The role of color during language-vision interactions

Falk Huettig

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands;

Donders Institute for Brain, Cognition, and Behavior, Radboud University, Nijmegen, The Netherlands

To appear in: R. K. Mishra & N. Srinivasan (Eds.) *Language and Cognition: State of the Art.* Munich: Lincom.

Correspondence should be addressed to:

Falk Huettig

Max Planck Institute for Psycholinguistics

P.O. Box 310

6500 AH Nijmegen

The Netherlands

E-mail: falk.huettig@mpi.nl

phone: +31-24-3521374

1. Introduction

Objects in our visual environment can be identified using many different features (e.g., shape, luminance, color, texture, relative size, stereo depth, and motion; see Treisman & Kanwisher, 1998). Color is one type of information that provides useful cues for object recognition. Color for instance increases the visibility of objects and makes it easier for us to detect them (e.g., Gouras & Eggers, 1983). It has frequently been suggested that fruit and primate trichromacy co-evolved, i.e. one of the evolutionary advantages of color may be that it has enabled us to detect ripe fruits colored in orange, red, or yellow against a background of green leaves (e.g., Caine & Mundy, 2000). Consistent with this claim it has been observed that humans with dichromatic vision have difficulty detecting fruit against dappled backgrounds where lightness is varying randomly (Mollon, 1989).

This example shows that color is one type of information that mediates the capture of attention by salient regions in the visual environment (e.g., detecting a red apple among green leaves). In this case attention is captured with no involvement of memory (i.e. through bottom-up activation). Attention however can also be top-down (i.e., involving long-term memory) to enable goal-directed visual search (see Desimone & Duncan, 1995; Wolfe, 1998). Spoken language is one of the factors that affect the top-down guidance of attention.

The attentional effects of the integration of linguistic information with information from the visual environment have recently sparked considerable interest from cognitive psychologists. Psycholinguists are interested in how and when visual context affects linguistic processing (e.g., Tanenhaus et al., 1995), for instance, to assess claims that language processing is modular and informationally encapsulated (e.g., that syntactic processing is autonomous and cannot be influenced by semantic or visual context; Fodor, 1983). Within the area of visual attention, investigators have become interested in how visual search for a pre-specified target (e.g., a red object among green objects) is mediated by linguistic representations (e.g., Soto & Humphreys, 2007). Thus, attention researchers are interested in the language-vision interface to investigate how search is affected by the interaction of higher order representations (e.g., linguistic representations) with the bottom-up salience of a visual stimulus.

In this chapter I will focus on how color-relations between information accessed from spoken words and information accessed from objects in the visual environment mediate the guidance of overt attention (i.e. eye gaze). First I will briefly review relevant research about influences of color on object recognition, object classification, word recognition, and visual search, and then describe recent experimental results that shed light on how stored and perceived color information mediate language-based guidance of overt attention in the visual world.

2. Background

Shape is a much stronger diagnostic feature for *object recognition* than color (e.g., balls, like most man-made objects, tend to vary in their color but have a distinctive shape). Influential theories of object recognition (e.g., Biederman, 1987; Marr & Nishihara, 1978) thus have focused on the role of edge-based information (i.e. shape) rather than surface information (e.g., color). Biederman (1987) for instance proposed that objects are spatial arrangements of shape-primitives. Every object is assumed to consist of these component parts, so-called geons. During object recognition the arrangements of geons detected is matched against structural models of objects stored in long-term memory. Biederman does not argue that color perception is delayed relative to detection of edge-based information. In fact, changes in the surface characteristics (e.g., color) provide crucial information for edge-based recognition of objects. But although surface information is fundamental for edge detection, it is assumed to play only a secondary role for object recognition (at least when edges can be readily determined). In a series of experiments Biederman and Ju (1988) compared participants' task performance using color photographs and black and white line drawings. Consistent with Biederman's theoretical account, color did not improve performance in object classification tasks. Similarly, Davidoff and Ostergaard (1988) found that adding color to black and white line drawings did not improve performance in a categorization task (i.e. deciding whether objects were living or non-living).

Price and Humphreys (1989) however argued that these findings may be task specific, i.e. masking and brief durations used in verification tasks may favor edge-based information over (surface) color information. Support for this notion comes from many studies which have found that object surface color facilitates performance in naming tasks (Biederman, & Ju, 1988; Humphreys, Goodale, Jakobson, & Servos, 1994; Ostergaard & Davidoff, 1985). Ostergaard and Davidoff (1985) for instance observed that adding color to black and white photographs facilitated naming times. Moreover, Price and Humphreys (1989) found that color influenced performance when participants had to classify birds or fruits, i.e. members of a category that are similar in shape. Thus surface color information appears to aid performance in naming tasks, and in verification tasks when edge-based cues are less informative about object identity.

The object recognition studies mentioned so far contrasted the influence of shape information with the influence of object surface color information. Object surface color (i.e. the perceived yellowness of a displayed banana) however may be distinct from stored object color knowledge (i.e. our knowledge that bananas are typically yellow). Thus it may not only be the surface color of an object that we process when we look at an object; we may also access the stored knowledge we have about the typical color of the particular type of object. Naor-Raz and Tarr (2003) explored whether color is an intrinsic property of object representation, i.e. whether stored color knowledge is routinely associated with an object's representation. Note that the question is not whether we know that bananas are yellow (evidently we do) but whether this knowledge is a component of object representation. Naor-Raz and Tarr (2003) obtained a Stroop effect, i.e. faster color naming times when participants saw a yellow banana than when they saw a purple banana. They interpreted this result as evidence that color is an intrinsic component of visual representations of objects. Most relevant for the present subject matter, Joseph and Proffitt (1996) directly contrasted the influence of stored object color knowledge with the influence of object surface color. They found that stored object color knowledge had a greater effect than object surface color in a verification task in which participants had to decide whether a briefly presented masked picture matched a subsequently presented object name.

In sum, object verification and classification tasks have tended to find that surface color information plays a negligible role during object recognition (at least when edges can be readily determined and category members have distinct shapes). In contrast, object surface color appears to influence the time participants take to name objects. A possible explanation for these contrasting results is that object verification and classification tasks favor edge-based information whereas naming tasks require deeper semantic processing of objects. The one study that directly contrasted surface with stored color knowledge of objects found that stored color knowledge had a greater influence than surface color during a verification task. Thus, the literature on color influences during object recognition does not provide a clear picture; influences of color appear to depend on the properties of the particular task used.

Similarly conflicting are the data on the access of perceptual (e.g., color) information during *word recognition*. These types of studies have investigated whether perceptual knowledge (i.e. knowledge about the typical visual form of an object, e.g., that an aspirin is white) facilitates responses in priming tasks (e.g., that the prime aspirin speeds up responses to white). Schreuder, Flores d'Arcais, and Glazenborg (1985) found that perceptually-related word pairs facilitated naming times. They proposed that perceptual information is rapidly and routinely accessed during word recognition. Pecher, Zeelenberg, and Raaijmakers (1998) however failed to replicate these results. Pecher et al. only obtained significant perceptual priming when they gave participants practice in categorizing primes and targets in a perceptual categorization task prior to the priming task. Moss, McCormick, and Tyler (1997) observed significant perceptual priming using the lexical decision task. Their perceptual priming effect however occurred significantly later than the facilitation found for related functional (e.g., information about what objects are used for) knowledge. Kellenbach, Wijers, and Mulder (2000) in contrast found no significant perceptual priming using the lexical decision task but did observe a robust perceptual priming effect using event-related potentials (ERPs).

When language is used in real-life situations to refer to an object in the listener's visual environment (e.g., to a frog; "that frog is really ugly"), the listener usually has had some pre-exposure to the relevant part of the visual surroundings. In many situations (and in contrast to most priming studies in psycholinguistics, e.g., those which used an auditory prime followed by an orthographic target), target lexical representations are therefore activated by the visual input *before* the critical spoken input is processed (i.e. the lexical concept *frog* is already somewhat active if the listener has looked at the frog before it is referred to). Thus the psycholinguistic studies on the activation of color representations during word recognition (e.g., that on hearing frog the listener may or may not activate that frogs are typically green) in absence of 'visual priming' (i.e. without priming by visual objects) may not directly relate to many real-world language-vision interactions. Similarly, the data from object verification and classification studies are of limited relevance in a situation in which the listener has already recognized the objects in the visual surroundings. In short, though interesting and suggestive, past research in psycholinguistics and visual perception has to be interpreted with caution with regard to how color influences language-based guidance of visual orienting.

More relevant appears to be some recent research in the visual search paradigm (Wolfe, 1998, 2003; Palmer, Verghese, & Pavel, 2000) within the field of visual attention. Soto and Humphreys (2007) for instance asked their participants to read aloud printed words such as "red square". Participants were then required to search for a target in a visual display. Importantly, these targets were unrelated to the items used during the reading task (e.g., unrelated to "red square"). Participants' search for the unrelated targets however was slowed when a red square was co-present in the visual display. Thus the prior printed descriptions of the color and shape of objects interfered with the subsequent visual search. This suggests that information about red squares was retrieved from long-term memory (and most likely kept active in working memory) and affected attentional guidance during the search task. Olivers (submitted) recently had participants search a display for grayscale versions of common traffic signs. On each trial one of the signs was in full color. Critically, the sign presented in color was never the target. Olivers (submitted) found that these color distractors interfered more with the visual search for the target when their color was related to the target (e.g., a distractor sign in red when the target was a grayscale stop sign). Thus, stored color knowledge (about the normal color of the grayscale stop sign) influenced performance even though color was irrelevant to the task.

3. Color-mediated language-based guidance of visual orienting

In a series of studies we explored systematically how color information affects language-based guidance of visual orienting. In our studies we used a variant of the visual world paradigm (Cooper, 1974; Tanenhaus et al., 1995). In our version of the paradigm we present four spatially-distinct pictures of common objects (e.g., a frog, a mitten, a pipe, and a suitcase) on a computer screen (see Huettig & Altmann, 2005; Huettig & McQueen, 2007). These visual displays are divided in four virtual quadrants (upper right, lower right, lower left, and upper left) with one picture presented in each quadrant. Objects are randomly assigned to these quadrants. The spoken sentences typically refer (or a related) to one of the objects in the display. We are particularly interested in what happens during or immediately after the presentation of a critical word (e.g., “frog”) in the sentence. To avoid anticipatory eye movements based on the context of the sentence (see Altmann & Kamide, 1999) each critical word is placed in a neutral sentence context (e.g., “The boy turned around carefully, and then he saw the frog and looked happy”). Participants typically are given a one second preview of the display before the spoken sentences starts to unfold. Moreover since critical words (e.g., “frog”) are not the first word in the sentence participants have a few seconds to familiarize themselves with the objects prior to its onset. We use arrays of objects (rather than more realistic visual scenes) to minimize that scene gist or world knowledge about particular types of scenes affect listeners’ understanding of the spoken sentences. When arrays of objects are used the impact of such knowledge is minimized; therefore arrays are well suited for studying the influence of lexical knowledge associated with particular words and objects.

Participants are not asked to perform an explicit task. They are told that they should listen to the sentences carefully, that they can look at whatever they want to, but that they should not take their eyes off the screen during the experiment (see Huettig & McQueen, 2007, for further discussion of the listening-only task). Participants’ eye movements are measured throughout the experiment.

In our first color study (Huettig & Altmann, 2004) study we presented participants with two experimental conditions. In the target condition the objects in the display (four line drawings) included the target object (e.g., a line drawing of a frog) and three unrelated distractors. The target was colored according to its typical color (e.g., green in the case of the frog); the distractors were in different colors (e.g., a blue mitten, a red suitcase, and a brown pipe). In the competitor condition the objects were identical except that the target object (e.g., the frog) was replaced by a color competitor, i.e. an object (e.g., a lettuce) that is typically associated with the same color as the target object. The spoken sentences ("… frog …") were identical in both conditions, therefore in the target condition the object mentioned in the sentence (e.g., "frog") was present in the display but in the competitor condition there was only a color relation between the critical spoken word (e.g., "frog") and one of the object in the display (e.g., lettuce). Unsurprisingly, we found that on hearing "frog" participants shifted their eye gaze towards the frog. More importantly, however we also observed that in the competitor condition participants directed their eye gaze towards the competitor object (that is on hearing "frog" they looked at the lettuce; or on hearing "lips" they looked at a strawberry; both lips and strawberries have a similar typical color). This study therefore suggested that language-mediated visual attention (i.e. eye gaze) is directed immediately towards objects that match on color even when they mismatch on most other dimensions (i.e. lips and strawberries have little in common other than their similar typical colors).

The color competitor effects in the Huettig and Altmann (2004) study were contingent upon access of stored color knowledge on hearing the *spoken words* (e.g., "frog"), otherwise we could not have observed any language-mediated color effects. However, as the critical *objects* we used were all associated with a prototypical color and also presented in their (surface) color (e.g., a green lettuce), it is unclear whether that color competition effect was contingent upon the stored knowledge about the typical color of the displayed object (i.e. the fact that we know lettuces are green) or whether the effect was contingent purely upon the perceived surface color of the visual object (i.e. our perception of the color green in the display) in absence of any stored object color knowledge. In our next study (Huettig & Altmann, in press) we examined whether color-mediated language-based guidance of visual orienting is driven primarily by the stored color knowledge associated with visual objects or whether it is primarily driven by the perceived surface colors of the objects in the display. Color is an object property that allows such an investigation since conceptual attributes (i.e. the stored color knowledge about visual objects) and perceptual attributes (the perceived but non-diagnostic color of an object) can be dissociated.[1]

3.1. Black and white objects

In Experiment 1 (Huettig and Altmann, in press) we presented black and white line drawings to participants. There were no other colors present in the visual displays. Shifts in eye gaze towards black and white color competitors (e.g., a black and white frog; frogs are typically green) would be due to stored object color knowledge since with black and white line drawings no surface color is depicted. There were three experimental conditions that differed with respect to the critical word that was presented in the spoken sentences. The same visual stimuli were used in all three conditions (a target picture, e.g., a frog, and three unrelated distractor objects). In the target condition the sentence contained the name of the target object (e.g., "The man thought about it for a while and then he looked at the frog and decided to release it back into the wild"). In the competitor condition the sentence was identical up to the critical word (e.g., "spinach"; "The man thought about it for a while and then he looked at the spinach and decided to try out the recipe"); the critical word, i.e. spinach, was not depicted but functioned as a color

[1] The question whether shape competition effects (Dahan & Tanenhaus, 2005; Huettig & Altmann, 2004, 2007) are contingent upon the perceived shape of the objects in the visual environment or our stored knowledge about the shape of a particular object is more difficult to address experimentally since shape is not as arbitrary as color to the function of most objects. It may however be possible to explore this issue by selecting visual objects that can have a typical and an atypical form (e.g., a closed vs. an open umbrella).

competitor to the depicted frog (both frogs and spinach are typically green). We also included a control condition in which critical words (e.g., "radish"; "The man thought about it for a while and then he looked at the radish and decided to try out the recipe") and depicted objects (e.g., frog) were completely unrelated (there was neither a visual referent for radish nor a color match between radish and frog; radishes are typically red). We included the control condition because when generating the materials it became apparent that most objects with a typical color are fruits, vegetables, and animals. The control condition allowed us to asses whether any obtained color competitor effect could be confounded with some other semantic/conceptual similarity between animals, fruit, and vegetables (e.g., radish is neither a color match nor a semantic/conceptual match for frog but belongs to the same category as spinach, i.e. both are vegetables).

As in previous research we found that when participants heard the target word (e.g., "frog") they immediately shifted their eye gaze towards the target object (the frog). In the competitor condition when they heard "frog" there was a tendency to look at the (black and white) lettuce but this tendency was not statistically reliable. In the control condition there was no hint of a shift in eye gaze towards the control object (e.g., the radish) on hearing "frog". To recap, in the competitor condition, the concepts accessed by the spoken words were associated with a typical color. One of the objects in the display was also associated with the same color but all of the displayed objects were presented as black and white line drawings. Any shift towards this competitor object therefore would have been due to stored object color knowledge and not object surface color. However, we did not observe a statistically reliable shift in eye gaze towards competitors.

Why was there no color competition effect with black and white stimuli? One possibility is that participants accessed stored object color knowledge from the black and white objects but did not use this information to direct eye gaze. However, this explanation of the results seems unlikely. Our past research (Huettig & Altmann, 2004, 2005; Huettig et al., 2006; Huettig & McQueen, 2007) has shown that people shift their eye gaze towards the object in the display that matches best the (conceptual and perceptual) specification of the spoken word even if it mismatches on most other dimensions. The color competitors (e.g., the black and white frog) were the only objects in the display with any match (a match in stored color knowledge) between spoken word and the visual display. Such a situation maximizes the likelihood of obtaining an effect (though it does not completely rule out that stored object color knowledge was accessed because eye gaze is a measure of overt attention and not a direct measure of activation of underlying mental representations).

There is however a second possibility. The fact that there was no color (other than black and white) present in the experiment may have induced attentional biases in participants that resulted in the null effect of color relations. There is no visual world research that has explored such potential situational task demands. In the visual search paradigm (Wolfe, 1998, 2003; Palmer, Verghese, & Pavel, 2000) in the field of visual attention however such biases have been explored. Folk and colleagues (Folk, Remington, & Johnston, 1992, 1993) have provided evidence that a particular visual feature may be more or less likely to draw attention depending on the attentional control setting adopted by the participant. Similarly, Műller and colleagues (e.g., Found & Műller, 1996; Műller, Heller, & Ziegler, 1995) have proposed that attentional selection is based on the dimensional properties of objects in the visual field, i.e. top-down information of target defining features is assumed to influence feature-processing stages. There are many differences between visual world and visual search studies (see Huettig, Olivers, & Hartsuiker, submitted, for review), it is however at least conceivable that black and white stimuli reduced attention to color because it was of little task-relevance. In Experiment 2 (Huettig & Altmann, in press) we explored therefore whether stored object color knowledge can ever be used to direct language-mediated eye movements.

3.2. Atypical color

To address this issue, we used the same materials as in the competitor condition of Experiment 1 but presented color and black and white photographs rather than line drawings. There were two conditions, a color condition and a black and white condition. In the color condition, the target objects (e.g., frog) were presented in an appropriate but atypical color (e.g., a yellow frog). Frogs can be all sorts of color (e.g., yellow) but in the Netherlands (our participant population) they are typically green. By using atypically colored objects we were able to use the same objects as in Experiment 1 (e.g., the same frog) but could avoid presenting the (prototypical) surface color to participants (i.e. green). We hypothesized that if people in principle can use stored object color knowledge to direct eye gaze, but in the (black and white) Experiment 1 adopted an attentional setting in which color was not relevant, then with the (atypical) color manipulation of Experiment 2 it should be possible to observe a stored object color effect. On hearing "spinach" (in the sentence "The man thought about it for a while and then he looked at the spinach and decided to try out the recipe") participants should shift overt attention to the (yellow) frog because they have stored color knowledge that frogs in the Netherlands are typically green.

In the black and white condition everything was identical to the color condition except that the photographs were not in color but in black and white. This condition served as an additional control. Black and white photographs contain more visual information than the simple line drawings (e.g., light and dark contrasts) used in Experiment 1. The use of black and white photographs thus served as another test of whether stored object color knowledge is ever used with black and white stimuli. We chose a between-participants design, one set of participants was assigned to the color condition, another set of participants to the black and white condition. This way the experimental design did not create any artificial focus on the presence or absence of color, i.e. color was present on all trials for the participants who were assigned to the color condition.

We found that participants in the color condition shifted their eye gaze more to the atypical color competitors than unrelated distractors. This effect however was marginal and occurred more than one second after the onset of the critical spoken word. This suggests that stored object color knowledge can mediate the language-based guidance of visual orienting but this influence appears to be very small. Participants who were assigned to the condition with the black and white photographs did not shift their eye gaze towards the competitors. This replicated the results of Experiment 1 that with black and white stimuli stored object color knowledge is not used to direct eye gaze.

3.3. Surface color

In Experiment 3 (Huettig & Altmann, in press) we presented pure surface color competitors (e.g., a green blouse) to participants. Blouses, for instance, are not typically green and thus shifts in overt attention to a green blouse on hearing "pea" (peas are typically green) could only be due to the perceived surface color in the display. We observed that participants started to look at the surface color competitors (e.g., the green blouse) as soon as acoustic information from the critical spoken word (e.g., "pea") became available, indicating that hearing "pea" activated prototypical color information which overlapped with the perceived surface color of the visually concurrent blouse. Thus it is not only stored knowledge but also the perceived surface attributes of visual objects which mediate language-based guidance of visual orienting. Notably, the surface color effect was also of much greater magnitude than the stored color knowledge effect of Experiment 2. This suggests therefore that the surface attributes of visual objects are a more important determinant of language-mediated eye movements than the stored knowledge about typical object color.

3.4. The role of color labels: Two -year-olds color-mediated shifts in eye gaze

The only reason why participants shifted their eye gaze to surface color competitors is that they accessed the colour knowledge (e.g., the greenness of a frog) from the *spoken word* because the conceptual make-up of the critical visual stimuli in Experiment 3 (Huettig & Altmann, in press) did not contain stored colour knowledge (e.g., the green blouse). Activation of stored color knowledge from *spoken words* may be direct or indirect. Hearing the name of a concept associated with a typical color (e.g., "frog") may lead to direct activation of stored color knowledge which in turn leads to shifts in eye gaze to green things in the environment. It is however possible that activation is indirect, i.e. mediated by the activation of stored color labels (e.g., the label “green” accessed on hearing the spoken words). When people hear the spoken word “frog”, they may automatically access the word *green*, which may then cause them to shift their eye gaze to anything that is green in the display. This notion is supported by responses in free word association tasks. When participants are asked to write down the first word they think of when reading the word *frog*, they typically give the response *green* (e.g., Nelson, McEvoy, & Schreiber, 1998)[2].

We (Johnson and Huettig, in preparation) have recently investigated this issue with young children who have yet to learn any color terms. We pre-tested two-year-olds on their understanding of color terms (e.g., "point to the red thing"). Two-year-olds were asked to look at one of two smiley faces that were identical in all ways besides color, e.g., the children heard ‘look at the red one’ while viewing a red smiley face and a yellow smiley face. Note that this test of color label knowledge was unlike most past studies of color label knowledge in young children in that it did not require children to actually verbally produce color labels or produce a motor response such as pointing to the correct object. Rather, all our participants had to do was look towards the correctly colored

---

[2] Note however that we do *not* argue that language-mediated eye movements reflect primarily simple associative relationships. Indeed, Huettig and Altmann (2005) found that participants directed overt attention towards a depicted object (such as a trumpet) when a semantically related but not associatively related target word (e.g., "piano") was heard (see also Yee & Sedivy, 2006). Importantly, the probability of fixating the semantic competitor correlated significantly with the semantic similarity between spoken word (e.g., "piano") and competitor object (e.g., trumpet) as derived from semantic feature norms (Cree & McRae, 2003). These data suggested that the increased attention directed to semantically related items (relative to distractor objects) is a function of the degree of semantic overlap. Huettig et al. (2006; see also Yee, Overton, & Thompson-Schill, 2009) provided further evidence for this conclusion. They found that several corpus-based measures of word semantics (Latent Semantic Analysis, Landauer & Dumais, 1997; Contextual Similarity, McDonald, 2000) each correlated well with fixation behavior. These data provide strong evidence that language-mediated eye movements are driven by semantic similarity between spoken word and visual object rather than all-or-none categorical knowledge.

object. Thus, due to the low task demands, this testing procedure provided a very sensitive measure of children's color label knowledge.

We observed that two-year-olds who did not understand color labels showed similar color-mediated eye movements to adults (see also the research within the preferential looking paradigm often used in developmental studies, e.g., Arias-Trejo & Plunkett, 2010; Golinkoff et al., 1987; and Johnson & Huettig, in press, for color-mediated eye movements in three-year-olds). Two-year-olds who were unable to verbally encode the specific relationship between a spoken referent and a seen object nonetheless recognized the perceptual-conceptual commonality between the two to direct eye gaze during online listening. Our results thus suggest that color-mediated language-based guidance of visual orienting is direct, i.e. not mediated by color labels. When two-year-olds who did not understand the color term red were asked to find a strawberry, they were more likely to look at the red plane than the yellow plane even though they could not verbally encode the color property shared by the named and seen objects. These results suggest that the color effect, at least in young children, is not mediated by activation of color labels.

3.5. Surface color vs. conceptual category

When confronted with objects in their visual surroundings and concurrent spoken language, people actively combine information from the utterance and the visual environment to best interpret the situation at hand (see Altmann & Kamide, 2007; Huettig & McQueen, 2007; Huettig, Rommers, & Meyer, under revision; Knoeferle & Crocker, 2006). A situation in which people are exposed to only one competitor (e.g., a surface colour competitor) in the visual surroundings (among unrelated objects) is therefore very different from a situation in which people are exposed to two types of competitors (e.g., a surface colour competitor and a conceptual competitor). In the first case only one type of competitor competes for overt attention but in the second case people's attention may be pulled in different directions. If only one type of competitor is present, this information may be used immediately to direct overt attention because no other information is available for the mapping process. But when two types of competitors are co-present one type of information may be weighed more strongly than the other.

In Huettig and Altmann (in press) we also contrasted the time-course of language-mediated eye movements contingent upon surface color information with language-mediated eye movements contingent upon semantic/conceptual information. We presented surface color competitors (e.g., green blouse) and semantic/conceptual competitors (e.g., mushroom) in the *same visual display* (among two unrelated distractors). The intention was to investigate whether one type of information is prioritized when the system is faced with different types of competitors at the same time. Moss and colleagues (e.g., Moss et al., 1997) for instance have shown that functional/semantic knowledge is a particular salient aspect of lexical knowledge. It is conceivable that the strength of activation of a particular feature determines the probability of attending towards whatever shares those features in the visual environment.

We observed that, on hearing words associated with a typical color such as “pea”, participants shifted their eye gaze towards the semantic/conceptual competitors (e.g., mushroom) about 400 ms earlier than they shifted their eye gaze towards the surface color competitors (e.g., the green blouse). In a condition in which the surface color competitor was the only competitor in the display, however, participants directed their eye gaze immediately towards the surface competitors on hearing the critical spoken word. These findings suggest that semantic/conceptual knowledge is a particular salient aspect of knowledge even for natural objects that are high in color diagnosticity. When semantic/conceptual information is co-present with a competing surface color match, it is semantic/conceptual information that is prioritized during language-based guidance of visual orienting.

4. Open questions and future directions

Color is one of many different factors (e.g., shape, size, motion, etc.) that mediates language-based guidance of visual orienting. The research reviewed above has shown that both object surface color and stored object color knowledge influence language-driven eye movements around the visual world. It appears however that object surface color is a far more important determinant of language-based guidance of visual orienting than stored object color knowledge. One may ask why this is the case.

A reason may be that access of surface color is faster and requires less cognitive resources than access of stored object color knowledge. According to feature integration theory (Treisman & Gelade, 1980) attention is needed for the binding of object features. There appears to be a pre-attentive state in which features such as color are bound to locations (and even objects) but without being bound together (Wolfe & Bennett, 1997). Thus surface color of an object can be processed at a pre-attentive stage but retrieval of stored object color knowledge requires the binding of features and thus attentional processing. It may be that this distinction between pre-attentive processing and attentional focus is the cause of the primacy of object surface color in the present research. Of course in our experiments the visual objects (e.g., the frog) usually are inspected prior to the acoustic unfolding of the critical words (e.g., prior to hearing "spinach"), and thus were previously in the focus of attention. However, Wolfe (1999) has presented evidence that even previously attended objects may revert to a state similar to pre-attentive processing once attention is withdrawn (see also the literature on change blindness, e.g., Simons, 1996).

This raises the question about the role of working memory during the present language-vision interactions. In the field of visual attention it has been suggested that long-term-memory representations of visual objects are bound to locations (so-called object files, Kahneman, Treisman, & Gibbs, 1992). We (Huettig, Olivers, & Hartsuiker, under revision) have recently argued that language-based guidance of visual orienting is mediated by working memory. There is a wealth of evidence that working memory is limited to about four objects that can be tracked, cued, prioritized, and actively remembered (Atkinson, Campbell, & Francis, 1976; Burkell & Pylyshyn, 1997; Cowan, 2001; Luck & Vogel, 1997; Mandler & Shebo, 1982; Pashler, 1988; Phillips, 1974; Pylyshyn & Storm, 1988; Trick & Pylyshyn, 1994; Sperling, 1960; Yantis & Johnson, 1990). The prediction from this research is that increasing the number of objects in the visual display should eliminate the stored object color effect; the object surface color effect however may be little affected by processing load and working memory capacity limits if we accept the argument that only pre-attentive processing of the display is necessary for it to occur. Clearly more empirical research is needed to evaluate this prediction.

One related issue to be resolved concerns the issue of attentional sets (or dimensional weighting). We have suggested that the absence of a stored colour knowledge effect with *black and white* line drawings and photographs may be due to an attentional setting that made retrieval (and attentional use) of colour information less likely. How attentional sets are established in visual world tasks such as the one employed by us is yet unknown. But we have already begun to address this issue. Huettig and McQueen (2007) showed that when pictures in the visual display were replaced with printed words (i.e. the names of the objects) attentional shifts were made only to the phonological competitors (but not semantic or shape competitors) irrespective of preview time. Huettig and McQueen suggested that this was because phonological information is the most relevant for a search among printed words. In other words, the search task with printed-word displays led participants to focus attention on the possibility of phonological matches in the situation where the display consists of orthographic representations of the sound forms of words. In a more recent study (Huettig & McQueen, submitted), we presented displays which consisted of semantic and visual-feature competitors of critical spoken words, and two unrelated distractors (Experiment 1). There were significant shifts in eye gaze towards the semantic but not the visual competitors. Thus participants can use semantic knowledge to direct attention on printed-word displays when phonological matches are impossible. In Experiment 2 semantic competitors were replaced with a further set of unrelated distractors but here was still no hint of preferential fixations towards the visual-feature competitors. In Experiment 3, semantically more loaded sentences were presented to encourage deeper semantic processing and visual imagery but again no shifts in eye gaze to visual competitors occurred. It appears therefore that attentional sets depend on the nature of the information in the visual environment (printed words or visual objects) and the situation in which participants find themselves (see also the research on context-dependent access of semantic/conceptual features, e.g., Barsalou, 1982; Glucksberg & Estes, 2000; Tabossi, 1988).

In another recent study (Huettig & Altmann, in preparation) we further investigated attentional sets by increasing the salience and thus relevance of colour in the experimental setting. We used a within-participants counter-balanced design and alternated colour and black and white trials randomly throughout the experiment. Therefore on one trial our participants heard a word such as '*spinach*' and saw a frog (coloured in green) in the visual display. On the next trial however they saw a banana (in black and white) on hearing '*canary*' (bananas and canaries are typically yellow), and so on. The presence (or absence) of colour was thus a salient characteristic. If the absence of a colour effect with black and white stimuli (Huettig & Altmann, 2004, in press) was simply due to a *general* experimental context which reduced (or even eliminated) an attentional focus to colour features then such a manipulation should greatly increase the likelihood of attentional shifts to the colour competitors in the black and white trials. However, again there was no hint of a color effect in the black and white trials (in contrast to the color trials). This indicates that attentional settings can be established on a trial-by-trial basis. On black and white trials, color is not relevant and thus it is not used; on color trials it is relevant and thus participants use this information to guide language-based visual orienting. More research could also usefully be directed at this issue.

5. Conclusions

Language-mediated eye movements are *not* always conscious or intentional (at least not in the sense that participants always consciously plan in advance which objects they will fixate). In such situations eye gaze is directed to objects that share a color relation with concurrently processed linguistic representations even if these objects mismatch on most other dimensions. Interestingly, such language-based guidance of visual orienting is not mediated by color when black and white stimuli are presented. When color is present, object surface color (rather than stored object color knowledge) is the primary determinant of language-based attentional guidance. These color effects, at least in young children, are not mediated by activation of color labels. Two-year-olds were more likely to look at a red plane than a yellow plane on hearing "strawberry" even though they could not verbally encode the color property (i.e. the property red) shared by the named and seen objects. Though we have made much progress in our understanding of how color mediates language-vision interactions, it is clear that there are still many unknowns about how color influences the integration of linguistic and visual representations. Some of the most interesting questions to be addressed concern the role of pre-attentive processing, the effect of processing load and working memory capacity limits, the influence of task constraints, and the nature of attentional settings.

References

Atkinson, J., Campbell, F. W., & Francis, M. R. (1976). The magic number 4 +/- 0. *Perception, 5*, 327-334.

Altmann, G. T. M., & Kamide, Y. (1999) Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*, 247-264.

Altmann, G. T. M. and Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*, 502-518.

Arias-Trejo, N., & Plunkett, K. (2010). The effects of perceptual similarity and category membership on early word-referent identification. *Journal of Experimental Child Psychology*, *105*, 63-80.

Barsalou, L.W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition*, *10*, 82-93.

Biederman, I. (1987). Recognition-by-Components: A Theory of Human Image Understanding. *Psychological Review, 94,* 115-147.

Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology*, *20*, 38 – 64.

Burkell, J. A., & Pylyshyn, Z. W. (1997). Searching through subsets: A test of the Visual Indexing Hypothesis. *Spatial Vision, 11*, 225-258.

Caine, N. G., & Mundy, N. I. (2000). Demonstration of a foraging advantage for trichromatic marmosets (Callithrix geoffroyi) dependent on food color. *Proceedings of the Royal Society of London B*, *267*, 439-444.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6*, 84-107.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87-185.

Dahan, D. & Tanenhaus, M. K. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychonomic Bulletin & Review*, *12*, 453-459.

Davidoff J. B., Ostergaard, A. L. (1988). The role of colour in categorical judgements. *Quarterly Journal of Experimental Psychology*, *A 40*, 533–544.

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience, 18*, 193-222.

Fodor, J. (1983) *The Modularity of Mind* MIT Press Cambridge, Mass.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 1030–1044.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1993). Contingent attentional capture: A reply to Yantis (1993). J*ournal of Experimental Psychology: Human Perception and Performance*, *19*, 682–685.

Found, A., & Müller, H. J. (1996). Searching for unknown feature targets on more than one dimension: Investigating a "dimensional-weighting" account. *Perception & Psychophysics*, *58*, 88-101.

Glucksberg, S., & Estes, Z. (2000). Feature Accessibility in Conceptual Combination: Effects of Context-induced Relevance. *Psychonomic Bulletin and Review*, *7*, 510 - 515.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*, 23-45.

Gouras, P., & Eggers, H. (1983). Responses of primate retinal ganglion cells to moving spectral contrast. *Vision Research*, *23*, 1175-1182.

Huettig, F., & Altmann, G. T. M. (2004). The online processing of ambiguous and unambiguous words in context: Evidence from head-mounted eye-tracking. In M. Carreiras & C. Clifton (Eds.), *The On-line Study of Sentence Comprehension: Eyetracking, ERP and Beyond* (pp. 187-207). New York, NY: Psychology Press.

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition*, *96*, B23-B32.

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

Huettig, F., & Altmann, G. T. M. (in press). Looking at anything that is green when hearing 'frog' - How object surface color and stored object color knowledge influence language- mediated overt attention. *Quarterly Journal of Experimental Psychology*.

Huettig, F., & McQueen, J.M. (2007). The tug of war between phonological, semantic, and shape information in language-mediated visual search. *Journal of Memory and Language, 54*, 460-482.

Huettig, F., Olivers, C. N. L., & Hartsuiker, R. J. (under revision). Looking, language, and memory: Bridging research from the visual world and visual search paradigms. *Acta Psychologica*.

Huettig, F., Quinlan, P. T., McDonald, S. A., & Altmann, G. T. M. (2006). Models of high-dimensional semantic space predict language-mediated eye movements in the visual world. *Acta Psychologica, 121*, 65-80.

Huettig, F., Rommers, J., & Meyer, A. S. (under revision). The visual world paradigm: What it does and doesn't tell us about linguistic processing. *Acta Psychologica*.

Humphreys, G. K., Goodale, M. A., Jakobson, L. S., & Servos, P. (1994). The role of surface information in object recognition: Studies of a visual form agnosic and normal subjects. *Perception, 23*, 1457-1481.

Johnson, E. K., & Huettig, F. (in press). Eye movements during language-mediated visual search reveal a strong link between overt visual attention and lexical processing in 36-months-olds. *Psychological Research.*

Johnson, E. K., & Huettig, F. (in preparation). Toddlers' verbally mediated shifts in attention: There are no words to describe it.

Joseph, J. E., & Proffitt, D. R. (1996). Semantic versus perceptual influences of colour in object recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 407– 429.

Kellenbach, M. L., Wijers, A. A., & Mulder, G. (2000). Visual semantic features are activated during the processing of concrete words: event-related potential evidence for perceptual semantic priming. *Cognitive Brain Research*, *10*, 67-75.

Knoeferle, P. & Crocker, M.W. (2006). The coordinated interplay of scene, utterance, and world knowledge: evidence from eye tracking. *Cognitive Science*, *30*, 481-529.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato’s problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review, 104*, 211-240.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*, 279-281.

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General, 111*, 1-22.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three dimensional structure. *Proceedings of the Royal Society of London B*, 200:269-294.

McDonald, S. A. (2000). *Environmental determinants of lexical processing effort*. Unpublished doctoral dissertation, University of Edinburgh, Scotland. Retrieved December 10, 2004, from http://www.inf.ed.ac.uk/publications/thesis/online/IP000007.pdf

Mollon, J. D. (1989). "Tho' she kneel'd in that place where they grew..." The uses and origins of primate color vision. *Journal of Experimental Biology*, *146*, 21-38.

Moss, H. E., McCormick, S. F., & Tyler, L. K. (1997). The time course of activation of semantic information during spoken word recognition. *Language and Cognitive Processes*, *12*, 695-731.

Müller, H. J., Heller, D., & Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics*, *57*, 1-17.

Naor-Raz, G., Tarr, M. J., & Kersten, D. (2003). Is colour an intrinsic property of object representation? *Perception*, *32*, 667-630.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/.

Olivers, C. N. L. (submitted). Long-term visual associations affect attentional guidance. *Acta Psychologica*.

Ostergaard, A. L., & Davidoff, J. B. (1985). Some effects of colour on naming and recognition of objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 579 –587.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40*, 1227-1268.

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics, 44*, 369-378.

Pecher, D., Zeelenberg, R., & Raaijmakers, J. G. W. (1998). Does pizza prime coin? Perceptual priming in lexical decision and pronunciation. *Journal of Memory and Language*, *38*, 401-418.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16*, 283-290.

Price C. J., & Humphreys, G. W. (1989). The effects of surface detail on object categorization and naming. *Quarterly Journal of Experimental Psychology, A 41*, 797 –828.

Pylyshyn, Z., & Storm, R. W. (1988). Tracking Multiple Independent Targets: Evidence for a Parallel Tracking Mechanism. *Spatial Vision, 3*, 179-197.

Schreuder, R., Flores D'Arcais, G. B., & Glazenborg, G. (1984). Effects of perceptual and conceptual similarity in semantic priming. *Psychological Research, 45*, 339-354.

Simmons, D. J. (1996). In sight out of mind. *Psychological Science*, *7*, 301-305.

Soto, D., & Humphreys, G. W. (2007). Automatic guidance of visual attention from verbal working memory. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 730-757.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74*(11), whole issue.

Tabossi, P. (1988). Effects of context on the immediate in-terpretation of unambiguous words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 153-162.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632-1634.

Treisman, A. M., & Kanwisher, N. G. (1998). Perceiving visually presented objects: recognition, awareness, and modularity. *Current Opinion in Neurobiology*, *8*, 218-226.

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101*, 80-102.

Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention*. Hove, UK: Psychology Press.

Wolfe, J. M. (1999). Inattentional amnesia. In Coltheart, V. (Ed.) *Fleeting memories* (pp. 71-94). Cambridge, MA: MIT Press.

Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences, 7*(2).

Wolfe, J. M., & Bennett, S. C. (1997). Preattentive object files: shapeless bundles of basic features. *Vision Research*, *37*, 25-44.

Yantis, S., & Johnson, D. N. (1990). Mechanisms of attentional priority. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 812-825.

Yee, E., Overton, E., & Thompson-Schill, S.L. (2009). Looking for meaning: Eye movements are sensitive to overlapping semantic features, not association. *Psychonomic Bulletin and Review,16(5),* 869-874.

Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 1-14.