Developmental Science

Developmental Science 18:1 (2015), pp 65–79 DOI: 10.1111/desc.12179

PAPER

Words, shape, visual search and visual working memory in 3-year-old children

Catarina Vales and Linda B. Smith

Department of Psychological and Brain Sciences, Indiana University, USA

Abstract

*Do words cue children's visual attention, and if so, what are the relevant mechanisms? Across four experiments, 3-year-old children (*N = *163) were tested in visual search tasks in which targets were cued with only a visual preview versus a visual preview and a spoken name. The experiments were designed to determine whether labels facilitated search times and to examine one route through which labels could have their effect: By influencing the visual working memory representation of the target. The targets and distractors were pictures of instances of basic-level known categories and the labels were the common name for the target category. We predicted that the label would enhance the visual working memory representation of the target object, guiding attention to objects that better matched the target representation. Experiments 1 and 2 used conjunctive search tasks, and Experiment 3 varied shape discriminability between targets and distractors. Experiment 4 compared the effects of labels to repeated presentations of the visual target, which should also influence the working memory representation of the target. The overall pattern fits contemporary theories of how the contents of visual working memory interact with visual search and attention, and shows that even in very young children heard words affect the processing of visual information.*

Introduction

A large literature suggests that language – and particularly labeling – has on-line effects on visual processes of attention (Huettig & Altmann, 2011; Huettig & Hartsuiker, 2008), categorization (Lupyan, Rakinson & McClelland, 2007), and stimulus detection (Lupyan & Spivey, 2010a) in adults, and perhaps also in infants and children (Fernald, Thorpe & Marchman, 2010; Ferry, Hespos & Waxman, 2010; Johnson, McQueen & Huettig, 2011; Mani, Johnson, McQueen & Huettig, 2013). However, the on-line mechanisms through which heard words influence visual attention and visual processing are not well understood (Huettig, Olivers & Hartsuiker, 2011a). In this paper, we will provide evidence regarding one possible mechanistic route by bringing together two distinct literatures: How basic-level category names influence young children's categorization by object shape and how visual working memory representations affect adult visual search.

Explicitly naming objects has been shown to increase children's attention to the shapes of the named thing over other properties, such that children are more likely to group objects by shape in labeling than non-labeling conditions (Landau, Smith & Jones, 1992). Several studies further suggest that basic-level names may alter how children represent the shapes of both novel and known things, biasing them to pay more attention to the aspects of shape relevant to determine category membership (Gershkoff-Stowe, Connell & Smith, 2006; Yoshida & Smith, 2003a). According to one account of these phenomena, the effects arise because category names are associated with, and predict, specific shapes. As a consequence, heard names cue attention to category shape and bias how those visual shapes are encoded and represented (Jones & Smith, 1993; Gershkoff-Stowe *et al.*, 2006; Yoshida & Smith, 2003b; Smith, Jones, Yoshida & Colunga, 2003).

Labeling has also been shown to influence adult performance in visual search tasks, in which participants are asked to find a target object in an array of

Address for correspondence: Catarina Vales, Department of Psychological and Brain Sciences, Indiana University, 1101 E. 10th St., Bloomington, IN 47405, USA; e-mail: cvales@indiana.edu

distractors. Adults are faster when the target is labeled prior to search (Lupyan & Spivey, 2010b). Adults are also faster at finding the target when they are holding information in memory that matches the target (Soto, Heinke, Humphreys & Blanco, 2005; Soto, Humphreys & Heinke, 2006) or when they have been presented with a visual preview of the specific target (Schmidt & Zelinsky, 2009; Vickery, King & Jiang, 2005; Yang & Zelinsky, 2009). Likewise, adult search is slowed if the information held in memory matches the distractors (Soto & Humphreys, 2007). Working memory representations are believed to guide visual search by automatically biasing visual attention to items in the array that match the contents of visual working memory (Kristjánsson, Wang & Nakayama, 2002; Soto, Hodsoll, Rotshtein & Humphreys, 2008; Soto & Humphreys, 2007), with more robust or more accurate representations of the target leading to faster search.

If we put these two ideas together – that basic-level category names bias children's encoding and representation of object shape, and that the contents of visual working memory bias where one looks – then we arrive at the hypothesis tested in the following experiments: Naming objects in a visual search task should bias children to attend to items in an array that match the named entity. The participants in the experiments were 3-year-old children who typically show a shape bias in novel noun learning tasks (Landau, Smith & Jones, 1988) and the visual search arrays were composed of pictures of instances of basic-level categories. The linguistic cues were the basic-level category name for the pictured item. Consistent with traditional measures of visual search (Wolfe, 1998; see also Gerhardstein & Rovee-Collier, 2002), we asked children to find a target in an array of distractors that varied in number, and search time was measured as a function of the number of distractors. Search time on any trial is conceptualized as being the product of several processing steps: Encoding and representing the target in visual working memory, searching the array to find the matching target, and selecting a response (see Solman, Cheyne & Smilek, 2011). The intercept of the search function relating search time to number of distractors is conceptualized as reflecting processes that do not depend on the number of elements in the array, whereas the slope of the search function measures the cost of each added distractor to the time to decide if a member of the array matches the target (Solman *et al*., 2011; Vickery *et al*., 2005; Woodman, Vogel & Luck, 2001). Past research with adults indicates that labeling affects overall search time (i.e. the intercept; Lupyan & Swingley, 2011; Soto *et al*., 2006; see also Lupyan & Spivey, 2010b; Soto & Humphreys, 2007), a result consistent with an effect on target representation in working memory.

In order to fit the cognitive skills of 3-year-old children, our visual search procedure differed in several ways from the usual approaches in adult studies. First, children searched for the very same target within a block of trials. We took this approach because past research indicates that young children show strong trial-to-trial carry-over effects, cannot readily switch rule assignments and also need continual reminding of the response rule (Chevalier & Blaye, 2009; Garon, Bryson & Smith, 2008). Second, in all conditions – Label and Silent – children were visually shown the search target on every trial, a procedure that helps these young children stay on task and is also similar to the visual preview of the target used in adult studies. Thus, the experiments compare performance in a Silent condition in which children are shown the target on every trial with performance in a Labeling condition that *adds* the spoken basic-level name of the target to the visual information. By hypothesis, the label should bias encoding of the shape of the previewed target over other properties such as color (Experiments 1 and 2) and enhance encoding of category-relevant aspects of shape (Experiments 3 and 4). If names for basic-level categories increase children's attention to shape in the sense of leading to more robust representations of target shape in visual working memory, then providing the basic-level name for a shown target should lead to better representation of category-relevant shape and thus more rapid detection of the target in an array of distractors.

Experiment 1

In Experiment 1, the target was defined by both its basic-level category shape and by its color. On every trial, half the distractors matched the target in shape and half matched in color. For example, if the target were a red bed, half the distractors were green beds and half were red couches. In both the Label and Silent conditions, the target was visually displayed at the start of each trial. In the Label condition, children heard the displayed target named with a noun (e.g. 'bed') prior to each search trial; in the Silent condition, they just saw the displayed target (see Figure 1). If hearing the name biases working memory representations of the target shape and if these stronger representations preferentially guide attention to the shape-matching objects (the beds) over the non-shape-matching objects (couches), then children should be able to find the specific target (the red bed) more rapidly in the Label condition.

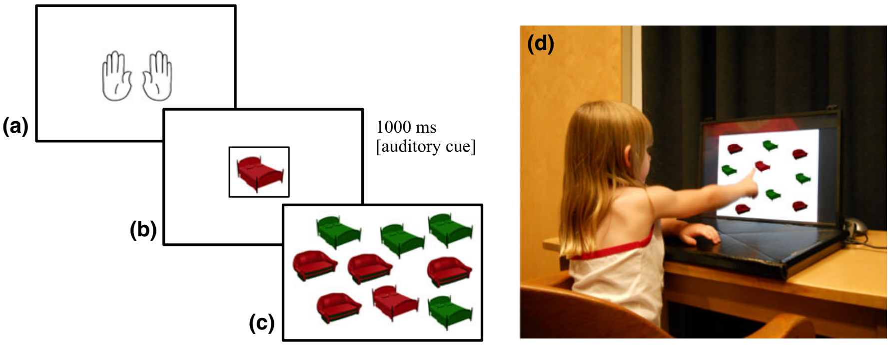

Figure 1 *Left: Main structure of a trial (the stimuli depicted were used in Experiment 1 and 2). Right: Experimental set up for all experiments.*

Method

Participants

Thirty-two children between 31 and 43 months of age (18 males; mean age: 37 months, *SD*: 2.9) were randomly assigned to either the Silent or the Label condition. Ten additional children were recruited but not included in the final sample due to refusal to participate in the study ($N = 3$), not finishing the familiarization phase ($N = 1$), or selecting a non-target object on most test trials and thereby not meeting the criterion of at least two correct responses per distractor set size (N = 6). Children had no known developmental disorders, and were reported to have normal (or corrected to normal) visual acuity and color vision. English was the main language spoken by all families. Parental consent was obtained for all participants in compliance with the IRB of Indiana University, and all children received a toy for participating.

Apparatus and stimuli

Stimuli were presented on a 17″ monitor equipped with a touchscreen (MagicTouch, Keytec, Garland, TX). Stimuli were presented and responses (location and latency) were recorded using E-Prime (PST, Pittsburg, PA). Each test stimulus was rendered in a 180 × 140 pixel area on a white background and could be placed in 16 different locations. Across test trials, the target appeared equally often on the left and right side of the screen. The audio files used in the Label condition were recorded using an artificial speech creator at a sample rate of 16KHz.

Procedure

Figure 1 shows the experimental set up and the temporal order of events on each trial. The child was seated at approximately 35 cm from the screen. On each test trial, a 'fixation' slide encouraged the child to rest their hands on the table (Figure 1a) before the target object was displayed on the center screen for 1 sec (Figure 1b). The search array was then displayed and the child asked to find the target picture as fast as possible (Figure 1c). Each child was assigned one search target and searched for the same target throughout 32 test trials. Four different objects served between-subjects as targets: a red bed, a red couch, a green bed, and a green couch. For each target, the distractors were selected so that half had the same shape and half had the same color as the target – so if the target object was a red bed, on each trial half the distractors would be red couches and the other half would be green beds (see Figure 1c). The number of distractor objects was 2, 4, 8 or 12 distractors. The order of the 32 trials, with eight occurrences of each set size, was randomly determined for each subject. The experimenter started each trial ensuring that the child was looking at the screen; no time limit was set for finding the target and no feedback was given. In the Label condition, a sound file containing the name of the target object (e.g. 'bed') played at the onset of the target cuing display (Figure 1b). No sound file was played in the Silent condition. None of the objects were labeled by the experimenter or the caregiver at any point during the session; in giving task instructions, the experimenter would say: 'Find this one' or 'Which one did you see?'

Prior to the test phase just described above, children were familiarized with using the touchscreen and with the idea of search. They were shown how to hold their hands on the table during the fixation slide and taught to watch the target preview and then, when the search array appeared, to touch the object that looked like the one they had just seen as soon as they saw it. The objects used during familiarization were unrelated to those used during test (a smiley face, a crayon and a bicycle). Prior to the testing session, children completed 20 familiarization trials with distractor set sizes varying from 1 to 8.

Results and discussion

Mean reaction times (RT) per distractor set size were calculated for each child. Only correct responses were included. Some participants did not complete all 32 test trials; their data were retained for the trials they did complete. The mean number of completed trials was 31 for both conditions ($SD_{\mathrm{Silent}} = 1.55$, $SD_{\mathrm{Label}} = 3.75$; see Table 1) and no reliable differences were found between conditions in the total number of trials completed [$t(30) = -0.37$, $p = .71$]. Analysis of accuracy revealed no significant main effects of condition [$F(1, 30) = 0.03$, $p = .86$] or distractor set size [$F(3, 90) = 0.40$, $p = .75$; see Table 1]. The interaction of these two factors was not significant [$F(3, 90) = 1.33$, $p = .27$].

Figure 2A depicts mean RT for the Silent and the Label conditions as a function of distractor set size. A mixed 2 × 4 analysis of variance with condition as the between-subjects factor and set size as the within-subjects factor yielded a significant main effect of set size [$F(3, 90) = 27.30$, $p < 0 .001$], reflecting the fact that RT increased as the number of distractors increased. A

Table 1 *Mean RT (ms) for correct responses and mean accuracy per set size, mean slope and intercept of the linear best-fit lines, and mean number of trials completed for each condition of Experiments 1, 2, 3 and 4. For the children who did not met the criterion for contributing data to the RT analysis in Experiment 2, mean percentage of errors is reported instead of the mean RT, accuracy, slope and intercept*

Condition		Distractor set size	Mean RT (*SE*)	Accuracy (*SE*)	Slope (*SE*)	Intercept (*SE*)	Trials completed (*SD*)
Exp. 1 Silent		2	3495 (317)	87 (3)	212 (28)	3264 (212)	31 (1.6)
		4	4327 (442)	89 (3)			
		8	5023 (413)	84 (4)			
		12	5735 (499)	85 (3)			
Exp. 1 Label		2	2721 (201)	88 (4)	233 (5)	2284 (37)	31 (3.8)
		4	3257 (217)	85 (3)			
		8	4129 (317)	86 (5)			
		12	5076 (434)	89 (3)			
Exp. 2 ‘Go’	Criterion met (*N* = 16)	2	3600 (349)	88 (3)	223 (15)	3085 (113)	30 (2.9)
		4	3854 (329)	88 (3)			
		8	4948 (459)	80 (4)			
		12	5744 (407)	80 (5)			
	Criterion not met (*N* = 11)		Mean percentage of errors (*SE*)				30 (4.7)
		2	94 (2)				
		4	97 (1)				
		8	98 (1)				
		12	99 (1)				
Exp. 3 Low Discriminability Silent		3	3849 (296)	79 (4)	160 (51)	3280 (451)	33 (6.3)
		9	4454 (181)	81 (4)			
		12	5376 (348)	79 (4)			
Exp. 3 High Discriminability Silent		3	3898 (340)	90 (2)	39 (6)	3792 (49)	31 (8.1)
		9	4168 (480)	88 (3)			
		12	4236 (337)	93 (2)			
Exp. 3 Low Discriminability Label		3	3281 (165)	94 (2)	168 (39)	2710 (342)	35 (2.3)
		9	4021 (212)	92 (2)			
		12	4860 (243)	90 (2)			
Exp. 3 High Discriminability Label		3	2476 (159)	94 (3)	108 (18)	2122 (159)	34 (4.5)
		9	2996 (166)	93 (2)			
		12	3475 (225)	91 (2)			
Exp. 4 Silent	1st Block	3	3284 (349)	92 (2)	128 (41)	2927 (270)	36 (0)
		9	3985 (611)	93 (2)			
		12	4540 (800)	88 (4)			
	2nd Block	3	2918 (376)	92 (2)	125 (29)	2399 (218)	
		9	3224 (330)	96 (1)			
		12	4110 (498)	95 (2)			
Exp. 4 Label	1st Block	3	2775 (315)	96 (2)	136 (23)	2374 (195)	36 (0)
		9	3594 (491)	91 (3)			
		12	3961 (476)	93 (3)			
	2nd Block	3	2839 (414)	90 (3)	130 (31)	2436 (256)	
		9	3397 (508)	89 (4)			
		12	4036 (637)	93 (3)			

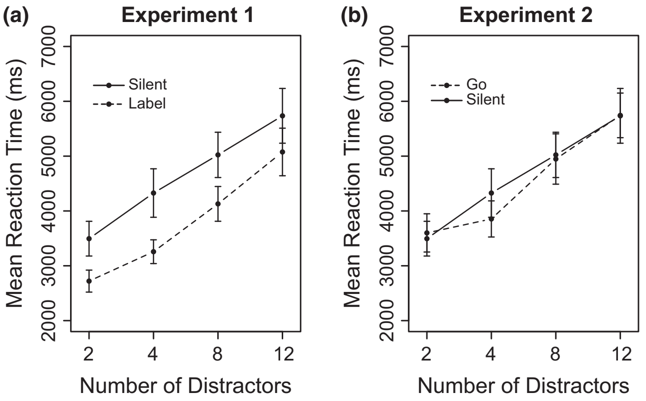

Figure 2 *Left: Mean RT for correct responses across number of distractors for the Silent and Label conditions of Experiment 1. Right: Go condition of Experiment 2 (for comparison purposes, the Silent condition of Experiment 1 is also depicted). Error bars display standard errors of the mean.*

significant main effect of condition was also found [$F(1, 30) = 4.48$, $p < .05$], reflecting a decrease in overall RT for the Label condition. The interaction between set size and condition was not significant [$F(3, 90) = 0.21$, $p = .89$]. The slopes and intercepts of the linear best-fit lines were also calculated for each child. Independent samples t-tests showed that while the slopes of the two conditions were not significantly different [$t(30) = 0.39$, $p = .70$], there was a significant reduction in the intercept of the Label condition when compared to the Silent condition [$t(30) = -2.40$, $p < .05$], reflecting the overall faster search times in the Label condition.

These results thus show a clear effect of labels on 3-year-old children's search time. The positive benefit of naming the search target emerged despite the fact that children in both conditions were visually presented with a preview of the search target on every trial. This suggests that the label does not merely tell children what to search for (information provided by the preview of the target) but influences the way that children encode the search target. The label affected overall search time, but did not affect the slope of the search function. However, by one line of reasoning, labeling might have been expected to reduce the slope of the search function given the present design. Flat search functions (no effect of number of distractors) characterize search tasks in which the target and distractors differ by a single feature (e.g. red versus green; Treisman & Gelade, 1980). If labeling with the basic-level name directed attention in an all-or-none fashion to only the shape-matching items in the array, then the search task would reduce to a one-feature task in which the participant 'saw' only the named shapes (e.g. the beds) and then the one odd-colored bed (the red bed target) would be expected to 'pop out'. Such an all-or-none effect of labeling on the encoding of the target or on search may not have been observed because the shapes of beds and couches are composed of many overlapping line segments. That is, in terms of the shapes alone, the children are presented with a conjunctive search (see Wolfe & Bennett, 1997). This pattern – an effect of labeling on overall search time but not on the slope of the search function – was observed in all the experiments reported in this paper. We consider the broader implications of this pattern in the General Discussion.

Experiment 2

A growing literature shows multimodal influences on visual attention and search such that auditory cues (even non-meaningful ones) may lead to more rapid visual search (Iordanescu, Grabowecky & Suzuki, 2011; Van der Burg, Olivers, Bronkhorst & Theeuwes, 2008). It is thus possible that the effects observed in Experiment 1 were due to the addition of a spoken word – potentially any word – and not to the target's name nor increased attention to category-specific shape. Accordingly, Experiment 2 replicated the Label condition of Experiment 1 but replaced the target name on each trial with the word 'go'.

Method

Participants

Twenty-seven children between 32 and 42 months of age (15 males; mean age: 36 months, *SD*: 2.3) were recruited from the same population as in Experiment 1; none of these children participated in the previous experiment. As in Experiment 1, the criterion for contributing data to RT analyses was at least two correct responses per distractor set size. In contrast to Experiment 1 (and also in contrast to Experiments 3 and 4), a large number of children ($N = 11$) did not meet this criterion and recruitment continued until a sample of 16 children met the criterion for reaction time analyses.

Apparatus, stimuli and procedure

All aspects were the same as in the Label condition of Experiment 1, except that the sound file presented at the onset of the target cuing display played the word 'go'.

Results and discussion

A substantial proportion of children (41% of the sample) in this experiment did the task but failed to reach the criterion set for contributing reaction time data to the analyses, as they selected a distractor object on most test trials. The proportion of children failing to reach criterion in this experiment is reliably greater than in Experiment 1 [$X^2(1, N = 69) = 4.85, p < .05$], suggesting that presenting a word that is not the name of the seen object disrupts children's performance. We first present the analyses of the reaction time data for the 16 children who met criterion and then consider the error patterns for the 11 children who did not.

Mean RT for correct responses was calculated for each child who met the criterion ($N = 16$). On average, these children completed 30 trials ($SD = 2.98$), with an overall mean accuracy of 83% (see Table 1). Figure 2b presents RT for correct responses per distractor set size for the Go condition. For comparison purposes, results from the Silent condition from Experiment 1 are also shown. A mixed 2 × 4 analysis of variance with distractor set size as the within-subjects factor and condition as the between-subjects factor yielded no reliable differences in RT between the Go condition of Experiment 2 and the Silent condition of Experiment 1 [$F(1, 30) = 0.06$, $p = .82$]. A significant main effect of set size was found [$F(3, 90) = 23.82, p < .001$], reflecting the increase in RT as a result of increasing the number of distractors. There was no significant interaction between condition and set size [$F(3, 90) = 0.41, p = .75$]. The analyses of the individual slopes and the intercepts confirmed the trends found for RT: No significant differences were found between the Go condition of Experiment 2 and the Silent condition of Experiment 1 in the slope [$t(30) = 0.25$, $p = .80$] or the intercept [$t(30) = -0.38, p = .71$]. In brief, for these children who found the target on most trials, an auditory word that was not the name of the target did not result in more rapid search than the presentation of no sound at all.

However, for a substantial proportion of the children, an auditory word that was not the name of the target appears to have disrupted their understanding of the task or their ability to keep the target in mind. That is, in contrast to Experiment 1, a substantial proportion of children performed so poorly in this task that they were unable to find the target on most trials (see Napolitano & Sloutsky, 2004; Sloutsky & Napolitano, 2003, for potentially related results). For these children, the overall correct performance was only 13% (see Table 1). Nonetheless, and despite their overall high rate of errors, proportion of errors was reliably related to distractor set size [$F(3, 30) = 5.01, p < .01$] with these children better able to find the target in smaller than in larger search arrays.

In sum, the better performance of children in the Label than Silent condition of Experiment 1 does not appear to be due to a generalized benefit of an auditory signal just prior to search but instead appears to reflect the specific benefit of hearing the target's name.

Experiment 3

Experiments 1 and 2 presented children with targets specified by both their shape and color and with distractors that matched the target on one of those properties. The spoken name of the target enabled children to more rapidly find the specific target in Experiment 1, presumably by guiding their attention to the shape-matching over the color-matching items in the array. However, it is also possible that hearing the name of the target guided children's attention to the color-matching objects in the array. But if category name cues attention to shape because of its association with category-relevant shape properties, hearing the name of the target object should also benefit attention to category-specific shape (and not just to shape over color). Thus, Experiment 3 examined this possibility by asking children to search for targets (pictures of basic-level category instances) among distractors (instances of other categories) that differed only in shape. We employed two stimulus conditions as shown in Figure 3, one in which the target and distractor shapes were very different overall and one in which target and distractor shapes were highly similar. In both cases, the objects were composed of multiple line elements and thus might be considered as instances of a conjunctive search task, as finding the target depends on attending to multiple line elements in the right configuration for the category-relevant shape (e.g. balloon versus ice cream cone). While the Low Discriminability condition clearly requires attending to the configural properties of multiple elements to discriminate between the distractors and the target, the High Discriminability might not as the targets and distractors could be discriminated on a single shape feature (e.g. vertically elongated versus round). Through this manipulation we sought evidence on whether labels are more helpful when more features – and specifically category-relevant configurations of features – are required to discriminate target from distractors. A second question was whether labels would interact with discriminability to produce effects on both the intercept and the per item cost of distractors. Past research with adults has shown steeper slopes when the target and the distractors are hard to discriminate (Duncan & Humphreys, 1989; see Scerif, Cornish, Wilding, Driver & Karmiloff-Smith, 2004, for a similar finding in a non-RT task with children); if our stimulus manipulation is valid for

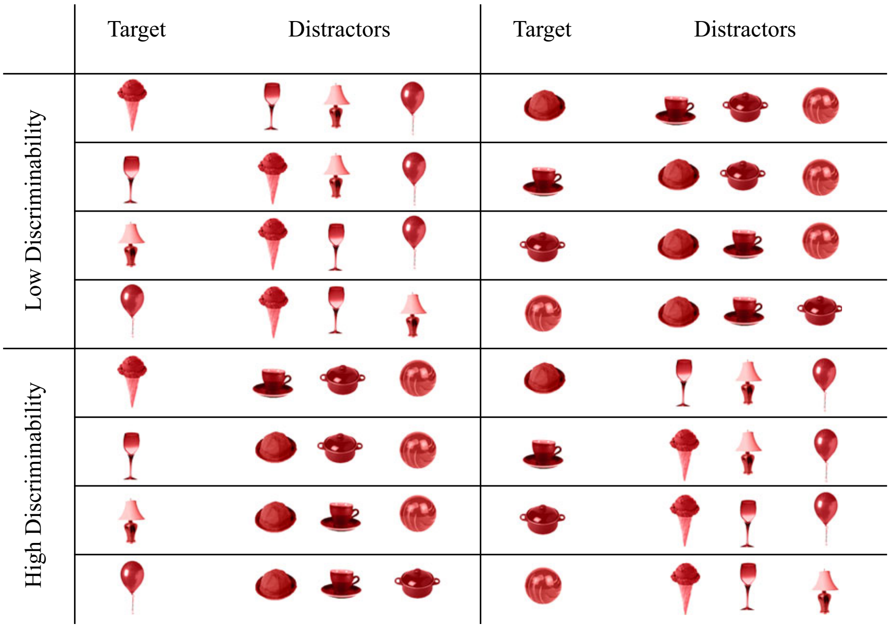

Figure 3 *Set of stimuli used in Experiment 3. Each target could be placed amidst low discriminability (top) or high discriminability (bottom) distractors. Each row indicates the stimuli pairings used for the two blocks of trials (see text for details).*

children, steeper slopes in the High than Low Discriminability condition are expected, at least in the Silent condition. The open question is how the addition of the label affects the intercept and slope measures of performance. If the label activates category-relevant shape representations, then search in both the High and Low discriminability conditions might be similar, as in both cases the label should cue children to attend to the configuration of features that defines the category shape.

In sum, the experiment consisted of a two-by-two (all between-subjects) design in which the Silent and Label conditions were each realized in two stimulus conditions, one in which the shapes of target and distractors were easily discriminable and the other in which the shapes were more difficult to discriminate.

Method

Participants

Sixty-four children between 31 and 41 months of age (35 males; mean age: 36 months, *SD*: 2.6) were randomly assigned to one of four conditions (Low Discriminability – Silent, Low Discriminability – Label, High Discriminability – Silent, or High Discriminability – Label). None of these children participated in the two previous experiments. Twelve additional children were recruited but not included in the final sample due to refusal to participate in the study ($N = 4$), selecting a non-target object on most test trials and thereby not meeting the criterion of at least two correct responses per distractor set size ($N = 2$), not finishing the familiarization phase ($N = 2$), parental interference ($N = 3$) and experimenter error ($N = 1$). Recruitment and informed consent procedures were the same as in the previous experiments.

Apparatus, stimuli and procedure

In contrast to Experiments 1 and 2, targets and distractors differed only in shape. To ensure that the labeling effects observed in Experiment 1 were generalizable to other basic category shapes and category-level names, we used pictured instances of eight different basic-level categories and their common names. Each child participated in two blocks, one block with one target and another block with a second target (order counterbalanced across subjects). For each child, no stimuli were repeated across the two blocks (see Figure 3) and both blocks were instantiations of the same Labeling and Discriminability conditions. We used two blocks with different targets and distractors to increase the number of trials per array set size without increasing practice in the search for specific target (the issue addressed in Experiment 4). In order to increase the number of trials per set size, we also used only three distractor set sizes (3, 9, and 12). Equal numbers of these trials yields 18 total search trials per block, with an order randomly determined for each subject.

The eight pictures of the eight different categories were taken from the 'Massive Memory' database (Konkle, Brady, Alvarez & Oliva, 2010). They were recolored in red scale and rendered in a 100 × 90 pixel area on a white background. The pictures were selected to yield two groups of four images each, elongated shapes (e.g. ice cream cone, glass) versus round shapes (e.g. ball, hat). In the High Discriminability conditions, the target was placed amidst distractors of different overall shape, while in the Low Discriminability conditions the same targets were placed amidst distractors with a similar overall shape (see Figure 3). The differences in shape were confirmed by calculating the amount of shape overlap (i.e. number of pixels shared) between target and distractors when centers were aligned: The mean overlap ratio in the Low Discriminability condition was 0.89 (SD = 0.03) and the mean overlap ratio in the High Discriminability condition was 0.73 (*SD* = 0.05). Prior testing using a forced-choice procedure ensured that children in this age range recognized the stimulus pictures by name ($N = 9$, $M_{Accuracy} = 0.86$, $SD = 0.15$). All other aspects of the procedure were the same as in Experiment 1.

Results and discussion

Mean RT as a function of number of distractors was calculated for each child, collapsed across the two blocks. Only correct responses were included. The slopes and intercepts of the linear best-fit lines were also calculated for each child. Accuracy was above 80% for all conditions (see Table 1 for accuracy per condition). Analyses of accuracy yielded a significant main effect of Labeling [$F(1, 60) = 7.35$, $p = .009$], with accuracy higher in the Label condition, and a main effect of Discriminability [$F(1, 60) = 4.67$, $p = .03$], with accuracy higher in the High Discriminability condition. There was an interaction between Labeling and Discriminability on accuracy [$F(1, 60) = 4.34$, $p = .04$], as the difference in accuracy between the two labeling conditions was larger for the Low Discriminability condition. There was no significant main effect nor interactions with distractor set size (all $p > .05$) on accuracy. No reliable main effects of Labeling or Discriminability and no interactions were found for number of trials completed (all $p > .05$; see Table 1 for number of trials completed per condition).

We first considered the effect of Discriminability in the RT of the Silent condition, to ensure that this stimulus manipulation was effective and that the pattern in this condition replicated adult findings. Figure 4A shows the mean RT in the High and Low Discriminability arrays in the Silent conditions, and Table 1 provides the mean slopes and intercepts. The pattern in the Silent condition shows clear differences between the Low and High Discriminability sets, indicating the effectiveness of our manipulation. Moreover, the pattern is consistent with findings from adults: The discriminability of the target from the distractors affects the cost of additional distractors, showing reliable differences in slopes [$t(30) = 7.43$, $p = .01$] but not intercepts [$t(30) = 1.04$, $p = .32$] in the absence of labels.

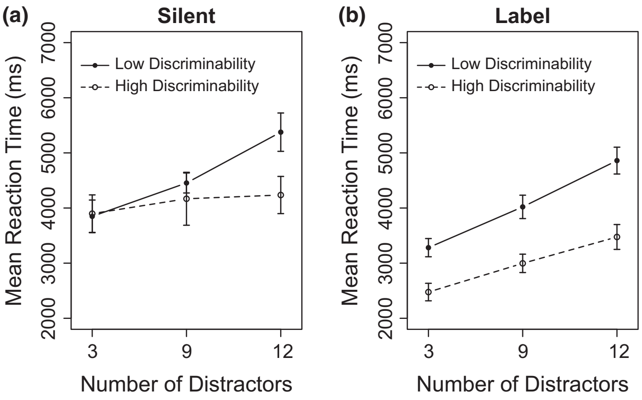

Figure 4 *Mean RT for correct responses across number of distractors for Silent (left) and Label (right) conditions of Experiment 3. Within each condition, targets could be placed amidst low discriminability (solid lines) or high discriminability (dashed lines) distractors (see text for details). Error bars display standard errors of the mean.*

To assess the effects of labeling on this pattern, the mean RT for each participant was entered into a mixed 2 × 2 × 3 analysis of variance with Discriminability (High, Low) and Labeling (Silent, Label) as the between-subjects factors, and distractor set size as the within-subjects factor (see Figure 4). The analysis yielded a significant main effect of Labeling [$F(1, 60) = 11.10$, $p < .01$], resulting from the overall lower RT in the Label conditions – replicating the main finding from Experiment 1 that labels decrease overall search time. This specific result thus extends those of Experiment 1 by showing that the labeling effect occurs even when targets and distractors differ only in shape. The analysis also yielded a significant main effect of Discriminability [$F(1, 60) = 9.86$, $p < .01$], as participants were overall faster when the target was placed amidst distractors that were easier to discriminate from the target, and a significant main effect of distractor set size [$F(2, 120) = 44.15$, $p < .001$] reflecting the increase in RT as number of distractors increased. There was no reliable interaction between Labeling and Discriminability [$F(1, 60) = 1.58$, $p = .21$], showing that labeling the target object benefited both High and Low Discriminability sets. The only reliable interaction was between discriminability and set size on RT [$F(2, 120) = 7.32$, $p < .001$]: Search times were less affected by the number of distractors in the High Discriminability conditions, a finding that implicates differences in the slope of the search functions between the High and Low discriminability conditions.

An analysis of variance with slope as the dependent variable yielded only a significant main effect of Discriminability [$F(1, 60) = 12.93$, $p < .001$]. The main effect of Labeling was not reliable [$F(1, 60) = 2.31$, $p = .13$], nor was the interaction [$F(1, 60) = 1.45$, $p = .23$]. Thus, discriminability *but not labeling* showed clear effects on cost of additional distractors to search time. In contrast, the analysis of intercepts yielded a reliable main effect of Labeling [$F(1, 60) = 16.18$, $p = .00$], but no reliable main effect of Discriminability [$F(1, 60) = 0.02$, $p = .89$]. However, for the intercept measure, the interaction between Labeling and Discriminability approached conventional standards of significance [$F(1, 60) = 3.90$, $p = .05$]. This marginal effect is likely due to the steeper slope in the High Discriminability – Label condition than in the High Discriminability – Silent condition. However, pairwise comparisons of the mean slopes did not yield reliable differences in slope between the Silent and Label conditions for both High [$t(30) = 1.78$, $p = .17$ with Bonferroni correction] and Low Discriminability [$t(30) = 0.24$, $p = 1.00$ with Bonferroni correction] arrays. Thus, labeling the target speeded overall search but may not affect the per item decision time.

In sum, and within the limits of these measures with young children, the results of Experiment 3 support three conclusions: First, labeling the target decreases overall search time in a task in which only shape varied, a result consistent with the hypothesis that labeling the target enhances the working memory encoding and representation of category-relevant shape. Second, labeling affects overall search time but not the slope for both easy and hard to discriminate targets and distractors. This suggests that children were not treating the High Discriminability condition as a single feature search and that even in this easy-to-discriminate condition the label may have led to the encoding of category-relevant shape (see Lupyan, 2008, Experiment 3). Third, discriminability of the target and distractor principally affected the slope of the search function, a result consistent with previous findings in adults (Duncan & Humphreys, 1989).

Experiment 4

Our hypothesis is that hearing the basic-level category name of the search target leads to better encoding of the category-relevant object shape in working memory. This active representation of the target is hypothesized to guide attention to the items in the array that better match that representation, thereby decreasing the overall search time. However, children were shown a preview of and searched for the very same target on every trial, in both the Silent and Label conditions. One might expect that the repeated visual presentations in the Silent conditions would lead to progressively more robust representations of the target and thus to faster search, a result that has been found in adult studies (Kristjánsson & Campana, 2010; Rabbitt, Cumming & Vyas, 1979; Schmidt & Zelinsky, 2009; Vickery *et al.*, 2005; Yang & Zelinsky, 2009). One possibility is that the repeated presentations of the target in Experiments 1 through 3 were leading to progressively more rapid (i.e. more 'label-like') search times in the Silent conditions, but that it took some time for these effects to emerge. This possibility could be addressed by examining children's performance over time (e.g. first vs. second half of the experiments). However, in the prior experiments, the designated distractor set size for a trial was randomly determined for each participant and therefore set size was not equated across the two halves. Accordingly, Experiment 4 replicated the Low Discriminability conditions of Experiment 3 but asked children to search for the same target throughout the entire experiment, with the trials partitioned into two blocks such that there were equal numbers of trials at each set size in the first and second half.

If object names rapidly lead to robust representations of object shape that then drive more rapid search, a labeling effect should be clearly evident even in the first half of the experiment. If repetitions of the visual target lead, more slowly, to robust representation of the target, then search times in the Silent condition should improve from first to second half. The principal effect of labels may be that they shortcut visual learning from repeated presentations.

Methods

Participants

Forty children between 30 and 42 months of age (23 males, mean age: 36 months, *SD*: 3.1) were randomly assigned to either the Silent or the Label condition. None of these children had participated in the previous experiments. Ten additional children were recruited but not included in the final sample due to refusal to participate in the study ($N = 5$), selecting a non-target object on most test trials and thereby not meeting the criterion of at least two correct responses per distractor set size ($N = 3$), not finishing the familiarization phase ($N = 1$) and experimenter error ($N = 1$). Because this experiment was designed to address the role of repeated search on the working memory representation of the target, the final sample included only children who finished all test trials – 10 additional children did not meet this criterion and were therefore not included in the analysis; on average, this group of children completed 28 test trials ($SD = 3.9$). Recruitment and informed consent procedures were the same as in the previous experiments.

Apparatus, stimuli and procedure

Children were asked to find the same target picture across 36 test trials. In order to investigate performance over time, and in contrast with the previous experiments, there were an equal number of trials at each distractor set size (3, 9 and 12) on each block of 18 trials. Because both the Low and High discriminability conditions of Experiment 3 yielded labeling effects, this experiment replicated only the Low discriminability conditions (Silent and Label). Moreover, to increase the likelihood of detecting changes in RT over time, we used only two targets (ice cream cone and ball). All other aspects of the procedure were the same as in Experiment 3.

Results and discussion

Mean RT as a function of distractor set size, and the slopes and the intercepts of the linear best-fit lines, were calculated for each child. Only correct responses were included. Mean accuracy was above 90% for both conditions (see Table 1). Analyses of accuracy revealed no significant main effects of condition [$F(1, 38) = 0.10$, $p = .75$] or set size [$F(2, 76) = 0.01$, $p = .99$]. The two factors did not interact [$F(2, 76) = 2.72$, $p = .07$].

Figure 5 shows the mean RT in the Silent and Label conditions for the first and second half of the task, and Table 1 provides the mean slopes and intercepts. In the first block of 18 trials, the presentation of the target name seems to have speeded up search, similar to the previous experiments. However, by the second block, the time it took to find the target was comparable in the Silent and Label conditions. The mean RT for each child and block was entered into a mixed 2 × 3 × 2 analysis of variance with Labeling (Silent, Label) as the between-subjects factor, and distractor set size and Block (First, Second) as the within-subjects factors. There was a reliable main effect of distractor set size [$F(2, 76) = 39.9$, $p < .001$], reflecting the increase in RT with increasing number of distractors. There was also a reliable main effect of Block [$F(1, 38) = 4.0$, $p = .05$], as RT was overall lower on the second block. Although there was no reliable main effect of Labeling [$F(1, 38) = 0.9$, $p = .34$], there was a significant interaction between Labeling and Block [$F(1, 38) = 4.4$, $p < .05$], suggesting that the difference between the two labeling conditions was modulated by the repetition of the visual information. There were no other significant interactions. An analysis of variance with the intercept as the dependent variable failed to find any significant main effects or interactions as did the corresponding analysis of the slopes. This likely reflects the lack of power when the trials are partitioned into first and second half with just 18 trials (and 6 per set size) per half.

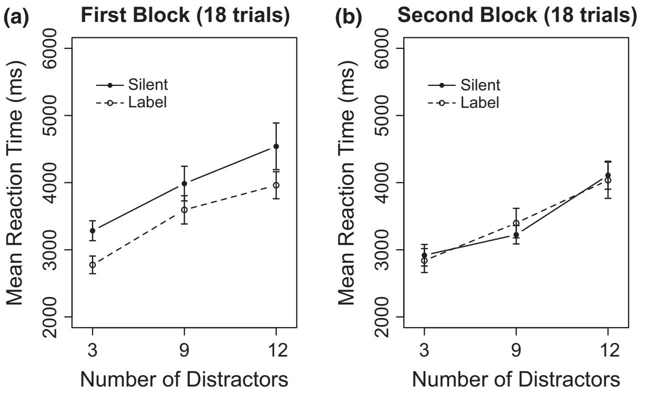

Figure 5 *Mean RT for correct responses across number of distractors for the Silent (solid lines) and the Label (dashed lines) conditions for the first (left) and second (right) block of 18 trials in Experiment 4. Error bars display standard errors of the mean.*

The results of Experiment 4 thus offer converging evidence to the hypothesis that the label influences the robustness of visual representations. Within the limits of our measures with 3-year-old children (for whom 36 total trials is quite demanding), the results suggest that hearing a label quickly enabled children to more rapidly find the target object amidst distractors and that repeated visual exposures to the target more incrementally led children to just as rapid search. The pattern fits the hypothesis that labeling, by activating category-specific shape features of the target in visual working memory, resulted in faster performance with less repetition of the visual information.

General discussion

The four experiments reported here show that hearing the name of an object improves 3-year-old children's ability to find that object in an array. The effect of the object name in speeding children's performance in visual search was found in all experiments that manipulated labeling, with targets and distractors that varied in shape and color or that varied in shape alone, and when the target and distractors were of high and low discriminability. The results are the first showing labeling effects on performance in visual search in children this young and they indicate that the influence of language on visual processing begins early. The working hypothesis behind the design of the experiments was motivated by previous research with adults on the role of visual working memory representations in visual search (Kristjánsson *et al*., 2002; Schmidt & Zelinsky, 2009; Soto *et al*., 2005; Soto *et al*., 2006; Soto & Humphreys, 2007; Vickery *et al*., 2005; Yang & Zelinsky, 2009) and also by developmental evidence on the influence of common nouns on the visual encoding of objects (Gershkoff-Stowe *et al*., 2006; Yoshida & Smith, 2003a). Although the present results are consistent with these interpretations, strong conclusions – given the paucity of prior work on visual search in very young children – are not warranted. Nonetheless, the present results are a first step toward understanding the mechanisms through which language influences visual attention and they raise new testable hypotheses about these mechanisms.

Within the limits inherent to collecting RT data from 3-year-old children, the pattern to be explained is this: Labeling affected overall search time as measured by the intercept but did not affect the additional cost of each added distractor. In this way, and as shown in Experiments 3 and 4, the effect of labeling does not mimic the effect of target and distractor discriminability but does mimic the effect of repeated presentations of the visual target, with labeling accomplishing at the outset what repeated visual presentations accomplish only after some number of repetitions. What might explain this pattern? We have proposed that hearing the target name in some way strengthens the representation of the visual target in working memory and that this stronger representation guides attention to the target item in the array. Such an effect would lead to faster overall search. But why don't more robust representations also not lead to easier discrimination of target from distractor and thus an effect on the slope of the search function?

One way to think about these issues is in terms of two possible ways that children could compare an item being fixated in the array to the target being represented in working memory. These are illustrated in Figure 6. In the approach illustrated in Figure 6a, the child randomly fixates items in the search array. The item upon which the child is fixating at any moment is the driver of the comparison to the target held in memory. If the item being fixated at a given moment is sufficient to remind the child of the target, the two are compared and a decision is made about whether the item is the target or not; if it is not similar enough to activate the target in working memory, then the child moves on to the next item in the array. Given this approach – an inactive memory of the target that is activated only by a similar-enough fixated item – the slope of the search function and decision time per item should depend on discriminability; we propose this approach might best describe children's performances in the Silent conditions. In the approach illustrated in Figure 6b, on the other hand, the target in visual working memory is continually active and is the driver of which items are fixated, either by pulling visual attention to matching objects or by suppressing attention to non-matching objects. If hearing a label fosters this second approach and the Silent conditions foster the first approach, then labeling would result in faster overall search without necessarily changing per item cost to the decision time about each distractor. This hypothesis fits findings using eye-tracking methodology

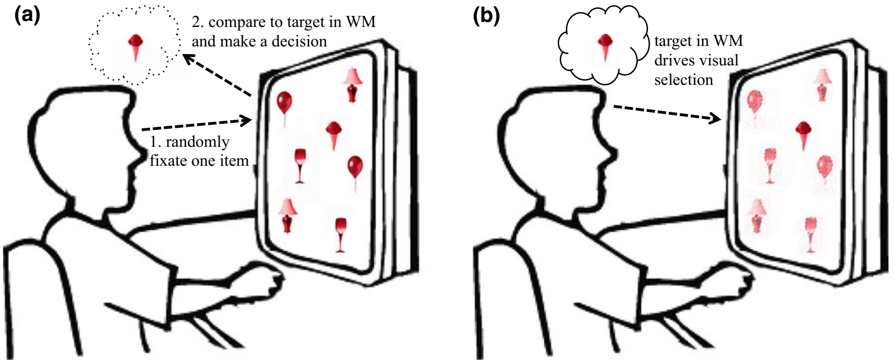

Figure 6 *Illustration of two possible ways to compare an item in the search array to the working memory representation of the target. Left: Items are randomly fixated and then compared to the target held in memory. Right: The continually active target representation increases the likelihood of fixating the target in the array (see General Discussion for details).*

in adults (Soto *et al*., 2005; Soto *et al*., 2006) and a direct test of where children look in search tasks (when the target is labeled versus not) is clearly the next step in the present research program.

The present hypothesis, based on the effects of labeling on children's categorizations (Gershkoff-Stowe *et al*., 2006; Yoshida & Smith, 2003a), is that the label affected visual encoding of the target when the visual preview was displayed, which decreased search times. But clearly words can have effects on visual expectations – and where one looks in an array – without a visually presented target. For example, adults listening to spoken sentences look at a possible visual referent even when the visual array is irrelevant to the task (for a review see Huettig, Rommers & Meyer, 2011b), and even look to shape-similar items when that item is clearly not the referent of the uttered word (e.g. to a rope when hearing a sentence about a snake; Huettig & Altmann, 2007). This effect of words on where people look is used to study on-line language comprehension in adults (Huettig *et al*., 2011b), young children (Fernald *et al*., 2010), and even infants (Bergelson & Swingley, 2012). By hypothesis, the underlying mechanism for these effects may be fundamentally the same as the one proposed here: Heard words yield representations (either expectations or biased encoding of seen things) in visual working memory and these active representations then drive where one looks in a scene.

The two hypothesized approaches to search illustrated in Figure 6 differ principally in whether the target representation is active throughout search or whether it is activated upon seeing a similar enough item in the array. That is, the key effect of hearing a label may be to keep the target active during search and thus able to influence where the participant looks in the search array. Labeling, in this way, might be viewed as akin to active rehearsal in maintaining working memory representations (Baddeley & Hitch, 1994). Interestingly, an adult study has shown that participants were overall faster in a visual search task if they were instructed to actively rehearse the object name (Lupyan & Swingley, 2011). One might hypothesize that in the present experiments, hearing a label encouraged children to covertly repeat the object name and this active rehearsal was key to their keeping the target active in memory, and thus able to guide search. Although we cannot rule out this possibility, it seems unlikely as verbal rehearsal is a late developmental achievement, not robust until middle childhood (Flavell, Beach & Chinsky, 1966; Gathercole, 1998; Jarrold & Tam, 2011), and nearly impossible to teach young children to do (Keeney, Cannizzo & Flavell, 1967). Still, this is a possibility that merits future consideration.

The current findings also have implications for understanding why children are more likely to group objects by shape when they are named (Landau *et al*., 1992). The shape bias in children's noun learning does not just concern the effect of known names on categorization but also the effect of novel names on the categorization – and name generalization of novel things. By one account, this shape bias emerges as a second-order generalization across known names and categories and is cued by the common linguistic contexts of naming things (Colunga & Smith, 2005; Smith, Jones, Landau, Gershkoff-Stowe & Samuelson, 2002). Evidence for this account derives from correlational findings showing a developmental relation between knowing object names and attending to object shape (e.g. Gershkoff-Stowe & Smith, 2004; Smith, 1995; Smith, Colunga & Yoshida, 2010), from experimental findings teaching naming biases to very young children (Smith *et al*., 2002; Samuelson, 2002; Perry, Samuelson, Malloy & Schiffer, 2010), and from computational models showing how a generalized shape bias could emerge as a higher-order generalization from the known names of specific categories (Colunga & Smith, 2005). The present findings offer a potential pathway to understand the in-task mechanisms that lead to biased attention to object shape in naming tasks: Once a child has learned the names of a sufficient number of basic-level categories, naming – even with novel names – may lead to the biased encoding of shape and more active visual working memory representations that then guide attention in novel name learning tasks.

The present findings and discussion are also relevant to a large literature on the development of working memory in children. This literature shows a protracted developmental course characterized by two critical changes: An increase in the number of items that can be stored in working memory (Cowan & Alloway, 2009) and an increase in the precision and stability of those representations (Heyes, Zokaei, van der Staaij, Bays & Husain, 2012). These developmental changes, which appear to characterize both auditory and visuospatial working memory, have also been linked to a variety of developing cognitive skills – including reading, mathematics, executive control and language learning (Archibald & Gathercole, 2007; Bull & Scerif, 2001; Cowan & Alloway, 2009; Gathercole, Alloway, Willis & Adams, 2006). Individual differences in working memory have also been implicated in a number of developmental disorders (Alloway, Gathercole, Kirkwood & Elliott, 2009), including in children with language delays (Montgomery, 2003; Weismer, Evans & Hesketh, 1999) who also do not show a shape bias in early noun learning (Jones, 2003; Jones & Smith, 2005; cf. Weismer & Evans, 2002). The present results – by implicating a role for words in the quality of visual working memory representations – may provide new paths for understanding the development and individual differences in working memory processes.

In conclusion, the present results document for the first time a role for object names in directing visual attention in young children in a visual search task. The results also document visual search processes in 3-year-olds that include a dissociation of the effects of labels and target–distractor discriminability, with the labels affecting the intercept of the search function but not its slope and discriminability affecting the slope but not the intercept. The pattern fits the hypothesis that labels influence the encoding and the maintenance of the target in working memory, an idea that has broad implications for understanding how heard words affect visual processing and performance in many cognitive tasks.

Acknowledgements

We would like to thank the members of the Cognitive Development Lab at IU for useful discussions on this project, the two anonymous reviewers for their very helpful comments on the paper, and Angela AuBuchon for references on verbal rehearsal. We are also thankful to Anna MacKinnon, Blakely Meyer and Tracy Kelsey for their help with stimuli creation, recruitment and data collection, and the parents and children who participated in these studies. This work was supported by a grant from the National Institute of Child Health and Development (HD28675) to LBS and a Graduate Fellowship from the Portuguese Foundation for Science and Technology (SFRH/BD/68553/2010) awarded to CV.

References

Alloway, T.P., Gathercole, S.E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development*, **80** (2), 606–621.

Archibald, L.M.D., & Gathercole, S.E. (2007). The complexities of complex span: specifying working memory deficits in SLI. *Journal of Memory and Language*, **57**, 177–194.

Baddeley, A.D., & Hitch, G.J. (1994). Developments in the concept of working memory. *Neuropsychology*, **8** (4), 485–493.

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences, USA*, **109** (9), 3253–3258.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: inhibition, switching, and working memory. *Developmental Neuropsychology*, **19** (3), 273–293.

Chevalier, N., & Blaye, A. (2009). Setting goals to switch between tasks: effect of cue transparency on children's cognitive flexibility. *Developmental Psychology*, **45** (3), 782–797.

Colunga, E., & Smith, L.B. (2005). From the lexicon to expectations about kinds: a role for associative learning. *Psychological Review*, **112** (2), 347–382.

Cowan, N., & Alloway, T. (2009). Development of working memory in childhood. In M.L. Courage & N. Cowan (Eds.), *The development of memory in infancy and childhood* (pp. 303–342). Hove, East Sussex: Psychology Press.

Duncan, J., & Humphreys, G.W. (1989). Visual search and stimulus similarity. *Psychological Review*, **96** (3), 433–458.

Fernald, A., Thorpe, K., & Marchman, V.A. (2010). Blue car, red car: developing efficiency in online interpretation of adjective-noun phrases. *Cognitive Psychology*, **60** (3), 190–217.

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** (2), 472–479.

Flavell, J.H., Beach, D.R., & Chinsky, J.M. (1966). Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development*, **37**, 283–299.

Garon, N., Bryson, S.E., & Smith, I.M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychological Bulletin*, **134** (1), 31–60.

Gathercole, S.E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry*, **39** (1), 3–27.

Gathercole, S.E., Alloway, T.P., Willis, C., & Adams, A.M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology*, **93** (3), 265–281.

Gerhardstein, P., & Rovee-Collier, C. (2002). The development of visual search in infants and very young children. *Journal of Experimental Child Psychology*, **81**, 194–215.

Gershkoff-Stowe, L., Connell, B., & Smith, L. (2006). Priming overgeneralizations in two- and four-year-old children. *Journal of Child Language*, **33** (3), 461–486.

Gershkoff-Stowe, L., & Smith, L.B. (2004). Shape and the first hundred nouns. *Child Development*, **75** (4), 1098–1114.

Heyes, B.S., Zokaei, N., van der Staaij, I., Bays, P.M., & Husain, M. (2012). Development of visual working memory precision in childhood. *Developmental Science*, **15** (4), 528–539.

Huettig, F., & Altmann, G.T. (2007). Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. *Visual Cognition*, **15** (8), 985–1018.

Huettig, F., & Altmann, G.T.M. (2011). Looking at anything that is green when hearing 'frog': how object surface colour and stored object colour knowledge influence language-mediated overt attention. *Quarterly Journal of Experimental Psychology*, **64** (1), 122–145.

Huettig, F., & Hartsuiker, R.J. (2008). When you name the pizza you look at the coin and the bread: eye movements reveal semantic activation during word production. *Memory & Cognition*, **36** (2), 341–360.

Huettig, F., Olivers, C.N., & Hartsuiker, R.J. (2011a). Looking, language, and memory: bridging research from the visual world and visual search paradigms. *Acta Psychologica*, **137** (2), 138–150.

Huettig, F., Rommers, J., & Meyer, A.S. (2011b). Using the visual world paradigm to study language processing: a review and critical evaluation. *Acta Psychologica*, **137** (2), 151–171.

Iordanescu, L., Grabowecky, M., & Suzuki, S. (2011). Object-based auditory facilitation of visual search for pictures and words with frequent and rare targets. *Acta Psychologica*, **137** (2), 252–259.

Jarrold, C., & Tam, H. (2011). Rehearsal and the development of working memory. In P. Barrouillet & V. Gaillard (Eds.), *Cognitive development and working memory: A dialogue between neo-Piagetian theories and cognitive approaches* (pp. 177–199). New York: Taylor & Francis.

Johnson, E.K., McQueen, J.M., & Huettig, F. (2011). Toddlers' language-mediated visual search: they need not have the words for it. *Quarterly Journal of Experimental Psychology*, **64** (9), 1672–1682.

Jones, S.S. (2003). Late talkers show no shape bias in a novel name extension task. *Developmental Science*, **6** (5), 477–483.

Jones, S.S., & Smith, L.B. (1993). The place of perception in children's concepts. *Cognitive Development*, **8** (2), 113–139.

Jones, S.S., & Smith, L.B. (2005). Object name learning and object perception: a deficit in late talkers. *Journal of Child Language*, **32** (1), 223–240.

Keeney, T.J., Cannizzo, S.R., & Flavell, J.H. (1967). Spontaneous and induced verbal rehearsal in a recall task. *Child Development*, **38**, 953–966.

Konkle, T., Brady, T.F., Alvarez, G.A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long--term memory for real-world objects. *Journal of Experimental Psychology: General*, **139** (3), 558–578.

Kristjánsson, Á., & Campana, G. (2010). Where perception meets memory: a review of repetition priming in visual search tasks. *Attention, Perception, & Psychophysics*, **72** (1), 5–18.

Kristjánsson, Á., Wang, D., & Nakayama, K. (2002). The role of priming in conjunctive visual search. *Cognition*, **85**, 37–52.

Landau, B., Smith, L.B., & Jones, S.S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, **3** (3), 299–321.

Landau, B., Smith, L.B., & Jones, S.S. (1992). Syntactic context and the shape bias in children's and adults' lexical learning. *Journal of Memory and Language*, **31** (6), 807–825.

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

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** (12), 1077–1083.

Lupyan, G., & Spivey, M.J. (2010a). Making the invisible visible: verbal but not visual cues enhance visual detection. *PloS ONE*, **5** (7), e11452.

Lupyan, G., & Spivey, M.J. (2010b). Redundant spoken labels facilitate perception of multiple items. *Attention, Perception, & Psychophysics*, **72** (8), 2236–2253.

Lupyan, G., & Swingley, D. (2011). Self-directed speech affects visual processing. *Quarterly Journal of Experimental Psychology*, **65**, 1068–1085.

Mani, N., Johnson, E., McQueen, J.M., & Huettig, F. (2013). How yellow is your B\banana? Toddlers' language-mediated visual search in referent-present tasks. *Developmental Psychology*, **49** (6), 1036–1044.

Montgomery, J.W. (2003). Working memory and comprehension in children with specific language impairment: what we know so far. *Journal of Communication Disorders*, **36** (3), 221–231.

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.

Perry, L.K., Samuelson, L.K., Malloy, L.M., & Schiffer, R.N. (2010). Learn locally, think globally: exemplar variability supports higher-order generalization and word learning. *Psychological Science*, **21** (12), 1894–1902.

Rabbitt, P., Cumming, G., & Vyas, S. (1979). Modulation of selective attention by sequential effects in visual search tasks. *Quarterly Journal of Experimental Psychology*, **31** (2), 305–317.

Samuelson, L.K. (2002). Statistical regularities in vocabulary guide language acquisition in connectionist models and 15–20-month-olds. *Developmental Psychology*, **38** (6), 1016–1037.

Scerif, G., Cornish, K., Wilding, J., Driver, J., & Karmiloff-Smith, A. (2004). Visual search in typically developing toddlers and toddlers with Fragile X or Williams syndrome. *Developmental Science*, **7** (1), 116–130.

Schmidt, J., & Zelinsky, G.J. (2009). Search guidance is proportional to the categorical specificity of a target cue. *Quarterly Journal of Experimental Psychology*, **62** (10), 1904–1914.

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.

Smith, L.B. (1995). Self-organizing processes in learning to learn words: development is not induction. In C.A. Nelson (Ed.), *Basic and applied perspectives on learning, cognition, and development* (pp. 1–32). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Smith, L.B., Jones, S.S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, **13** (1), 13–19.

Smith, L.B., Jones, S.S., Yoshida, H., & Colunga, E. (2003). Whose DAM account? Attentional learning explains Booth and Waxman. *Cognition*, **87** (3), 209–213.

Solman, G.J., Cheyne, J.A., & Smilek, D. (2011). Memory load affects visual search processes without influencing search efficiency. *Vision Research*, **51** (10), 1185–1191.

Soto, D., Heinke, D., Humphreys, G.W., & Blanco, M.J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance*, **31** (2), 248–261.

Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G.W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences*, **12** (9), 342–348.

Soto, D., & Humphreys, G.W. (2007). Automatic guidance of visual attention from verbal working memory. *Journal of Experimental Psychology: Human Perception and Performance*, **33** (3), 730–753.

Soto, D., Humphreys, G.W., & Heinke, D. (2006). Working memory can guide pop-out search. *Vision Research*, **46** (6), 1010–1018.

Treisman, A.M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, **12**, 97–136.

Van der Burg, E., Olivers, C.N.L., Bronkhorst, A.W., & Theeuwes, J. (2008). Pip and pop: nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, **34** (5), 1053–1065.

Vickery, T.J., King, L., & Jiang, Y. (2005). Setting up the target in visual search. *Journal of Vision*, **5**, 81–92.

Weismer, S.E., & Evans, J.L. (2002). The role of processing limitations in early identification of specific language impairment. *Topics in Language Disorders*, **22** (3), 15–29.

Weismer, S.E., Evans, J., & Hesketh, L.J. (1999). An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, **42** (5), 1249–1260.

Wolfe, J.M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, **9** (1), 33–39.

Wolfe, J.M., & Bennett, S.C. (1997). Preattentive object files: shapeless bundles of basic features. *Vision Research*, **37** (1), 25–43.

Woodman, G.F., Vogel, E.K., & Luck, S.J. (2001). Visual search remains efficient when visual working memory is full. *Psychological Science*, **12** (3), 219–224.

Yang, H., & Zelinsky, G.J. (2009). Visual search is guided to categorically-defined targets. *Vision Research*, **49** (16), 2095–2103.

Yoshida, H., & Smith, L.B. (2003a). Known and novel noun extensions: attention at two levels of abstraction. *Child Development*, **74** (2), 564–577.

Yoshida, H., & Smith, L.B. (2003b). Shifting ontological boundaries: how Japanese- and English-speaking children generalize names for animals and artifacts. *Developmental Science*, **6** (1), 1–17.

Received: 9 March 2013
Accepted: 13 December 2013