Title
A Question of Timing: The Impact of Label Synchrony on Infants' Categorisation

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 35(35)

ISSN
1069-7977

Authors
Althaus, Nadja
Plunkett, Kim

Publication Date
2013

Peer reviewed

A Question of Timing: The Impact of Label Synchrony on Infants' Categorisation

Nadja Althaus (nadja.althaus@psy.ox.ac.uk)
Department of Experimental Psychology, University of Oxford
South Parks Road, Oxford OX1 3UD, United Kingdom

Kim Plunkett (kim.plunkett@psy.ox.ac.uk)
Department of Experimental Psychology, University of Oxford
South Parks Road, Oxford OX1 3UD, United Kingdom

Abstract

Recent research into the impact of labelling on infants' visual category formation has led to controversial results, with some findings indicating a beneficial role and others pointing to interference effects in the presence of labels. Here we present an eye tracking study with 12-month-olds investigating the impact of the label's timing on categorisation. We find that synchronous presentation of words and objects leads to a decreased novelty preference, creating the impression of a dramatic detrimental impact on learning. Asynchronous presentation of the word one second after the image onset does not appear to interfere with processing. Detailed analyses of infants' gaze patterns with respect to object parts reveal that even synchronous labels do not hinder learning but slow down infants' shift from familiarity to novelty preference. Besides offering detailed insight into the effects of labelling on infants' attention our findings offer the potential to reconcile previous contradictory results.

Keywords: Categorisation; cognitive development; language development; eye tracking; attention.

Introduction

The idea of linguistic influences on cognitive processes has been a heavily debated subject over the past century, with very extreme positions like Whorfian determinism (Whorf, 1956) gradually being replaced by less radical points of view (e.g. Boroditsky, 2001). From a developmental perspective, the question is fundamental: do infants use language, and words in particular, as cues to learn about the complex structure of the world? The almost universal presence of labels in an infant's environment, both in speech directed at the infant and in conversation between adults overheard by the infant, makes the hypothesis that labels may serve as meaningful cues very compelling. Shared labels indicate, after all, that dissimilar looking things may share attributes or function (e.g., a bonnet and a boater may both simply be called a "hat"). However, labels are not always readily identified by infants in their first year of life: language development is a gradual process involving learning about relevant dimensions in rhythm, prosody and phonetics before individual words are segmented from the speech stream and mapped onto referents (Tincoff & Jusczyk, 1999; Bortfeld, Morgan, Golinkoff, & Rathbun, 2005; Bergelson & Swingley, 2012). In spite of this, several studies in the past 20 years have found facilitative effects of labelling on categorisation in pre-linguistic infants between six and twelve months (e.g., Waxman & Markow, 1995; Balaban & Waxman, 1997; Waxman & Braun, 2005; Fulkerson & Waxman, 2007), and more recently even in infants as young as three months (Ferry, Hespos, & Waxman, 2010). This work suggests that even infants who are just at the beginning of language development can make use of labels when learning about objects and similarities between them. In fact, Plunkett, Hu, and Cohen (2008) and Althaus and Westermann (in prep.) have demonstrated that labels can serve to guide the formation of category boundaries when the structure of visual space is ambiguous, i.e. labels can cause infants to merge or split visual clusters depending on whether the visual exemplars are encountered with identical or differing labels. However, contradicting results which report "auditory overshadowing" effects in the presence of labels (as well as other auditory stimuli) have also been reported (Robinson & Sloutsky, 2007), calling into question whether labelling has uniformly beneficial effects. It is as of yet unclear under what circumstances labels can facilitate learning, and what factors may contribute to labels attenuating learning.

From an information-processing point of view, labels provide additional information that may help learning – e.g., by increasing perceived similarity between objects that share labels (Sloutsky, Lo, & Fisher, 2001), or by highlighting commonalities (Waxman & Markow, 1995). However, processing this additional signal comes at a cost: attention and processing resources have to be allocated to two modalities rather than one – a feat which may be particularly problematic for young infants. This factor in particular makes it seem likely that the exact circumstances of *how* labels are encountered will play a role in whether they are going to interfere with, or aid, processing.

Here we explore the hypothesis that whether or not labels and objects can both be processed depends on the timing of the label: if both are presented in exact synchrony, i.e. image and label occur simultaneously, this may impose high cognitive load, and processing in one or both modalities may be attenuated. By contrast, if there is a delay between visual onset and auditory onset, this may allow infants to process both stimuli equally well – simply because some visual object recognition processes will already have been completed by the time the label occurs (Quinn, Westerlund, & Nelson, 2006; Grossmann, Gliga, Johnson, & Mareschal, 2009).

Whereas synchronous label onsets have been used in experimental scenarios (e.g., Robinson & Sloutsky, 2007), delayed labelling scenarios appear to be much more likely to occur in a young child's experience (for example, a caregiver asking "Do you like the ball?" when the child is already attending to the toy). Even though some researchers have claimed that synchrony is beneficial for word-object association (Gogate, Bahrick, & Watson, 2003), and cross-modal synchrony has been demonstrated to facilitate discrimination of amodal signals such as tempo or rhythm (Bahrick & Lickliter, 2000), it is likely that synchronous picture-word pairings are unusual and surprising to infants at one year of age. These infants, after all, are at a stage in development where they may have learned some things about words (e.g., they often occur together with their referents, but not generally in synchrony like "causal" sounds, such as a hammer hitting a wall), but are far from being experts at processing speech sounds as phonetic units and mapping them to words.

In order to examine the impact of audiovisual synchrony on the interaction of labelling and categorisation, we familiarised three groups of 12-month-olds with a novel category either in silence (Silent condition), with labels presented one second after the picture onset (Asynchronous Label condition), or with labels and pictures occurring simultaneously (Synchronous Label condition). In order to gain further insight into *how* infants process objects, the target category was constructed such that each exemplar consisted of two spatially separate object parts (a shell and a leaf, see Figure 1 for example objects). The shells were highly variable across exemplars, whereas the leaves were quite similar. This allowed us to track infants' encoding of features with different similarity structure. On test, infants were presented with an out-of-category item, as well as an object consistent with the familiarised category. In familiarisation/novelty preference paradigms, novelty preference for the out-of-category object has been established as an indicator of category formation. However, the two out-of-category items occurring in the first two test trials were constructed by replacing just one of the two parts (shell or leaf) with an item that differed from the familiarisation exemplars. We were therefore able to track on a very fine-grained, almost featural, level whether infants responded to novelty. Our hypothesis was that if infants were able to learn the category in silence, but labels (asynchronous or synchronous) interfered with learning, then infants should not exhibit novelty preference in the relevant conditions. A difference between Asynchronous and Synchronous conditions with regard to novelty preference would further indicate that the timing of the label plays an important role in infants' ability to process and integrate both stimuli.

Methods

Participants

A total of 87 infants participated in this study (mean age: 372 days, 43 girls). Eight additional infants were not included in the analysis due to failure to reach the looking time criterion. Infants were recruited shortly after birth at the local maternity ward and English was the main language spoken in their home.

Stimuli

A novel category was constructed by assembling 11 "objects" from images of a shell, a leaf and a pipe-cleaner in the Gnu Image Manipulation Program (see Figure 1 for example stimuli). Across the different objects, the leaves were very similar, the shells highly variable, and the invariable pipe cleaner served as a connecting limb between these two parts. In addition, three "out-of-category" objects were constructed (see Figure 2): Test object 1 contained a shell consistent with the category, but an inconsistent type of leaf ("novel leaf"), Test object 2 contained a leaf consistent with the category but an inconsistent shell ("novel shell"), and Test object 3 contained a sea urchin and a starfish instead of shell and leaf. All images were depicted against a medium grey background. Objects subtended approximately 14 x 10° visual angle. On test displays, there was a gap of approximately 5° visual angle between out-of-category and within-category objects. A recording of the novel label "timbo", pronounced by a female British-English speaker in an infant-directed voice, served as the auditory stimulus in the Asynchronous and Synchronous Label conditions.

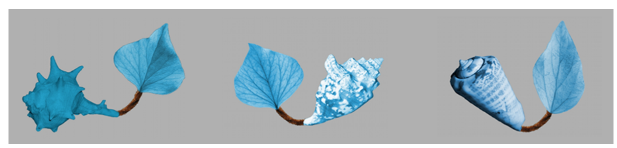

Figure 1: Example familiarisation stimuli.

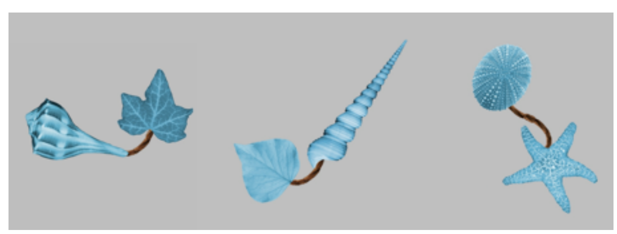

Figure 2: Test objects 1, 2 and 3.

Procedure

After a short warm-up phase during which written consent was obtained from the caregiver, infants were seated on the caregiver's lap at 75 cm distance from the eye tracker. A nine-point calibration sequence was performed up to three times or until all points had been calibrated successfully.

Infants were presented with eight familiarisation images in pseudo-randomised order which were on the screen for 6000 ms each. Four of the familiarisation images appeared on the left half of the screen, and four on the right, in no predictable order. Every image was preceded by an attention getter, a small animation at the centre of the screen (with a medium grey background) accompanied by an attractive chiming sound. Animation and sound lasted about 1.5 seconds, with the next trial beginning 2 seconds after the onset of the attention getter. In the Asynchronous Label condition, the sound file containing the label "timbo" was played 1000 ms after the picture onset. In the Synchronous Label condition, the label started at picture onset. Familiarisation was followed by three test trials, lasting 10 000 ms each. On the test trials, the three test objects described above were paired with one of the three remaining objects from the familiar category. Test trials in all conditions were silent. Infants' looking was recorded using a Tobii eye tracker sampling at 120 Hz throughout the familiarisation and test phase.

Results

We first focus on global measures of looking during familiarisation and test (i.e. with respect to whole objects), and then turn to a more detailed analysis of looking directed at individual object parts. In order to analyse gaze patterns with regard to object parts, areas-of-interest (AOIs) were defined to contain the area covered by the images of shell and leaf, respectively, plus a 30-pixel margin around the image outline (corresponding roughly to the eye tracker's 0.5 degree visual angle accuracy). Recorded gaze data were analysed using custom Matlab code.

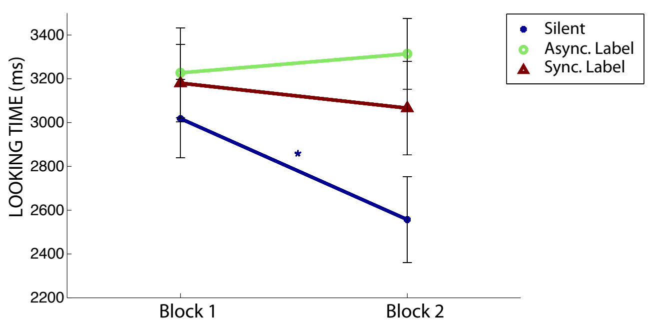

Figure 3: Looking time during familiarisation.

Looking time during familiarisation Looking time for each familiarisation trial was calculated as the sum of fixation time falling on the leaf and shell AOIs. Average looking times for Blocks 1 (Trials 1-4) and 2 (Trials 5-8) are shown in Figure 3. A mixed ANOVA with within-subjects factor Block (Block 1, Block 2) and between-subjects factor Condition (Silent, Asynchronous Label, Synchronous Label) revealed a near-significant main effect of Block ($F(1,84)=3.639$, $p=.06$). Neither the Block x Condition interaction ($F(2,84)=2.231$, $p=.114$) nor the main effect of condition were significant ($F(2,84)=1.76$, $p=.178$). Planned comparisons (paired t-tests) showed that infants' looking decreased in the Silent condition ($t(28)=2.864$, $p=.008$), but not in either of the Label conditions (Asynchronous Label: $t(28)=.46$, $p>.64$; Synchronous Label: $t(28)=1.246$, $p>.22$). This is consistent with previous research showing that auditory input causes infants to maintain their looking (e.g. Baldwin & Markman, 1989; Robinson & Sloutsky, 2007; Plunkett et al., 2008).

Part-based looking during familiarisation In order to investigate whether synchronous or asynchronous labels affected infants' processing of individual parts during familiarisation, we calculated a mean looking proportion for the "leaf" part by dividing the amount of looking at the leaves accumulated during familiarisation by the total looking time accumulated during familiarisation. These data were subjected to a one-way ANOVA with factor Condition, which revealed no significant effect ($F(2,84)=1.226$, $p=.299$). Across all three conditions, infants spent less time looking at the leaves than at the shells, indicating that they were sensitive to the greater variability of the shells (Proportion of looking at leaf, collapsed across conditions: $M=.33$, $SE=.01$; $t(86)=13.1$, $p<.001$).

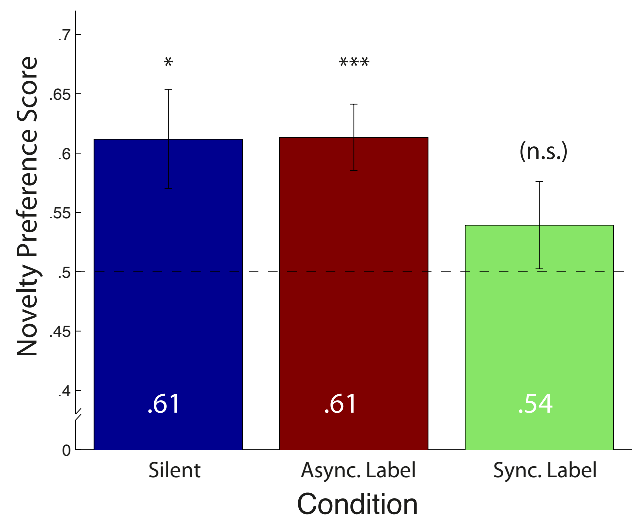

Figure 4: Novelty preference scores in Test 1.

Preferential looking at test (object-based preferences) Object-based novelty preference scores were obtained for all test trials by dividing the amount of looking at the out-of-category object by the total looking time accumulated for the trial (within-category and out-of-category object).

Separate one-way ANOVAs on novelty preference scores with factor Condition were conducted for all three test types, which did not reveal any significant differences between conditions (all Fs<1.4, ps$>.2$). However, planned comparisons were conducted to test each group's performance against the chance level of 0.5 . If infants failed to form a category and

Table 1: Novelty preference scores for test trials 2 and 3 in all conditions. * indicates significance at the .05-level, ** indicates significance at the .005 level, *** indicates significance at the .0005 level.

Condition	Test 2 $M(SE)$	Test 3 $M(SE)$
Silent	.63 (.05)*	.66 (.05)**
Async.	.64 (.04)**	.65 (.04)**
Sync.	.65 (.04)**	.68 (.04)***

did not discriminate between the two objects, we would expect them to spend 50% of their looking directed at each object. By contrast, if they successfully formed a category we would expect them to reliably prefer the out-of-category object.

The results for Test 1 are illustrated in Figure 4. Here, the out-of-category object contained a novel, inconsistent type of leaf and consistent shell. Infants in the Silent and Asynchronous Label conditions exhibited significant preference for the out-of-category object (Silent: $t(28)$=2.679, p=.01; Asynchronous Label: $t(28)$=4.04, $p<.001$). Infants in the Synchronous Label condition, by contrast, failed to prefer the out-of-category object systematically ($t(28)$=1.067, p=.29). Results for Test 2, where the out-of-category object contained a novel shell, and Test 3, where the out-of-category object contained two entirely novel parts, are provided in Table 1. In both Test 2 and 3, *all* infants exhibited systematic novelty preference, even those familiarised with synchronous labels.

In order to further understand the pattern of results in Test trial 1, we obtained the number of infants in each condition who spent more than 50% of looking time on the novel object. This confirmed that infants in the Silent and Asynchronous Label condition were mostly successful at learning the category (Silent: N=20, p=.06, two-tailed binomial test; Asynchronous label: N=22, $p<.01$), whereas the number of successful infants was at chance level in the Synchronous Label condition (N=16, $p>.7$).

Infants' failure to recognise the out-of-category stimulus on Test 1 as novel in the Synchronous condition suggests a detrimental impact of the synchronous label on learning: it seems as though infants in this condition have not encoded the category equally well as infants in the other conditions. While this is in line with the hypothesis that synchronous labels impose greater processing load and therefore visual stimuli may be processed in less detail, it is possible to achieve a more fine-grained insight into infants performance at test. The out-of-category stimulus in Test 1 was designed to be novel owing to the presence of a different type of leaf – its shell part, by comparison, is relatively consistent with the familiarisation category. Examining infants' looking patterns with regard to the two object parts and corresponding parts in the within-category object should therefore provide more information as to how infants process the object.

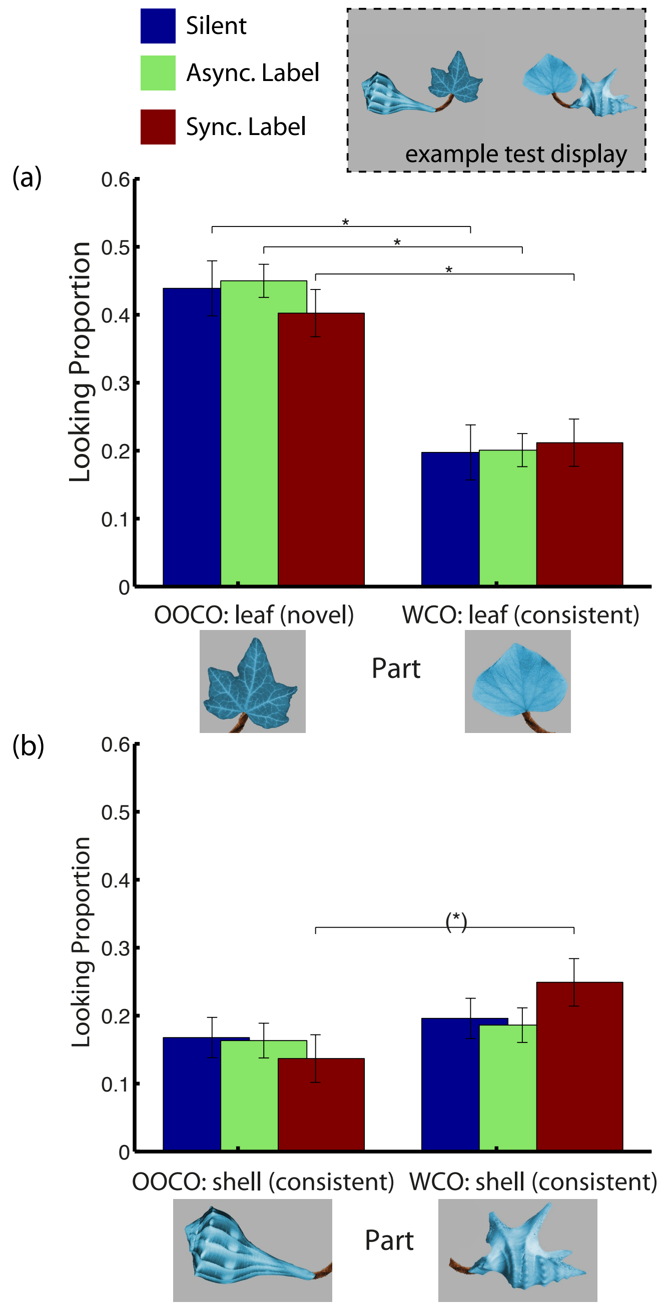

Figure 5: Looking proportions for individual object parts during Test 1: (a) Leaves belonging to out-of-category object (OOCO), here the "novel" part, and within-category object (WCO); (b) Shells belonging to OOCO and WCO. The inset shows an example test display with both objects. * indicates a statistically significant difference, (*) indicates a trend.

Part-based looking at test Figure 5 (a) shows looking proportions directed at the leaves belonging to the two objects in Test 1, i.e. out-of-category leaf and within-category leaf, for every condition. All infants in fact had a clear preference for the novel leaf in comparison to the familiar leaf, even those in the Synchronous Label condition (paired t-test: $t(28)=5.157$, $p<.001$, two-tailed). In contrast to what the standard analysis of global looking indicates, infants in the Synchronous Label condition as a group did not fail to encode the distributional properties of the "leaf", as they did not fail to perceive the novel leaf as unfamiliar. In fact, the only difference between the looking patterns across conditions is a trend for the "familiar shell" to be looked at more in comparison to the "novel shell" (illustrated in Figure 5 (b)) in the Synchronous condition ($t(28)=1.76$, $p=.09$, paired t-test, 2-tailed) – which does not exist in the other conditions.

Discussion

We familiarised infants with a novel object category, either in silence (Silent condition) or with novel labels. Labels were either presented one second after image onset (Asynchronous Label condition) or simultaneously with image onset (Synchronous Label condition). Global preferential looking results on three subsequent test trials indicated that infants in the Silent and Asynchronous Label conditions were highly successful at learning the target category, exhibiting novelty preference on all test trials. Infants in the Synchronous Label condition, however, did not exhibit novelty preference until the second test trial, pointing to a disruptive role of synchronous labels. Detailed looking patterns indicated, however, that infants in the Synchronous Label condition still preferred the novel leaf within the out-of-category object over the familiar leaf in the consistent test object.

While our present analysis of part-based looking during familiarisation did not reveal an effect of labeling, future analyses will focus on more fine-grained measures of looking (such as the time course of processing within a trial) in order to establish a link between individual infants' gaze patterns during familiarisation and their performance on test.

Preferential looking on the test trials shows that both infants in the Silent and the Asynchronous Label condition learned the category and showed novelty preference on Test 1 and 2, where either the highly variable or the less variable object parts were replaced. This indicates that they successfully encoded the feature distribution of both parts. By contrast, infants in the Synchronous condition did not exhibit novelty preference until Test Trial 2. Using the established measure of object-based novelty preference as a marker of successful category formation this appears to indicate that learning was attenuated in comparison to silence by the presence of the label at the start of the trial. However, analysing infants' looking patterns at the level of individual object parts revealed that infants in the Synchronous condition did appear to be sensitive to the replacement of the leaf part, even though they spent more time inspecting the within-category object compared to the other groups. In addition, on the subsequent test trial, in which the highly variable shell part was replaced by a novel type of shell, even infants in the Synchronous Label condition exhibited significant preference for the novel object. Taken together these data imply that rather than signifying a failure to learn, the looking patterns reflect a delay in the progression to novelty preference, or a lingering familiarity preference (Hunter & Ames, 1988). This finding is consistent with the hypothesis that synchronous labelling increases cognitive load, but it is inconsistent with the hypothesis that visual processing as such is compromised as a result. Infants hearing synchronous labels during familiarisation have learned about the category in question – but they have a tendency to prefer looking at familiar elements.

This finding has implications for the interpretation of studies investigating the impact of labelling on categorisation. First of all, a decrease in novelty preference scores at the object level does not necessarily imply a decrease in visual learning, but can potentially be explained by changes in the speed in which the shift from familiarity to novelty preference is obtained. Null preferences therefore have to be interpreted with caution, specifically when comparing conditions that inherently differ in terms of cognitive load, such as a silent condition vs. one that includes auditory stimuli.

Secondly, this result highlights the role of timing in infants' early learning: even if it is not the categorisation success *per se* that has been compromised due to the synchrony of the label, a lingering familiarity preference still indicates greater difficulty processing the stimuli. This could potentially play a role in reconciling previous contradictory findings regarding the impact of labelling on categorisation. Most studies cited above (e.g. Fulkerson & Waxman, 2007; Ferry et al., 2010; Plunkett et al., 2008; Althaus & Westermann, in prep.) used a delayed label onset similar to our Asynchronous label condition, but Robinson and Sloutsky (2007), who found auditory overshadowing in the presence of labels, presented labels at picture onset. Further research will be necessary in order to determine whether the timing of the label can indeed explain the discrepancies between the findings.

Finally, the fact that infants can deal with asynchronous labels and may even benefit from them, whereas synchronous presentation seems to cause problems with cognitive load, is an important cue to how words and images are processed. Clearly stimuli are processed on-line, rather than stored in short term memory and processed separately and independently of their presentation time. While, in the context of cross-modal processing, synchrony is often claimed to be beneficial at least for young infants (Gogate et al., 2003), the increased load due to synchronous presentation here appears to slow down learning. The actual processing load of any coupling of visual and auditory stimuli is also dependent on the visual and auditory complexity or novelty. However, the fact remains that the likelihood of a label in real life occurring at the same time as an object comes into view is rather small. In terms of learning, this may be an ecological advantage rather than a shortcoming. Perhaps having the opportunity to process visual and auditory information sequentially allows the learning of more complex visual structures. Specifically, the grouping of similar (but not identical) items into categories, for which more abstract visual processing may be needed than just for recognition of individual objects, could be facilitated in this way. Further research is needed in order to shed light on the relationship between exact time course of word-object integration and the ecological circumstances the auditory and visual signals occur in.

Acknowledgments

This research was supported by Wellcome Grant 084386/Z/07/Z to Kim Plunkett and the Winkler Career Development Fellowship at St Hugh's College, University of Oxford, to Nadja Althaus. We wish to thank all parents and infants who participated in this study.

References

Althaus, N., & Westermann, G. (in prep.). *Labels can cause infants to split a visual category.* (In preparation)

Bahrick, L., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental psychology*, *36*(2), 190.

Balaban, M., & Waxman, S. (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*, *60*, 381-398.

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*, *109*, 3253–3258.

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

Bortfeld, H., Morgan, J., Golinkoff, R., & Rathbun, K. (2005). Mommy and me familiar names help launch babies into speech-stream segmentation. *Psychological Science*, *16*(4), 298–304.

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

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

Gogate, L., Bahrick, L., & Watson, J. (2003). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child development*, *71*(4), 878–894.

Grossmann, T., Gliga, T., Johnson, M., & Mareschal, D. (2009). The neural basis of perceptual category learning in human infants. *Journal of Cognitive Neuroscience*, *21*, 2276-2286.

Hunter, M., & Ames, E. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*, *5*, 69-95.

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

Quinn, P. C., Westerlund, A., & Nelson, C. (2006). Neural markers of categorization in 6-month-old infants. *Psychological Science*, *17*, 59-66.

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

Sloutsky, V., Lo, Y., & Fisher, A. (2001). How much does a shared name make things similar? linguistic labels, similarity, and the development of inductive inference. *Child Development*, *72*, 1695–1709.

Tincoff, R., & Jusczyk, P. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*, 172–175.

Waxman, S., & Braun, I. (2005). Consistent (but not variable) names as invitations to form object categories: new evidence from 12-month-old infants. *Cognition*, *95*, 59–68.

Waxman, S., & Markow, D. (1995). Words as invitations to form categories: Evidence from 12-to 13-month-old infants. *Cognitive Psychology*, *29*, 257–302.

Whorf, B. (1956). In J. B. Carroll (Ed.). *Language, thought and reality: Selected writings of Benjamin Lee Whorf*, Cambridge, MA: MIT Press.