase and a testing phase. During familiarization, infants were familiarized to 10 different stimuli within the cat category. Each familiarization trial consisted of a cat with a white background that appeared for 8000 ms. The auditory stimulus was presented at the onset of the visual stimulus and lasted for 1000 ms. Infants heard nothing in the silent condition. After each familiarization trial, the cat disappeared and the projection screen blackened for 1000 ms prior to the next familiarization stimulus. After 80 s of familiarization, infants were presented with two test trials. Each test trial consisted of a novel in-category exemplar (i.e., novel cat) and a novel out-of-category exemplar (i.e., novel bear), presented simultaneously. Test stimuli were presented to the left and right sides of the projection screen and approximately 64 cm apart from each other. As with familiarization trials, the duration of each test trial lasted 8000 ms. No auditory input was presented at test. The orders of familiarization and of test trials were randomized and the left-right location of the test stimuli were counterbalanced within subjects. Fixation durations were recorded online during familiarization and test trials.

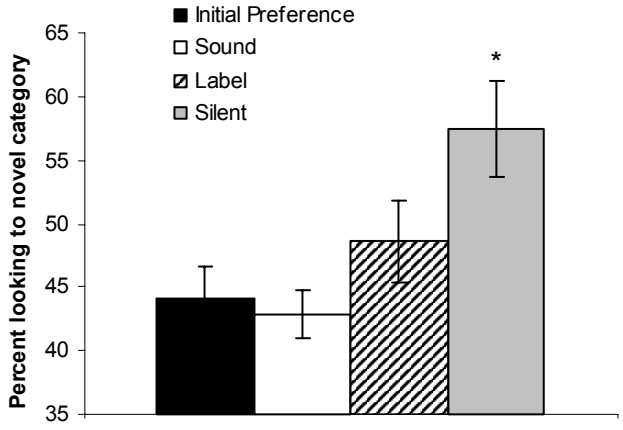
Results and Discussion

Analyses focused on accumulated looking during familiarization and the percent looking to the novel category at test. A one-way ANOVA with Stimulus Condition (i.e., Sound, Label, and Silent) as a between-subjects factor revealed a main effect of Stimulus Condition, F (2, 54) =

9.73, $p < .001$. Across the 80 s familiarization phase, infants accumulated more looking when the stimuli were accompanied by the nonlinguistic sound (M = 64.10 s, SE = 2.67 s) and the linguistic label (M = 61.91 s, SE = 2.27 s) than when the same visual stimuli were presented in the silent condition (M = 48.61 s, SE = 2.99 s), ts > 3.55, ps < .001. Furthermore, accumulated looking during familiarization did not differ between the label and sound conditions, t (37) = 0.61, p = .54, which suggests that the labels and sounds had comparable effects on increasing infants' attention.

Additional analyses focused on infants' relative looking to the novel category at test. As can be seen in Figure 2, 12-month-olds' looking to the novel category differed across the Stimulus Conditions (i.e., Initial Preference, Sound, Label, and Silent), one-way ANOVA with Stimulus Condition as a between-subjects factor, F (3, 67) = 6.05, $p < .001$. In particular, 12-month-olds in the silent condition (M = 57.89%, SE = 3.41%) increased looking to the novel category compared to infants' initial preference (M = 44.08%, SE = 2.51%), t (30) = 3.09, $p < .005$. In contrast, infants' looking to the novel category in the label condition (M = 48.62%, SE = 3.18%) and sound condition (M = 42.91%, SE = 1.88%) did not differ from infants' initial preference, ts < 1.07, ps > .29. In addition, planned comparisons revealed that 12-month-olds looked longer to the novel category in the silent condition compared to the sound condition, t (37) = 4.00, $p < .001$, and compared to the label condition, t (34) = 1.99, p = .055. Furthermore, similar to previous reports (e.g., Balaban & Waxman, 1997), these participants looked longer to the category in the label condition than in the sound condition, one-tailed p = .059.

Figure 2. Effect of Auditory Input on Categorization



Note: "*" different from initial preference, $p < .005$. Error bars reflect Standard Errors of the mean.

General Discussion

The results of the current study reveal several important findings. First, introducing sounds and labels during familiarization increased infants' overall attention, as indicated by more accumulated looking to the visual stimuli in the auditory conditions compared to the no auditory condition. However, this increase in looking did not correspond to an increase in categorization performance. Rather, infants in the current study were more likely to form a category when visual stimuli were presented in isolation than when the same stimuli were paired with sounds and labels.

Recall that we considered two views that may underlie effects of labels on categorization. The language-specific view suggests that linguistic input facilitates categorization because infants at the onset of word learning have broad assumptions that words and categories are linked (see Waxman, 2003 for a review). Thus, according to this view hearing the same label (but not the same sound) associated with different entities, directs infants' attention to the commonalties between these entities, which in turn, facilitates categorization (e.g., Balaban & Waxman, 1997; Waxman & Booth, 2003). Alternatively, it is possible that many of the effects of labels stem from the dynamics of attention in cross-modal processing (i.e., general auditory proposal). In particular, auditory input often overshadows visual input early in development (Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2004a; 2004b; Sloutsky & Napolitano, 2003). Furthermore, many of the effects of labels on cognitive tasks may stem from general auditory and familiarity effects, as opposed to assuming that infants at the onset of word learning understand the conceptual importance of labels (but see Xu, Cote, & Baker, 2005).

The results of the current study are consistent with the general auditory proposal: infants' increased attention in the sound and label conditions was primarily directed to the auditory input, and the auditory input hindered visual processing and category learning. However, since the current experiment did not assess encoding of auditory stimuli, it is also possible that the auditory stimulus simply interfered with forming categories (e.g., auditory stimulus added complexity to the task). While both of these explanations are consistent with the general auditory proposal, it is uncertain how the language-specific view can account for the current findings: infants in the label condition accumulated significantly more looking during familiarization compared to infants in the no auditory condition, however, these infants were less likely to form a category than infants in the no auditory condition.

If linguistic input overshadows visual input, how can linguistic input ever facilitate categorization? One possibility stems from attentional learning (Jones & Smith, 2002; Smith, Jones, & Landau, 1996; Smith, 1999; Yoshida & Smith, 2003). As infants start associating words with objects, they will eventually detect that each word is associated with a cluster of information (e.g., entities called *birds* often have

feathers and fly). Eventually children not only learn that specific words are linked to specific clusters, but they also learn a more abstract regularity: count nouns in general are linked to categories. Another possibility is that as children become progressively more familiar with words, or any auditory stimulus, overshadowing effects should decrease and possibly even facilitate visual processing.

Finally, it is also possible that effects of labels change in the course of processing. While the current study found no facilitative effects of labels on category learning after 80 s of familiarization, it is possible that labels may play a significant role later in the course of processing. For example, when auditory stimuli (e.g., familiar sounds and human speech) and visual stimuli are presented to children for short durations, the auditory stimulus often overshadows the visual stimulus (Napolitano & Sloutsky, 2004). In contrast, when familiar auditory stimuli are paired with visual stimuli for longer durations, overshadowing effects attenuate and under some conditions infants are actually more likely to encode the visual stimulus than when the same stimuli are presented in a unimodal baseline (Sloutsky & Robinson, 2005). Thus, understanding how infants allocate attention to cross-modal stimuli at various points in processing will not only shed light on attention and cross-modal processing, but will also highlight the component processes underlying word learning and the effects of labels on cognitive tasks such as categorization, individuation, and induction.

Acknowledgments

This research has been supported by grants from the National Science Foundation (REC # 0208103 and BCS # 0078945) to Vladimir M. Sloutsky.

References

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.

Jones, S. S., Smith, L. B., (2002). How children know the relevant properties for generalizing object names. *Developmental Science, 5,* 219-232.

Murphy, G. L., (2002). *The big book of concepts.* Cambridge, MA: MIT Press.

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.

Oakes, L. M., Madole, K. L., & Cohen, L. B. (1991). Object examining: Habituation and categorization. *Cognitive Development*, *6*, 377-392.

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.

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*, 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 XXVI Annual Conference of the Cognitive Science Society*, 1167-1172. Mahwah, NJ: Erlbaum.

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. (2005). *The role of words and sounds in visual processing: From overshadowing to attentional tuning.* Manuscript submitted for publication.

Smith, L. B., (1999). Children's noun learning: How general learning processes make specialized learning mechanisms. In B. MacWhinney (Ed.), *The emergence of language* (277-303). Mahwah, NJ: Erlbaum.

Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children: A dumb attentional mechanism. *Cognition, 60,* 143-171.

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, UK: Oxford University Press.

Waxman, S. R., & Booth, A. E. (2003). The origins and evolution of links between word learning and conceptual organization: New evidence from 11-month-olds. *Developmental Science*, *6*, 130-137.

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

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

Younger, B. A., & Furrer, S. D. (2003). A comparison of visual familiarization and object-examining measures of categorization in 9-month-old infants. *Infancy*, *4*, 327-348.

Yoshida, H., & Smith, L. B., (2003). Known and novel noun extensions: Attention at two levels of abstraction. *Child Development, 74,* 564-577.