Journal of Experimental Child Psychology 151 (2016) 5–17

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Labels constructively shape object categories in 10-month-old infants

Nadja Althaus [a,*], Gert Westermann [b]

[a] Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK
[b] Department of Psychology, Lancaster University, Lancaster LA1 4YF, UK

ARTICLE INFO

Article history:
Available online 6 January 2016

Keywords:
Categorization
Infancy
Eye-tracking
Cognitive development
Language acquisition
Language and cognition

ABSTRACT

How do infants' emerging language abilities affect their organization of objects into categories? The question of whether labels can shape the early perceptual categories formed by young infants has received considerable attention, but evidence has remained inconclusive. Here, 10-month-old infants ($N = 80$) were familiarized with a series of morphed stimuli along a continuum that can be seen as either one category or two categories. Infants formed one category when the stimuli were presented in silence or paired with the same label, but they divided the stimulus set into two categories when half of the stimuli were paired with one label and half with another label. Pairing the stimuli with two different nonlinguistic sounds did not lead to the same result. In this case, infants showed evidence for the formation of a single category, indicating that nonlinguistic sounds do not cause infants to divide a category. These results suggest that labels and visual perceptual information interact in category formation, with labels having the potential to constructively shape category structures already in preverbal infants, and that nonlinguistic sounds do not have the same effect.

* Corresponding author.
E-mail address: nadja.althaus@psy.ox.ac.uk (N. Althaus).

http://dx.doi.org/10.1016/j.jecp.2015.11.013
0022-0965/

Introduction

One of the central questions in early language and cognitive development is how the emergence of language affects infants' preverbal understanding of the world. During the past 20 years, many studies have shown that even very young infants can rapidly form categories on the basis of the static perceptual features of objects (Behl-Chadha, 1996; French, Mareschal, Mermillod, & Quinn, 2004; Mareschal & Quinn, 2001; Oakes, Coppage, & Dingel, 1997; Quinn, 2002, 2004; Quinn & Eimas, 1996) and that during the first year of life these early categories are gradually enhanced with more sophisticated knowledge such as feature correlations, sounds, motion, function, and animacy cues (Baumgartner & Oakes, 2011; Burnham, Vignes, & Ihsen, 1988; Pauen & Träuble, 2009; Perone & Oakes, 2006; Younger & Cohen, 1986). Other research has shown that in older children language can shape object categories by enabling the grouping together of perceptually dissimilar objects (e.g., dogs and whales as mammals) and the separation of similar objects (e.g., bats and birds) into different categories as well as forming the basis for inferences about hidden object properties (Graham, Kilbreath, & Welder, 2004; Welder & Graham, 2001). There is also evidence that labeling aids object individuation by 12 months of age (Xu, Cote, & Baker, 2005), and even younger infants at 9 or 10 months expect different labels to refer to different kinds of object (Dewar & Xu, 2007, 2009). However, although several studies have investigated the emerging effect of language on categorization during the first year of life, they have yielded little agreement and it is not yet clear how linguistic information interacts with object representations that have developed preverbally.

Two separate questions have been asked about the role of language in infants' object categorization during the first year of life. The first is whether labels can facilitate object categorization, that is, whether the labeling of objects enables infants to form categories that they would not form without labels. The second question is whether, like in older children, labels can override nonverbal perceptual information and change the structure of perceptual categories when visual similarity and category labels are in conflict with each other.

Much of the work exploring early categorization is based on the familiarization/novelty preference procedure (Fantz, 1964), which relies on the fact that infants tend to spend more time looking toward novel objects than toward familiar objects. In a typical study, infants are familiarized with a sequence of objects from one category and are then tested with two new objects—one a novel member of the familiarized category and one from a different category. If infants show a looking preference to the object from the new category, it can be concluded that they have formed a category that includes the novel within-category object but excludes the object from the different category. The emerging ability of language to shape categories has been studied in variants of this paradigm in which the objects presented during familiarization are accompanied by auditory labels or other sounds and the effect on infants' looking preferences during test are investigated.

The question of whether labels facilitate category formation has been addressed in several studies by Waxman and colleagues (Balaban & Waxman, 1997; Booth & Waxman, 2002; Ferry, Hespos, & Waxman, 2010; Fulkerson & Waxman, 2007; Waxman & Braun, 2005; Waxman & Markow, 1995). For example, in a seminal study, Balaban and Waxman (1997) showed that infants who were familiarized with a sequence of pig drawings exhibited at test a looking preference for a rabbit over a novel pig only when the familiarization items were accompanied by a labeling phrase ("A pig!") but not a tone sequence. Unsystematic labeling with different novel words did not lead to category formation in 12-month-olds (Waxman & Braun, 2005). Other studies have aimed to provide evidence that consistent novel labels facilitate categorization in infants at 6 months of age and even at 3 or 4 months (Ferry et al., 2010; Fulkerson & Waxman, 2007). This and other work has led to the claim that words serve as "invitations to form categories," enabling category formation by highlighting commonalities between objects with the same label (Waxman & Markow, 1995).

However, two considerations appear to weaken this claim. First, as described, many studies have shown that preverbal infants as young as 3 months can form perceptual categories in the absence of any auditory input, and these early categories can be at the basic or superordinate levels (Behl-Chadha, 1996; Quinn & Eimas, 1996). On this basis, it seems plausible that, in silence, 9-month-olds may be able to form a perceptual category of the stimuli used in the studies described above (e.g., a category of pigs that is different from rabbits). Because the described labeling studies did not involve a control condition in which the objects were presented in silence, it is difficult to assess whether labels had an effect above and beyond purely visual information (Plunkett, Hu, & Cohen, 2008; Robinson & Sloutsky, 2007a). Sloutsky and Napolitano (2003) and Sloutsky and Robinson (2008) have aimed to reconcile the results from categorization-in-silence studies and categorization-with-labels studies by arguing that, far from facilitating categorization, auditory information instead interferes with visual processing and, therefore, disrupts category formation when categories would have been formed in silence. According to this "auditory overshadowing hypothesis," labels disrupt category formation less than other auditory stimuli because they are highly familiar auditory signals. Therefore, visual categories that are learned in silence can still be formed in the presence of labels, but tone sequences usually disrupt visual category formation (Robinson & Sloutsky, 2007a, 2007b).

A second limitation of the described labeling studies is that they familiarized infants with only a single category. Even if we assume that labels have a facilitative effect above and beyond presentation in silence, such a paradigm cannot answer the question of whether labels constructively interact with preverbal representations to shape category structures. It is a well-known finding that infants' attention is maintained longer in the presence of labels compared with silence (Althaus & Mareschal, 2014; Baldwin & Markman, 1989; Robinson & Sloutsky, 2007a). The processes involved in this longer looking period may well have to do with more intensive visual processing, but not necessarily in the sense of using the information given by the label to adjust category boundaries. In familiarization with a single category, an increase in novelty preference might simply be the outcome of an effectively prolonged stimulus exposure.

The question of a constructive role of category labels can be addressed by familiarizing infants with two separate but visually similar categories with two distinct labels simultaneously. This is similar to word learning studies that test whether infants can associate words and (individual) objects (Mani & Plunkett, 2008; Schafer & Plunkett, 1998; Werker, Cohen, Lloyd, Casasola, & Stager, 1998). This paradigm can then be used to answer the two questions posed above. If infants do not form categories in silence but are successful under labeling, then labels facilitate category formation. If infants successfully categorize in silence but form *different* categories under labeling, then labels can serve to override visual information.

Plunkett and colleagues (2008) reported a two-category study, using a set of stimuli introduced by Younger (1985). These stimuli consisted of two sets of line drawings of schematic animals based on the same set of four distinctive features. In the "broad" set, features were combined such that there were no discernable clusters of stimuli; the eight stimuli formed a single large category. In the "narrow" set, the feature values were correlated (e.g., long-necked animals always had thick tails), and based on these correlations two subcategories of four stimuli could be formed. In Younger's study, 10-month-old infants were familiarized with the animal drawings in silence. Younger found that the infants were sensitive to the feature correlations and formed a single category for the broad set but formed two distinct categories for the narrow set. In a separate study, Younger and Cohen (1986) showed that 4-month-old infants were not able to make use of feature correlations, indicating that this ability develops between 4 and 10 months of age.

Plunkett and colleagues (2008) introduced labeling into Younger's (1985) paradigm. When 10-month-old infants were familiarized using the broad stimulus set and were provided with the same label ("Look! Dax!") for every stimulus, they formed one category, just like in silence. Likewise, when infants were familiarized with the narrow stimulus set and heard two different labels corresponding to the two subcategories, they formed two categories, again like in the silent case. However, importantly, when infants saw the narrow stimulus set but heard the *same* label for each stimulus, they formed a single category. Plunkett and colleagues argued that in this case the labels served to override visual categorization, leading infants to merge two perceptually distinct categories when they shared a common label.

However, these results leave open an alternative explanation, namely that forming two categories in the narrow condition depended on infants' ability to detect feature correlations, an ability that develops between 4 and 10 months of age. It is possible that the added complexity of the condition in which visual information (two categories) and labels (one label) were in conflict led to overloading the infants, resulting in a more shallow processing of the visual information. This explanation is compatible with results showing that when complex stimuli (e.g., conflicting visual and auditory information) overload the learning system, infants can regress to an earlier level of processing—here, that of 4-month-olds (Cohen, Chaput, & Cashon, 2002). In this case, instead of infants constructively using label information to merge two distinct categories, they could merely have failed to detect the feature correlations that define the two categories in the first place and, like 4-month-olds, formed a single category of the visual stimuli. One way in which to exclude the possibility that labels merely lead to shallower object processing is to conduct a study with a similarly ambiguous visual stimulus set that is in the absence of labels perceived as a single large category but that is parsed into two subsets when two distinct labels are presented. In other words, distinct labels cause infants to divide a large category rather than a single label causing them to merge two sets of objects. In this case, deeper processing is a prerequisite for the change in category boundary because a more detailed representation is necessary for subcategory formation. Such a result would constitute evidence for a constructive effect of labels. This is the approach taken in our study.

In sum, previous work with infants at the transition to language has not provided a coherent picture of how nonlinguistic and linguistic information interact in category formation. In particular, the central question of whether labels can guide the structure of object categories in preverbal infants remains open. Demonstrating such a constructive role requires the following methodological considerations. First, a labeling condition should be contrasted with a silent condition to ensure that labels have an effect beyond the visual information alone. Second, infants should be familiarized with a stimulus set that allows parsing into one or two categories to show that labels can shape the encoded category structure in a way that goes beyond merely enhancing attention. Third, the task should be designed so that category formation in the presence of labels cannot be explained by the processing of less visual detail than when familiarized in silence. A fourth point concerns the difficulty of ensuring that looking preferences represent novelty preference rather than familiarity preference. Previous work has demonstrated that infants' mode of preference can change as a function of task complexity or duration of exposure (Hunter, Ames, & Koopman, 1983; Roder et al., 2000). Because the addition of (one or more) auditory stimuli arguably changes task complexity, a test trial including a visually novel test item is necessary to establish whether infants display a novelty preference (longer looking at the novel test item) or a familiarity preference (shorter looking at the novel stimulus) (Cohen, 2004; Oakes, 2010). Note that Plunkett and colleagues (2008) did not include a novel test stimulus, which complicates interpretation of their results.

Here, we describe a study based on these considerations. We presented infants with a set of objects that we expected to be perceived as a single category when viewed in silence, and we asked whether pairing half of the objects with one label and half with another label would lead infants to split the objects into two categories according to their labels. This would require infants to use more detail for categorization with labeling than in the silent condition.

We constructed a set of stimuli by morphing a drawing of one novel animal into a drawing of a distinct second animal in a fixed number of steps, thereby obtaining a set of perceptually equidistant stimuli over which it was possible to define categories as well as subcategories (at either end of the morphing continuum) and their averages.

The overall logic of our study followed that of Younger (1985), but here differences between objects were based not on the variation in distinct features but rather on a holistic difference achieved by the morphing process. Infants were familiarized with eight objects either in silence, with one label, or with two labels. If labels serve to constructively shape category structure, infants should form a single category comprising all stimuli when familiarized in silence or when accompanied by a single label but should form two categories when two distinct labels are provided.

Experiment 1

Method

Participants

A total of 63 full-term 10-month-old infants participated in this study (mean age = 302 days; 28 girls). Infants were randomly assigned to either a silent condition ($n = 20$), a two-label condition

($n = 20$), or a one-label condition ($n = 23$). One additional infant was tested but excluded due to technical problems. Infants were recruited on a voluntary basis via local advertisements. Informed consent was received from the caregivers, and infants received a small gift for their participation.

Stimuli and design

Visual stimuli. The familiarization and test stimuli are depicted in Fig. 1. All stimuli were obtained by morphing Stimulus 1 into Stimulus 19 in 18 steps using the software MorphMagic. To ensure discriminability of the individual exemplars, only every second morphed image was used in the familiarization set. The gap between Stimuli 7 and 13 ensured that there was a visual grounding for a division into two separate subcategories (Stimuli 1, 3, 5, and 7 and Stimuli 13, 15, 17, and 19, respectively). Choice of test stimuli followed the design of Younger's (1985) study. It was assumed that infants forming one broad category would find the overall average stimulus (here, Stimulus 10) less novel than more peripheral stimuli (here, Stimuli 4 and 16). Conversely, infants who form two categories should find the overall average more novel because it falls between the categories but should not show preference for the peripheral stimuli because they now fall into the center of each of the two categories. A novel test animal was constructed to establish whether infants were showing novelty or familiarity preference at test.

The familiarization stimuli were depicted against a white background and presented on either the left or right half of the screen (counterbalanced such that no regular structure emerged). Stimuli subtended approximately 16 degrees visual angle on the screen. Roughly half of the infants saw the stimuli facing left, and the other half saw them facing right. The test stimuli were depicted pairwise facing in the same direction as the familiarization stimuli, either depicting the overall average (10) next to one of the subcategory averages (4 or 16) or depicting one of the average objects (10, 4, or 16) alongside the novel animal. This novel animal had been constructed to have the same configuration as the overall average (10) but different individual features. Features were chosen to correspond to those that were relevant to discriminate original animal drawings that served as the basis for morphing (e.g., tail, feet, ornament on tummy, wings, head shape) but looked different enough to be easily discriminable.

Auditory stimuli. The auditory stimuli were the phrases "Look, a geepee!" and "Look, a boota!" recorded digitally from a female native speaker of British English in enthusiastic, infant-directed speech. The label ("A geepee!"/"A boota!") was repeated twice, thereby occurring three times per stimulus. The onset of the first auditory phrase was at 2000 ms after the onset of the visual stimulus, with the labeling phrase ("A geepee!"/"A boota!") repeated at 5000 and 8000 ms after stimulus onset. In the two-label condition, half of the infants heard the label "geepee" with Stimuli 1, 3, 5, and 7 and heard the label "boota" with Stimuli 13, 15, 17, and 19—and vice versa for the other half of the infants. In the one-label condition, half ($n = 12$) of the infants heard "geepee" paired with all eight stimuli, and the other half heard "boota" paired with all eight stimuli.

Procedure

After a warm-up phase during which the experimenter explained the procedure and obtained written consent, infants were seated on the parent's lap at a distance of approximately 65 cm from a 22-inch screen. A 9-point calibration sequence was used to calibrate a Tobii X120 remote eye tracker (sampling frequency = 120 Hz, system accuracy = 0.5 degrees). During this procedure, a small animated object with accompanying sound was displayed at nine locations (left, center, and right for top, middle, and bottom row) on the screen. Calibration was repeated up to three times or until all nine points had been calibrated successfully.

After the calibration procedure, infants were presented with the eight familiarization stimuli, shown in a pseudorandomized order for on average 9750 ms ($SE = 22$) each. Individual sequences were obtained by using three Latin squares of three pseudorandom sequences, ensuring that no more than three consecutive stimuli were from the same subcategory. In the label conditions, infants heard the labels as described above. Prior to the first trial, and in between all subsequent trials, small animated video clips were shown at the center of the screen to direct infants' attention to this location. Their duration varied between 800 and 1800 ms.

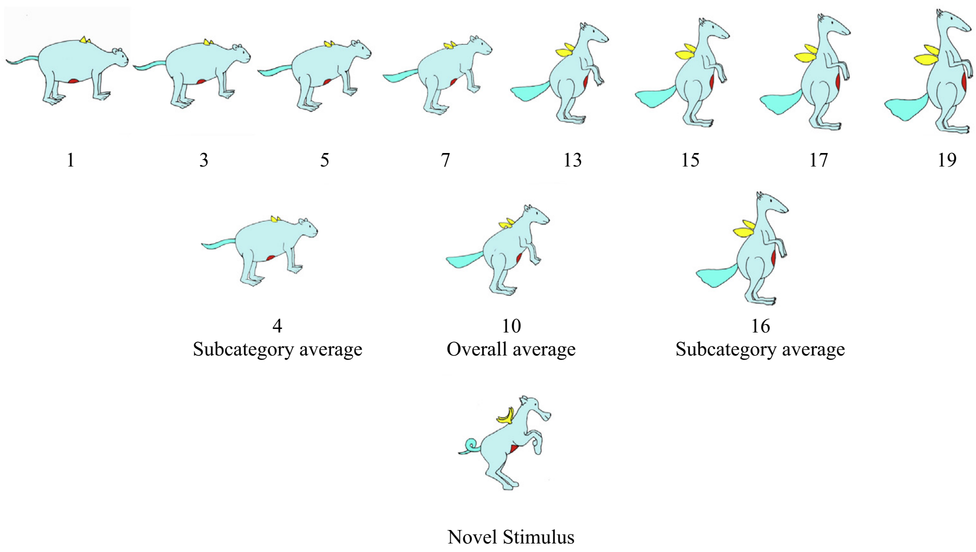

Fig. 1. Familiarization stimuli (top row) and test stimuli (bottom rows). Test stimuli fall into the center of the familiarization stimuli continuum (10) or into the center of each subcategory (4 and 16). There is also a novel stimulus that is visually distinct from the familiarization stimuli.

The familiarization phase was followed by a test phase consisting of six trials of 10 s each. The first and second, third and fourth, and fifth and sixth test trials were identical apart from left/right positioning of the test stimuli. The first two test trials always paired the overall average (Stimulus 10) with one of the subcategory averages (Stimulus 4 or 16, counterbalanced). The order of Test Trials 3/4 and 5/6 was counterbalanced; these presented infants with (a) the novel item paired with the overall average and (b) the novel item paired with the same subcategory average that was used on Test Trials 1 and 2 (i.e., each child either saw the subcategory average, 4, on Test Trials 1, 2, and either 3 and 4 or 5 and 6, or they saw the subcategory average, 16, on these four trials). The reason for always presenting the pair of overall average and subcategory average at the start of the test phase was that this is the most sensitive contrast and needs to be presented while infants are at the peak of familiarity with the target stimuli. In particular, there is the possibility that once the novel stimulus has been presented, the overall average and subcategory averages will appear equally uninteresting to participants regardless of the category structure established by the end of familiarization.

The location (left/right) of the items on Test Trials 1, 3, and 5 was counterbalanced across infants. The test trials were silent in all conditions.

Results

Looking time during familiarization

To assess whether familiarization had occurred, we compared average looking times during the first four and last four familiarization trials in all three conditions. The data were submitted to a mixed two-way analysis of variance (ANOVA) with the between-participants factor condition (silent, two-label, or one-label) and the within-participants factor block (Block 1 [Trials 1–4] or Block 2 [Trials 5–8]). There were significant main effects of condition, $F(2, 60) = 15.357$, $p < .001$, and block, $F(1, 60) = 14.563$, $p < .001$. The interaction between condition and block was not significant ($F = 0.087$, $p = .917$). Post hoc tests confirmed that infants in both label conditions had longer looking times than those in the silent condition (two-label: $M = 8275$ ms, $SE = 280$; one-label: $M = 8274$ ms, $SE = 271$; silent: $M = 6249$ ms, $SE = 271$; both $ps < .001$, Bonferroni-adjusted). Looking times between one- and two-label conditions did not differ significantly ($p = .59$). This result is consistent with previous studies showing longer looking under auditory input than in silence (Robinson & Sloutsky, 2007a). Furthermore, average looking times during Block 1 (M = 7763 ms, SE = 160) were higher than those during Block 2 (M = 7113 ms, SE = 189; p < .001, Bonferroni-adjusted), as demonstrated by a paired t-test on the collapsed data. This indicates that infants across all conditions had become familiarized by the end of the familiarization phase.

Preferential looking during test trials

Test trial results are given in Table 1 (proportions and test results) and Table 2 (absolute looking times). For the first two test trials, a preference score for the overall average object was obtained for each infant by dividing the total looking time at the overall average object (Stimulus 10) by the total looking time on that trial. For the remaining four test trials, a novelty preference score was obtained for each infant by dividing the looking time at the novel object by the total looking time on that trial. Preference scores for each pair of test trials (1 and 2, 3 and 4, and 5 and 6) were calculated as the average between the two.

A mixed effects ANOVA with the between-participants factor condition (silent, two-label, or one-label) and the within-participants factor test trial (novel vs. overall average or novel vs. subcategory average) revealed a significant interaction of condition and test trial, $F(2, 59)$ = 4.96, p = .010.[1] The main effects of condition, $F(2, 59)$ = 0.467, p = .629, and test trial, $F(1, 59)$ = 0.013, p = .91, were not significant. Planned comparisons against chance (.50) were conducted for each test trial in order to test whether infants preferred one stimulus over the other. In the silent condition, infants showed no preference for either stimulus when the subcategory averages and overall average stimuli were paired, $t(19)$ = 0.67, p = .512. Similarly, there was no looking preference in the pairing of the novel stimulus and the subcategory average, $t(18)$ = 1.10, p = .27. However, infants exhibited a preference for the novel stimulus when it was paired with the overall average stimulus, $t(18)$ = 4.36, p < .001, d = 1.00). This looking pattern indicates that the overall average, but not the subcategory averages, was more familiar than the novel exemplar. These results, therefore, indicate that in the silent condition infants formed a single category.

In the two-label condition, results for Test Trials 1 and 2 again showed no preference for either item, $t(19)$ = 0.279, p = .78. For the remaining test trials, the results from the silent and one-label conditions were reversed; infants showed no preference when the overall average and novel item were paired, $t(19)$ = 0.79, p = .43, but had a preference for the novel stimulus when paired with the subcategory averages, $t(19)$ = 4.17, p < .001, d = 0.94. These results indicate the formation of two categories in the two-label condition, resulting in the subcategory averages being perceived as familiar and the overall average stimulus being perceived as novel.

Results in the one-label condition resembled those in the silent condition; with no preference on the first two test trials, $t(22)$ = 0.97, p = .34, infants exhibited a significant novelty preference for the novel test item when it was paired with the overall average, $t(22)$ = 2.57, p = .017, d = 0.54, but not when it was paired with the subcategory averages, $t(22)$ = 0.438, p = .66. As in the silent condition, this pattern of results indicates that the infants formed one category when all objects were paired with the same label.

Taken together, these results indicate that infants in the silent condition and in the one-label condition formed a single category comprising all eight familiarization items, but they formed two categories when the objects were paired with two labels.

Discussion

Experiment 1 showed that infants formed a single category when objects were presented in silence and when they were paired with a single label. In contrast, infants split this category in two when the

[1] To ensure that results were not affected by either the order of test trials or the particular subcategory average used as the test item (4 or 16), we subjected the data from Experiments 1 and 2 to a mixed effects ANOVA with the repeated factor trial type (novel vs. overall average or novel vs. subcategory average) and the between-participants factors condition (silent, one-label, or two-label; two-tone, Experiment 2), trial order (overall average first or subcategory average first), and subcategory average number (4 or 16). This revealed a significant interaction between trial type and condition, $F(3, 63)$ = 4.765, p = .005, and a main effect of condition, $F(3, 61)$ = 3.441, p = .022. No other effects were significant (all Fs < 1.81, ps > .15). Therefore, we conclude that there were no effects of test order and subcategory average number.

Table 1
Average preference scores obtained in the test trials of Experiments 1 and 2.

Condition	Overall average vs. subcategory average (10 vs. 4/16)		Novel vs. overall average (novel vs. 10)		Novel vs. subcategory average (novel vs. 4/16)	
	M (*SE*)	*t*	*M* (*SE*)	*t*	*M* (*SE*)	*t*
Silent	.52 (.02)	0.67	.60 (.02)**	4.36	.54 (.03)	1.10
Two-label	.51 (.03)	0.26	.52 (.03)	0.78	.64 (.03)**	4.19
One-label	.52 (.02)	0.80	.55 (.03)*	2.07	.53 (.04)	0.64
Two-tone (Experiment 2)	.55 (.04)	1.48	.46 (.02) (*)	2.01	.48 (.02)	1.13

Note. Proportions given were directed at the underlined stimulus. Averages marked with two asterisks (**) were significant at the .001 level, and averages marked with one asterisk (*) were significant at the .05 level. Averages marked with an asterisk in brackets (*) were marginally significant.

Table 2
Average looking times obtained in the test trials of Experiments 1 and 2.

Condition	Overall average vs. subcategory average (10 vs. 4/16)	Novel vs. overall average (novel vs. 10)	Novel vs. subcategory average (novel vs. 4/16)
	M (*SE*)	*M* (*SE*)	*M* (*SE*)
Silent	2691 (276)	3530 (347)	3203 (395)
Two-label	3326 (341)	3601 (268)	3018 (281)
One-label	3471 (277)	3265 (260)	2824 (257)
Two-tone (Experiment 2)	3004 (333)	2991 (226)	2748 (336)

Note. Average looking times are given in milliseconds and were directed at the underlined stimulus.

objects were paired with two distinct labels. Infants who were familiarized with the visual stimuli in silence appeared to accept a large variety of head shapes, leg types, and body posture as possible realizations of the target category. Infants' preferences during the test phase indicated that in the silent condition they integrated all of the different familiarization exemplars into a single category representation whose prototype was the overall average (see also Younger & Gotlieb, 1988, for evidence that infants form prototype representation for perceptual categories). In contrast, infants distinguished between the two subsets at either end of the morphing continuum when the same objects were accompanied by two different labels, forming two more restrictive exclusive categories. If this result were due to labels merely enhancing visual processing overall (i.e., the trials simply being more interesting in the presence of speech), we would expect two categories to be formed regardless of how many different labels accompanied the visual stimuli. This is not the case; when a single label was presented with all stimuli, infants formed just one large category. Our results, therefore, suggest that infants tracked the correspondences between labels and object features. Furthermore, the current pattern of results is incompatible with the auditory overshadowing hypothesis, which predicts that visual processing in the presence of labels is less detailed than in silence; to form two categories, infants needed to process more detailed visual information to enhance the perceptual difference between the objects in the two subcategories. In particular, our findings indicate that the condition with the largest amount of novel auditory stimulation (two novel labels) leads to the most detailed representation (two separate subcategories). Together, therefore, these results provide unequivocal evidence that labels can have a constructive role in category formation in infants prior to their first productive use of language.

One open question with regard to infants' performance on the test trials is why infants in all three conditions did not exhibit any preference on the first two test trials contrasting the overall and subcategory averages. A possible explanation is that novelty preference depends on a sort of "feature pop-out." Because the stimuli were obtained using morphing, there were no clearly identifiable feature differences between stimuli presented side by side, that is, between the overall average stimulus (10) and the two subcategory averages (4 and 16). It seems plausible that the lack of clear novelty could have prompted infants to compare stimuli, resulting in overall similar looking times to both test items. Specifically, none of the stimuli used here (the overall average, 10, and the subcategory averages, 4 and 16) *violates* the familiarized category structure. Even if an infant has formed two categories during familiarization and, therefore, the overall average (10) is perceived as relatively novel, the infant could attempt to integrate this stimulus into the previously constructed category structure by merely extending the category boundaries without needing to completely re-form the categories. Similarly, if the infant has formed one category, the subcategory average is not implausible as a category member but merely further away from the overall category average.

Infants' looking patterns on test trials containing a novel stimulus suggested competition effects between the two test stimuli presented side by side; pop-out of the features of the novel stimulus creates a novelty effect, and if the other stimulus is perceived as highly familiar, it presents no competition to the novel item. In contrast, if the other item is perceived as less familiar, infants will spend time exploring the novel stimulus but also scan the other item because more processing is necessary to incorporate it into the established category. In other words, competition between the test items here, or lack thereof, reveals how different category boundaries have been formed.

One remaining open question is whether the observed constructive effect for labels in category formation is specific to language or whether it extends to other forms of auditory input. In the work by Waxman and colleagues (Balaban & Waxman, 1997; Ferry et al., 2010; Fulkerson & Waxman, 2007), labels were contrasted with tone sequences and differential effects were found. However, as discussed above, these differences were interpreted by Sloutsky and colleagues as being due to the different familiarity of speech sounds and nonspeech sounds. Even in studies that did not find positive effects for labels, nonlinguistic sounds had a more detrimental impact than labels (Robinson & Sloutsky, 2007a). Finally, the study by Plunkett and colleagues (2008), which is so far the best controlled study, did not involve a nonspeech sound condition. Therefore, to establish whether labels take a special role in shaping visual categories, we tested an additional group of infants on a nonlinguistic sound condition.

Experiment 2

The design and procedure for this experiment were identical to the two-label condition in Experiment 1 except that instead of auditory labels, nonlanguage sounds were played.

Method

Participants

A total of 17 10-month-old infants participated in this experiment (mean age = 300 days; 8 girls). A further 6 infants were excluded due to fussiness ($n = 1$) or technical problems ($n = 5$).

Materials

The visual stimuli were the same as in Experiment 1. Two sounds were used: a tingling bell sound and a wooden xylophone tone sequence. Each sound lasted 1 s. Each sound was played three times during the presentation of a picture, starting at 2000, 5000, and 8000 ms after stimulus onset. Half of the infants heard the bell sound with Stimuli 1, 3, 5, and 7 and the xylophone sound with Stimuli 13, 15, 17, and 19—and vice versa for the other half of the infants.

Procedure

The procedure was identical to that in Experiment 1.

Results

Familiarization

A paired *t*-test on average looking times for Blocks 1 and 2 of familiarization (Block 1 [Trials 1–4] or Block 2 [Trials 5–8]) revealed that there was a trend for infants to begin to habituate by the second block of familiarization (Block 1: $M = 6937$ ms, $SE = 446$; Block 2: $M = 5999$ ms, $SE = 573$), $t(16) = 2.033$, $p = .059$. An ANOVA on the average looking time per trial was conducted to compare looking time during familiarization in Experiment 2 with the conditions of Experiment 1. This revealed a significant effect of condition, $F(3, 76) = 10.13$, $p < .0005$. Post hoc comparisons showed that infants in Experiment 2 spent approximately equal amounts of time gazing at the stimuli ($M = 6468$ ms, $SE = 336$) as those in the silent condition of Experiment 1 ($p > .99$); that is, they exhibited shorter familiarization looking times than infants in the labeling conditions (one-label: $p = .023$; two-label: $p = .001$).

Test trials

Looking preference scores for all test trials are shown in Table 1. Test trials and calculation of looking preferences were the same as in Experiment 1. For Test Trials 1 and 2 (overall average vs. subcategory average), the preference score did not differ from chance, $t(16) = 1.48$, $p = .16$, two-tailed. In test trials pairing the novel stimulus with the overall average stimulus, there was a marginally significant preference for the familiar item (i.e., the overall average) over the novel item, $t(16) = 2.01$, $p = .061$, $d = -0.49$. On the test trials pairing the novel stimulus with a subcategory average, infants did not reliably prefer either stimulus on these two test trials (t-test on average preference scores: $t(16) = 1.13$, $p = .28$). In summary, infants' results differed from those found in Experiment 1. The marginal familiarity preference exhibited when the novel stimulus was paired with the overall overage stimulus indicates that infants formed a single category in the two-tone condition of Experiment 2.

Discussion

The two-label condition in Experiment 1 revealed that pairing objects from a single category with two labels led infants to split the category into two separate categories. Pairing the visual stimuli with two nonlinguistic tones did not have the same effect. The only (marginal) preference shown by the infants was one for the overall average when it was paired with the novel item. This looking pattern can only be interpreted as a familiarity preference. As described above, familiarity preferences can occur when the familiarization stimuli are complex and/or infants have not become sufficiently familiarized during training.

In line with our interpretation of the results in Experiment 1, we need to interpret the familiarity preference here as indicative of infants forming a single category over the familiarization stimuli. This means that despite the presence of two distinct auditory stimuli, infants are not persuaded that there are two visual categories—in contrast to the label condition, where their category representation was guided by the structure provided by the labels.

This result is compatible with the auditory overshadowing hypothesis in the sense that infants (a) seem to struggle with category formation, as indicated by the marginal test preference, and (b) do not use the auditory stimuli to group the visual objects. Therefore, nonlinguistic tone stimuli, unlike labels, appear to interfere with infants' processing of visual material. Although category formation per se does not appear to be disrupted in Experiment 2, the finding is consistent with results from previous studies showing that nonlinguistic sounds do not facilitate category formation (Balaban & Waxman, 1997; Ferry et al., 2010; Fulkerson & Waxman, 2007; Robinson & Sloutsky, 2007a; Sloutsky & Robinson, 2008).

General discussion

Our aim in this work was to investigate how object labels interact with visual information during category formation in preverbal 10-month-old infants. Although a body of literature that has addressed this question exists, we argued that methodological considerations left the results of these experiments inconclusive. We argued that to investigate whether labels can constructively guide category formation and override visual similarities in young infants it is necessary to (a) familiarize infants on a set of stimuli that could be parsed into one or two categories, (b) test infants under varying labeling conditions (silence, one label, and two labels), (c) include a novel test stimulus to ascertain whether at test infants show a novelty or familiarity preference, and (d) construct the stimulus set so that the effects of labels cannot be explained as processing less visual detail than without labels. We presented a paradigm that incorporates these considerations, familiarizing infants on a stimulus set that could be grouped into one large category or two smaller categories. We found that on presentation in silence or when all objects were paired with the same label, infants formed a single category. In contrast, when half of the objects were paired with one label and the other half were paired with another label, infants formed two separate categories in line with how the objects had been labeled. Conversely, when the objects were paired with nonlinguistic sounds, infants formed a single category but exhibited familiarity preference. In infants at the transition to language, these results present clear evidence that labels interact with nonlinguistic information to constructively shape categories and that these effects are specific to linguistic labels and do not extend to nonspeech sounds.

The results of our experiments reconcile arguments that sounds act to overshadow visual processing and that labels have a constructive effect on category formation. The results of Experiment 2 suggest that nonlinguistic tones indeed interfere with visual processing so that infants did not form a category in the tone condition, whereas in silence they did. However, whereas Sloutsky and Napolitano (2003) and Sloutsky and Robinson (2008) argued that labels merely interfere less with visual processing than nonlinguistic sounds but contribute nothing to category formation, our Experiment 1 clearly shows that labels do in fact constructively affect categorization. Therefore, depending on their nature, auditory signals can both interfere with and shape visual categories. Furthermore, recent results also indicate that the temporal relationship between presentation of visual and auditory stimuli matters so that interference is especially strong when visual and auditory stimuli have a common onset (Althaus & Plunkett, 2015).

Some authors have discussed the role of labels in terms of whether they function as features that increase similarity between objects in a bottom-up manner (Gliozzi et al., 2009; Sloutsky et al., 2001) or whether they act in a referential manner, serving as "names" (Waxman & Gelman, 2009). Waxman and Markow (1995) described the role of labels as "invitations to form categories" by highlighting common visual features, suggesting a top-down role for labels. Likewise, Waxman and Gelman (2009) argued that words are "referential" in nature and, as such, are more than associates or perceptual features. Gliozzi and colleagues (2009), in contrast, demonstrated that even results such as those presented by Plunkett and colleagues (2008), which show a change in category formation based on label–object correlations, can be simulated with a connectionist model that treats labels as features.

In our study, labeling caused infants to use more restrictive criteria for a classification of two items as "similar," effectively producing two smaller categories. The label, therefore, serves not so much as a mere additional feature but instead as a feature that modulates the way in which visual features are used. As such, our interpretation of the labels' role is somewhere between the extremes of "labels as features" and "labels as names"; labeling has a supervisory function in the sense that it modulates the processing of similarity, but whether this is enough to establish a "referential" relationship between the labels and objects in these infants at just 10 months of age remains open. This interpretation is in line with recent computational models (Althaus & Mareschal, 2013; Westermann & Mareschal, 2014) that simulate the role of words in the context of categorization. Althaus and Mareschal (2013) demonstrated how early interactions between word learning and learning about objects led to improved category representations compared with isolated learning, without labels explicitly acting as supervisory signals. The model by Westermann and Mareschal (2014) shows how labels can affect the similarity relations between objects in a top-down manner by warping the perceptual similarity space to increase perceptual distance between objects that have different labels. Such a mechanism could also account for the results presented here.

Our findings fit well with recent theories about interactions of language with other domains. Although the strong Whorfian hypothesis of linguistic determinism (Whorf, 1956) has been largely discounted over the last few decades, newer research is gradually establishing links between language processing and perception or cognition in adult processing (Boroditsky, 2001; Lupyan, 2008a, 2008b; Lupyan, Rakison, & McClelland, 2007). In particular, the "label feedback hypothesis" (Lupyan, 2012) suggests that language modulates cognitive processes, for example, benefitting visual search or recognition memory (Lupyan, 2008a, 2008b). It seems likely that the effects we observe in the experiments presented here reflect early stages of these interactions between language and cognition.

To summarize, the experiments presented in this article show that labels have a constructive impact on category formation in preverbal infants; when a set of objects is accompanied by labels, 10-month-old infants adapt their category structure in correspondence to how the objects are labeled. This ability necessitates a deeper processing of objects when they are paired with labels than when they are presented in silence, with different labels highlighting visual differences between objects. Our results further indicate that the effect of labels is not merely to enhance attention but that it is also a constructive process in the sense that infants' categorization behavior was contingent on the number of labels they heard and that infants aligned their categories with the labels. Object labels, thus, are powerful stimuli that are able to direct infants' attention to different visual dimensions in the construction of visual categories.

Acknowledgments

This work was supported by Economic and Social Research Council UK Grant RES-000-22-3394 awarded to Gert Westermann. We are grateful to the parents and infants who took part in this study.

References

Althaus, N., & Mareschal, D. (2013). Modeling cross-modal interactions in early word learning. *IEEE Transactions on Autonomous Mental Development, 5*, 288–297.

Althaus, N., & Mareschal, D. (2014). Labels direct infants' attention to commonalities during novel category learning. *PLoS One, 9*(7), e99670.

Althaus, N., & Plunkett, K. (2015). Timing matters: The impact of label synchrony on infant categorisation. *Cognition, 139*, 1–9.

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.

Baldwin, D., & Markman, E. (1989). Establishing word–object relations: A first step. *Child Development, 60*, 381–398.

Baumgartner, H. A., & Oakes, L. M. (2011). Infants' developing sensitivity to object function: Attention to features and feature correlations. *Journal of Cognition and Development, 12*, 275–298.

Behl-Chadha, G. (1996). Basic-level and superordinate-like categorical representations in early infancy. *Cognition, 60*, 105–141.

Booth, A. E., & Waxman, S. (2002). Object names and object functions serve as cues to categories for infants. *Developmental Psychology, 38*, 948–957.

Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology, 43*, 1–22.

Burnham, D. K., Vignes, G., & Ihsen, E. (1988). The effect of movement on infants' memory for visual compounds. *British Journal of Developmental Psychology, 6*, 351–360.

Cohen, L. B. (2004). Uses and misuses of habituation and related preference paradigms. *Infant and Child Development, 13*, 349–352.

Cohen, L. B., Chaput, H. H., & Cashon, C. H. (2002). A constructivist model of infant cognition. *Cognitive Development, 17*, 1323–1343.

Dewar, K., & Xu, F. (2007). Do 9-month-old infants expect distinct words to refer to kinds? *Developmental Psychology, 43*, 1227–1238.

Dewar, K., & Xu, F. (2009). Do early nouns refer to kinds or distinct shapes? Evidence from 10-month-old infants. *Psychological Science, 20*, 252–257.

Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science, 146*, 668–670.

Ferry, A. L., Hespos, S. J., & Waxman, S. R. (2010). Categorization in 3- and 4-month-old infants: An advantage of words over tones. *Child Development, 81*, 472–479.

French, R. M., Mareschal, D., Mermillod, M., & Quinn, P. C. (2004). The role of bottom-up processing in perceptual categorization by 3- to 4-month-old infants: Simulations and data. *Journal of Experimental Psychology: General, 133*, 382–397.

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.

Gliozzi, V., Mayor, J., Hu, J. F., & Plunkett, K. (2009). Labels as features (not names) for infant categorization: A neurocomputational approach. *Cognitive Science, 33*(4), 709–738.

Graham, S. A., Kilbreath, C. S., & Welder, A. N. (2004). Thirteen-month-olds rely on shared labels and shape similarity for inductive inferences. *Child Development, 75*, 409–427.

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.

Lupyan, G. (2008a). From chair to "chair": A representational shift account of object labeling effects on memory. *Journal of Experimental Psychology: General, 137*, 348–369.

Lupyan, G. (2008b). The conceptual grouping effect: Categories matter (and named categories matter more). *Cognition, 108*, 566–577.

Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology, 3*. http://dx.doi.org/10.3389/fpsyg.2012.00054.

Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking: Redundant labels facilitate learning of novel categories. *Psychological Science, 18*, 1077–1083.

Mani, N., & Plunkett, K. (2008). Fourteen-month-olds pay attention to vowels in novel words. *Developmental Science, 11*, 53–59.
Mareschal, D., & Quinn, P. C. (2001). Categorization in infancy. *Trends in Cognitive Sciences, 5*, 443–450.
Oakes, L. M. (2010). Using habituation of looking time to assess mental processes in infancy. *Journal of Cognition and Development, 11*, 255–268.
Oakes, L. M., Coppage, D. J., & Dingel, A. (1997). By land or by sea: The role of perceptual similarity in infants' categorization of animals. *Developmental Psychology, 33*, 396–407.
Pauen, S., & Träuble, B. (2009). How 7-month-olds interpret ambiguous motion events: Category-based reasoning in infancy. *Cognitive Psychology, 59*, 275–295.
Perone, S., & Oakes, L. M. (2006). It clicks when it is rolled and it squeaks when it is squeezed: What 10-month-old infants learn about object function. *Child Development, 77*, 1608–1622.
Plunkett, K., Hu, J. F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition, 106*, 665–681.
Quinn, P. C. (2002). Category representation in young infants. *Current Directions in Psychological Science, 11*(2), 66–70.
Quinn, P. C. (2004). Multiple sources of information and their integration, not dissociation, as an organizing framework for understanding infant concept formation. *Developmental Science, 7*, 511–513.
Quinn, P. C., & Eimas, P. D. (1996). Perceptual organization and categorization in young infants. In C. Rovee-Collier & L. P. Lipsitt (Eds.). *Advances in infancy research* (Vol. 10, pp. 1–36). Norwood, NJ: Ablex.
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., Roder, B. J., Bushneil, E. W., Bushneil, E. W., Sasseville, A. M., & Sasseville, A. M. (2000). Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy, 1*, 491–507.
Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development, 69*, 309–320.
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., & Robinson, C. (2008). The role of words and sounds in infants' visual processing: From overshadowing to attentional tuning. *Cognitive Science, 32*, 342–365.
Sloutsky, V. M., Lo, Y. F., & Fisher, A. V. (2001). How much does a shared name make things similar? Linguistic labels, similarity, and the development of inductive inference. *Child Development*, 1695–1709.
Waxman, S. R., & Braun, I. (2005). Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition, 95*, B59–B68.
Waxman, S., & Gelman, S. (2009). Early word-learning entails reference, not merely associations. *Trends in Cognitive Sciences, 13*, 258–263.
Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29*, 257–302.
Welder, A. N., & Graham, S. A. (2001). The influence of shape similarity and shared labels on infants' inductive inferences about nonobvious object properties. *Child Development, 72*, 1653–1673.
Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology, 34*, 1289–1309.
Westermann, G., & Mareschal, D. (2014). From perceptual to language-mediated categorization. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*. http://dx.doi.org/10.1098/rstb.2012.0391.
Whorf, B. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press.
Xu, F., Cote, M., & Baker, A. (2005). Labeling guides object individuation in 12-month-old infants. *Psychological Science, 16*, 372–377.
Younger, B. (1985). The segregation of items into categories by ten-month-old infants. *Child Development, 56*, 1574–1583.
Younger, B., & Cohen, L. (1986). Developmental change in infants' perception of correlations among attributes. *Child Development, 57*, 803–815.
Younger, B., & Gotlieb, S. (1988). Development of categorization skills: Changes in the nature or structure of infant form categories? *Developmental Psychology, 24*, 611–619.