The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t716100704

Looking at anything that is green when hearing "frog": How object surface colour and stored object colour knowledge influence language-mediated overt attention

Falk Huettig[a b]; Gerry T. M. Altmann[c]
[a] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands [b] Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, The Netherlands [c] University of York, York, UK

First published on: 01 June 2010

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
iFirst, 1–24

Psychology Press
Taylor & Francis Group

Looking at anything that is green when hearing "frog": How object surface colour and stored object colour knowledge influence language-mediated overt attention

Falk Huettig
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, and
Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, The Netherlands

Gerry T. M. Altmann
University of York, York, UK

Three eye-tracking experiments investigated the influence of stored colour knowledge, perceived surface colour, and conceptual category of visual objects on language-mediated overt attention. Participants heard spoken target words whose concepts are associated with a diagnostic colour (e.g., "*spinach*"; spinach is typically green) while their eye movements were monitored to (a) objects associated with a diagnostic colour but presented in black and white (e.g., a black-and-white line drawing of a frog), (b) objects associated with a diagnostic colour but presented in an appropriate but atypical colour (e.g., a colour photograph of a yellow frog), and (c) objects not associated with a diagnostic colour but presented in the diagnostic colour of the target concept (e.g., a green blouse; blouses are not typically green). We observed that colour-mediated shifts in overt attention are primarily due to the perceived surface attributes of the visual objects rather than stored knowledge about the typical colour of the object. In addition our data reveal that conceptual category information is the primary determinant of overt attention if both conceptual category and surface colour competitors are copresent in the visual environment.

Keywords: Attention; Colour; Conceptual category; Eye movements.

How language mediates our attention to objects in the concurrent visual environment has recently attracted great interest. In an early visual-world study, Cooper (1974) showed that participants tended to fixate spontaneously the visual referents of words concurrently heard. For instance, they were more likely to fixate the picture of a snake when hearing "*snake*" or part of "*snake*" than

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

F.H. was supported by a University of York doctoral studentship and a grant from FWO, the Fund for Research Flanders (G.0181.05). The second experiment was conducted while the first author was at the Max Planck Institute for Psycholinguistics, Nijmegen. The eye-tracking facilities and software/analysis routines were supported by the Medical Research Council Grants G9628472N and G0000224 awarded to G.A. A partial report of Experiment 1 appeared in Huettig and Altmann (2004). We thank Philip Quinlan, Gareth Gaskell, and Graham Hitch for their useful contributions to this research and Willemijn van den Berg and Holger Mitterer for assistance.

http://www.psypress.com/qjep
DOI:10.1080/17470218.2010.481474

pictures of referents of unrelated control words (see also Allopenna, Magnuson, & Tanenhaus, 1998). We (Huettig & Altmann, 2005) recently investigated whether semantic properties of individual lexical items can direct eye movements towards objects in the visual field. Participants who were presented with a visual display containing four pictures of common objects directed overt attention immediately towards a picture of an object such as trumpet when a semantically related but nonassociated target word (e.g., "*piano*") was heard. Different measures of semantic relatedness (semantic feature norms, Cree & McRae, 2003; latent semantic analysis, Landauer & Dumais, 1997; McDonald's contextual similarity measure, McDonald, 2000) each separately correlated with fixation behaviour (Huettig & Altmann, 2005; Huettig, Quinlan, McDonald, & Altmann, 2006). These data show that semantic similarity can drive language-mediated eye movements to objects in the concurrent visual environment.

Given the observed effects of *conceptual* overlap, it is interesting to explore how *perceptual* overlap interacts with language-mediated eye movements. Recent studies have shown that looks can be directed to visually related (i.e., by shape) but semantically inappropriate objects (Dahan & Tanenhaus, 2005; Huettig & Altmann, 2004, 2007). For example, we found that participants shift overt attention to a picture of a cable during the acoustic unfolding of the word "*snake*" (cable and snake have a similar global shape). Similarly, participants shift overt attention to a picture of a strawberry during the acoustic unfolding of the word "*lips*" (strawberries and lips have a similar typical colour; Huettig & Altmann, 2004).

A difficulty with researching perceptual synergies, however, is that eye movements that are contingent upon perceptual information cannot easily be dissociated from eye movements that are contingent upon stored conceptual information. The shape of an object—the long and thin form of a snake for example—can be *perceived* (perceptual information) but is also *known* (conceptual information). Therefore, it is unclear whether the shape-driven shifts in overt attention reported previously both by ourselves and by Dahan and Tanenhaus (2005) were driven by the stored (conceptual) knowledge of the shape of the displayed objects or by the perceived shape of the objects in the visual display. Similarly, it is unclear whether the colour effect was contingent upon the *stored* knowledge about the typical colour of the displayed object or whether language-mediated attention can also be contingent upon the *perceived* surface colour of the visual object in absence of any stored colour knowledge.

Colour is an object property that allows the investigation of this question directly since conceptual attributes (the stored colour knowledge about an object) and perceptual attributes (the perceived but nondiagnostic colour of an object, i.e., its surface colour) can be dissociated. It is a part of our general knowledge about frogs that they are typically green, and thus part of the make-up of the concept frog should be their prototypical colour—namely, green. If we hear the spoken word "*frog*" then part of the information that is accessed should be that prototypical colour information. 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). Similar, if we see a visual object, then part of the conceptual information that we retrieve should be the prototypical colour of the object (if the object is associated with a prototypical colour). Naor-Raz, Tarr, and Kersten (2003), for instance, investigated the role of colour in object representation using a task in which participants had to name the displayed colours of objects and words. They found that participants were faster to name the colours of objects if they were presented with the typical colour (a yellow banana) than with an atypical colour (a purple banana). Naor-Raz et al. concluded that colour is an *intrinsic* component of the mental representations of objects. When we see a visual object, however, it is not only conceptual information that we access, we also perceive the surface properties of the object such as its colour, which may (e.g., a line drawing of a frog coloured in green) or may not (e.g., a line drawing of a frog coloured in blue) be prototypical. The current research investigates whether language-mediated eye movements are driven primarily by the stored colour knowledge associated with the visual objects or whether they are primarily driven by the perceived surface colours of the visual objects.

The influence of colour on object recognition and classification

Although the role of colour during language–vision interactions has received little attention, a number of object identification studies have investigated the effect of colour information with respect to object recognition and object classification. Research into object recognition has typically found that surface colour facilitates object naming but does not facilitate performance in object classification or object verification tasks. For instance, Ostergaard and Davidoff (1985) found that presenting participants with colour photographs instead of black-and-white photographs facilitated object naming. Similarly, Williams and Tanaka (2000; as cited in Naor-Raz et al., 2003) observed that participants were faster to name an object with a typical colour such as a banana when presented with colour pictures than when presented with line drawings.

In contrast, Davidoff and Ostergaard (1988) found no facilitation of colour line drawings compared to black-and-white line drawings when participants had to classify the objects as living or nonliving. Price and Humphreys (1989), however, found that colour had an effect on classification performance among category members that were similar in shape (birds). They concluded that surface colour can be useful in tasks other than naming. On the other hand, Biederman and Ju (1988) found that colour photographs, compared to black-and-white line drawings, did not improve performance in a verification task in which participants had to decide whether a word presented after a picture referred to the same object. Price and Humphreys (1989), however, suggested that these findings may be task-specific; they argued that the masking and brief durations used in verification tasks may favour edge-based information over surface colour information.

Note that the studies summarized above did not compare the effect of surface colour versus the effect of stored colour knowledge. In this regard, Joseph and Proffitt (1996) conducted a verification task in which participants had to decide whether a briefly presented (and masked) picture matched the subsequently presented object name. They observed that when response interference could not be due to shape information, stored colour knowledge had a much greater effect than surface colour. For items that were similar in shape, however, they found some influence of surface colour on performance. Joseph and Proffitt concluded that, overall, stored colour knowledge is more influential than surface colour in object recognition.

In sum, the effect of stored colour knowledge and perceived surface colour in object recognition and object classification tasks seems to be largely task dependent. It is therefore difficult to interpret the colour effect in the visual-world paradigm (Huettig & Altmann, 2004) with regard to this literature. In our experiments participants are presented on a given trial with a visual display containing four pictures of common objects. The displays are divided into four (virtual) quadrants—upper right, lower right, lower left, and upper left—and one picture is presented in each quadrant. As participants view such displays their eye movements are monitored. Importantly, during the course of a trial a spoken sentence is presented to the participant, and the participant's eye movements are tracked as the sentence unfolds. Participants are told to listen to the sentences carefully. They are also told that they can look at whatever they want to, but are told not to take their eyes off the screen throughout the experiment. Participants are given no other task instructions (cf. Huettig & Altmann, 2005; Huettig & McQueen, 2007). Note that the observed competitor effects (e.g., that participants shift overt attention to a picture of a cable during the acoustic unfolding of the word "s*nake*") are unlikely to be limited to the visual-world task employed here. Dahan and Tanenhaus (2005) demonstrated similar visual-form competitor effects when participants were required to engage in an explicit physical task (moving the objects mentioned in spoken sentences using a computer mouse). Similarly, Yee and Sedivy (2006), using a task in which participants had to touch one of the displayed objects on a computer screen, observed similar semantic effects to those obtained by Huettig and Altmann (2005). Note also that these competitor effects occur when the entity mentioned in the spoken sentence is not present in the visual display as well as when it is. Competitor effects have also been observed in a language production task (Huettig & Hartsuiker, 2008). Moreover, Huettig and Altmann (2005) compared target-present and target-absent conditions and have found, other than the tendency for fixations to targets to dominate when targets are present, similar results across these conditions. Thus, these competitor effects are not limited to certain specific goal-directed task demands (see Huettig & McQueen, 2007, for further discussion). In other words these competitor effects are robust and occur when participants hear spoken language and view simultaneously a visual scene or display of objects. Participants map the unfolding spoken words onto the concurrent visual objects. The literature on perceptual priming during lexical processing is thus also relevant here.

Perceptual priming and word recognition

The question of whether perceptual information (such as an object's colour) becomes automatically available on hearing a spoken word such as "*aspirin*" has attracted some attention in language comprehension research although the data have at times been contradictory. Schreuder, Flores d'Arcais, and Glazenborg (1984; see also Flores d'Arcais, Schreuder, & Glazenborg, 1985) obtained significantly facilitated target-naming times for perceptually related word pairs and proposed a model of lexical activation in which perceptual representations are activated very rapidly during spoken-word recognition. Pecher, Zeelenberg, and Raaijmakers (1998), however, found perceptual priming only if participants were first given practice in categorizing primes and targets in a perceptual categorization task. Moss, McCormick, and Tyler (1997), using the lexical decision task, found a significant priming effect for functional properties of words early during the duration of the word but priming for perceptual targets (e.g., the colour overlap between *aspirin* and *white*) only at the offset of the prime word. Kellenbach, Wijers, and Mulder (2000) obtained robust perceptual priming as indexed by the event-related potential (ERP) N400 component but contrary to Moss et al. (1997) observed no effect for the same materials from the ERP study when used in a lexical decision task. Similarly, we (Huettig & Altmann, 2004, 2007) observed that our participants started to shift overt attention to a cable well before the offset of the spoken word "snake", suggesting that shape information can be accessed before word offset (in contrast to Moss et al., 1997, who did not find priming of perceptually related word pairs before the offset of the prime word).

The data from the perceptual-priming literature thus have at times been contradictory. Note, however, that as for the literature in the domain of visual-object processing these findings cannot be translated directly to the visual-world paradigm due to differing task constraints. Critically, in cross-modal priming, the auditory prime precedes the visual (orthographic) target. Thus, any visual properties associated with the conceptual representation activated by the spoken word could in principle become activated before the onset of the visual (orthographic) target. But looks to a visually or semantically related object in the visual display are not the equivalent of facilitatory priming of the visual target in cross-modal priming. This is because in the visual-world paradigm, the visual target precedes the auditory signal—visual properties associated with the conceptual representations ordinarily activated by the spoken word are not activated before the onset of the visual objects; those visual properties have already been activated, precisely because those visual objects precede the spoken word. The implication of this for time-course analyses in the visual-world paradigm is that these do not necessarily indicate the time-course of activation of visual-feature information associated with the unfolding word. Rather, they indicate the time-course of the process by which this information modulates the activation of the episodic (or "working memory") representations themselves activated, earlier, by the visual display.

Current study

Some care, therefore, must be taken when interpreting the present eye-tracking research on language–vision interactions with regard to results obtained using other methods (such as the lexical decision task or object verification tasks). Though the research presented here may be of relevance to an understanding of both visual-object processing and lexical processing, its primary importance is in increasing our knowledge about the interaction between linguistic, attentional, and visual processing. The main goal is to increase our knowledge of how language processing interacts with the properties of the visual environment in which language processing often takes place. Traditionally, language processing has almost exclusively investigated language in isolation from the environment in which it is usually used. Although one of the defining features of language is that it can be used to refer to objects in their absence, language is often used to refer to objects in the immediate environment. From an evolutionary perspective, the sensory environment is likely to have had an important impact on the evolution of language and, more generally, cognition. Similarly, from a developmental perspective, language acquisition appears to be reliant in a large part on a concurrent world with which the child can interact. Notwithstanding this reliance on an external world (and even in its absence, language does refer to objects and events in that external world), there has been little research that directly explores the interface of language and visual perception and its impact on our interactions with the visual world.

The specific aim of the present research was to investigate whether effects of visual overlap during language-mediated overt visual attention are contingent primarily upon purely perceptual information such as the perceived surface colour of the visual object or whether they are primarily contingent upon stored conceptual information (e.g., the stored knowledge about the typical colour of the displayed object). In Experiment 1 we presented black-and-white line drawings to investigate whether language-mediated eye movements can be driven by stored colour knowledge about concurrent objects in the visual field.

EXPERIMENT 1

Method

Participants

A total of 60 participants from the University of York student community took part in this study. All were native speakers of British English and either had uncorrected vision or wore soft contact lenses or glasses.

Materials

Three experimental conditions were generated: a target condition, a competitor condition, and a control condition. The same visual stimuli were used in all three conditions. Each display contained four line drawings (see Figure 1). The individual line drawings were selected from Snodgrass and Vanderwart (1980) and were presented in black and white. No other colours were physically present in the displays.

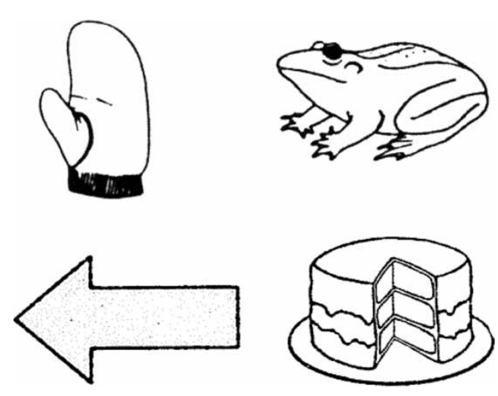

Figure 1. *Example of a visual stimulus in Experiment 1 (depicting: target frog and three unrelated distractors).*

The three conditions differed with respect to the critical word that was presented in spoken sentences. In the target condition the sentence contained the name of one of the pictures in the display. With respect to Figure 1 the relevant sentence in the target condition would be "*The man thought about it for a while and then he looked at the frog and decided to release it back into the wild*" (here "*frog*" is the target word). In the competitor condition the relevant sentence would be "*The man thought about it for a while and then he looked at the spinach and decided to try out the recipe*" (here "*spinach*" is a colour competitor to the depicted frog). Finally in the control condition the speech and the pictures were completely unrelated. Here a relevant sentence would be "*The man thought about it for a while and then he looked at the radish and decided to try out the recipe*"; the corresponding picture contained neither a visual referent for "*radish*" nor a colour match between frog and radish.

The control condition was included because, in generating the materials for this experiment, it became apparent that most of the objects associated with a prototypical colour are animals or fruits/vegetables. The control condition was therefore included to ensure that any possible competitor effect could not be confounded with some other conceptual similarity between animals and fruits/vegetables (e.g., "radish" is neither a colour nor a semantic match for frog but belongs to the same conceptual category as "spinach", i.e., both are vegetables).

In order to ensure that the referents of the critical words were associated with a prototypical colour, only words that elicit a colour according to the Nelson word association norms (Nelson et al., 1998) were chosen. In these norms the forward strength (FSG or cue-to-target strength) represents the proportion of participants who produced the target in the presence of the cue word. For example, the FSG for "grasshopper" and "green" is .240: In other words, 38 of the 154 participants produced "green" in response to "grasshopper". Table 1 shows the forward strength values for the critical words.

As shown in Table 2 the critical words were also matched for CELEX log word frequency, $F(2, 28) = 2.46$, $p = .103$, number of syllables, $F(2, 28) = 1.24$, $p = .306$, and number of letters, $F(2, 28) = 1.41$, $p = .261$.

Design

A total of 15 experimental trials and 30 filler trials were generated. Across the 15 experimental trials,

Table 1. *Forward strength values of target, competitor, and category control words in Experiment 1*

Target			*Colour competitor word*			*Category control word*		
Word	*Colour*	*FSG*	*Word*	*Colour*	*FSG*	*Word*	*Colour*	*FSG*
grasshopper	green	.240	broccoli	green	.250	carrot	orange	.180
frog	green	.070	spinach	green	.250	radish	red	.200
lettuce	green	.060	emerald	green	.370	sapphire	blue	.160
celery	green	.060	turtle	green	.020	salmon	pink	.030
tree	green	.040	lizard	green	.100	duck	yellow	.020
leaf	green	.040	crocodile	green	.010	shrimp	pink	.010
cherry	red	.210	lipstick	red	.240	jeans	blue	.250
tomato	red	.190	brick	red	.110	hose	green	.050
strawberry	red	.100	fox	red	.40	bear	brown	.020
lobster	red	.080	raspberry	red	.090	kiwi	green	.010
lips	red	.060	rose	red	.170	daisy	yellow	.040
banana	yellow	.140	canary	yellow	.110	robin	red	.050
lemon	yellow	.040	dandelion	yellow	.40	orchid	purple	.050
corn	yellow	.060	chick	yellow	.040	parrot	green	.020
pumpkin	orange	.160	tiger	orange	.020	beaver	brown	.010

Note: FSG: forward strength or cue-to-target strength (cf. Nelson, McEvoy, and Schreiber, 1998).

Table 2. *Properties of the critical words in Experiment 1*

	Target word	*Colour competitor word*	*Control word*
Celex log frequency	0.83 (0.54)	0.55 (0.56)	0.52 (0.45)
Mean length in syllables	2.0 (0.85)	2.13 (0.83)	1.73 (0.59)
Mean length in letters	6.07 (2.15)	6.47 (1.88)	5.40 (1.12)

Note: Standard deviations of the means in parentheses.

5 were from the target condition, 5 were from the competitor condition, and 5 were from the control condition. Subsets of 5 items taken from a critical set of 15 were counterbalanced across three counterbalancing groups but the same 30 filler items were used for all three groups. The same sets of visual displays were used for all three groups, but the assignment of sentences to these pictures varied across the groups. Filler items contained the name of a pictured object, and as a consequence 78% of the 45 trials included a named picture; hence across trials participants may well have built up an expectation that one of the pictures would be named.

Procedure

Participants were seated at a comfortable distance (with their eyes about 50 cm from the display) in front of a 17-inch display and wore an SMI EyeLink head-mounted eye-tracker, sampling at 250 Hz from the right eye (viewing was binocular). Participants were not asked to perform any explicit task. They were told that they should listen to the sentences carefully and that they could look at whatever they wanted to, but not to take their eyes off the screen throughout the experiment (cf. Huettig & Altmann, 2005; Huettig & McQueen, 2007).

The position of the four objects was randomized over trials but fixed across participants. The onset of the visual display occurred 1 s before the onset of the speech; the onset of the critical word was on average 4 s after the onset of the speech. Between adjacent trials participants were shown a single dot located in the centre of the screen, which they were asked to fixate prior to a fixation cross appearing in this position (this procedure allowed the eye-tracker to correct for drift). The termination of trials was preset and controlled by the experimental program, and thus participants could not terminate trials by themselves. The trial was automatically terminated after 9 seconds, which, typically, left 2 seconds after the end of the sentence. After every fourth trial, the eye-tracker was recalibrated using a 9-point fixation stimulus. The entire experiment lasted approximately 20 minutes.

Results

The change as the critical word acoustically unfolded in time in the proportion of trials on which the target and unrelated distractor were fixated is shown in Figure 2. The curves are synchronized to the acoustic onset of the critical word, and the *x*-axis shows time in milliseconds from this onset. The calculation excluded all movements prior to the acoustic onset, and thus negative values reflect moves away from objects that were already fixated at this onset; in effect, each data point reflects the proportion of trials with a fixation at that moment in time minus the proportion of trials with a fixation at the acoustic onset of the target word (cf. Huettig & Altmann, 2005). The time-course graph shows that in the target condition as soon as information from the acoustic target word (e.g., "*frog*") became available participants shifted their eye gaze towards the target picture (the frog). Figure 2 suggests that there may have been a small shift in overt attention in the competitor condition. Fixations in the category control condition did not diverge.

Statistical inferences

For the statistical analyses we computed mean fixation proportions for each type of object. We calculated the ratio between the proportion of fixations to the target picture and the sum of the target- and distractor-fixation proportions (cf. Huettig & McQueen, 2007). A ratio greater than .5 shows that, of all the fixations directed towards the target and the unrelated distractors, the targets attracted more than half of those fixations. We computed mean ratios per participant and item for the first 100 ms starting from the acoustic

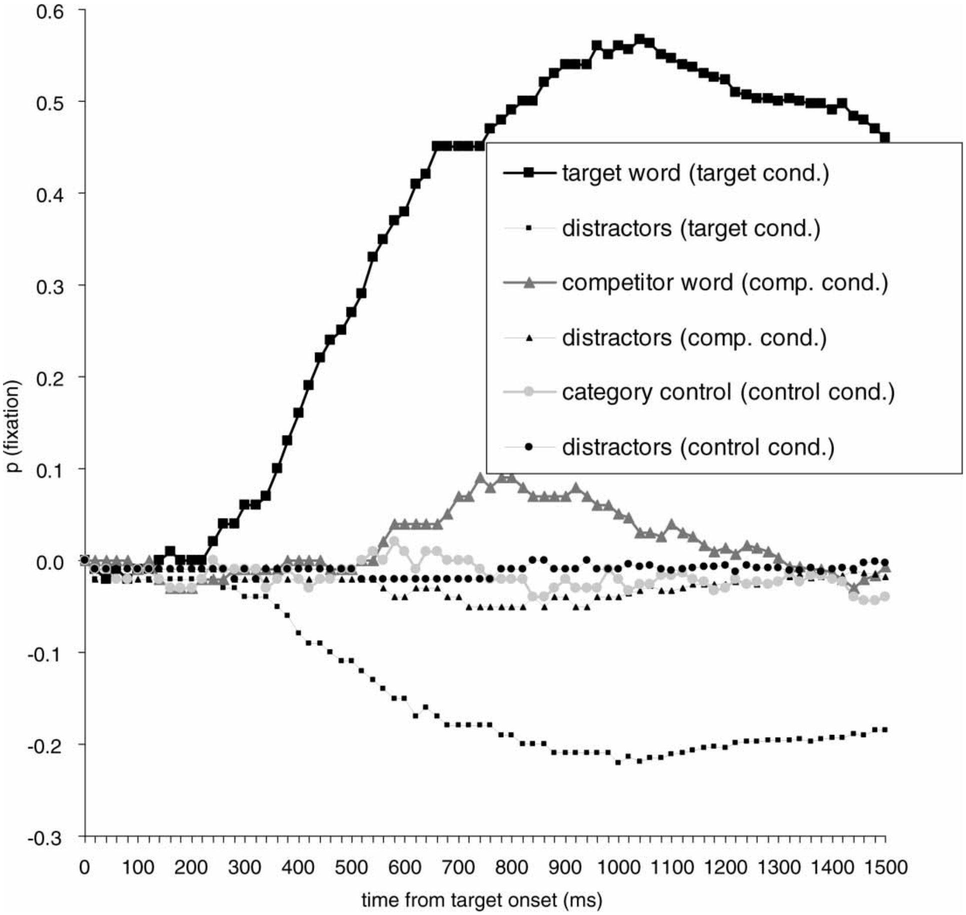

Figure 2. *Time-course graph plotting the change in fixation proportions directed to the target pictures (e.g., frog) and the unrelated distractors in Experiment 1. The curves are synchronized to the acoustic onset of the critical word, and the* x-*axis shows time in milliseconds from this onset.*

onset of the critical spoken word in order to obtain a baseline of target–distractor fixation ratios. Estimates about the time to programme and initiate a saccadic eye movement vary between 100 ms and 180 ms (see Altmann & Kamide, 2004, for review). We can therefore be sure that fixations during the baseline time region were not influenced by information from the critical word because of the time considered necessary for programming and initiating an eye movement. We calculated mean ratios during the baseline region to adjust for any bias in overt attention to a type of object before information from the critical word became available. Calculating a mean ratio for the baseline time regions (and then comparing these ratios with the mean target–distractor ratios during later time regions) allows us to test for any shifts in overt attention to particular types of object during the time of interest. We tested whether the target–distractor ratios during the 100-ms baseline time window were significantly different from the target–distractor ratios during the subsequent 100-ms time windows.

Paired t tests showed that in the target condition the mean target–distractor ratios during the baseline region (.49) first differed significantly from the mean ratios during the 300-ms to 399-ms time region (.59), $t_1(59) = 3.48$, $p = .001$, $t_2(14)$

$= 1.88$, $p = .08$. The shift in overt attention towards the target pictures was statistically significant in all subsequent time regions. Thus on hearing the spoken target word (e.g., "*frog*") in the target condition, participants shifted their attention towards the target picture.

Figure 2 suggests that there was a small shift in overt attention to the target pictures in the competitor condition between 600 ms and 1,099 ms after the acoustic onset of the competitor word. We split this 600-ms to 1,099-ms time window into five 100-ms time windows and tested, for each time window, whether the target–distractor ratio in that window was significantly different from the baseline. The mean target–distractor ratios during the baseline region did not differ significantly from the mean ratios during the 600-ms to 699-ms time region and the 1,000-ms to 1,099-ms time window. The difference between the mean target–distractor ratios during the baseline region (.50) and the 700-ms to 799-ms time region (.59), $t_1(59) = 1.91$, $p = .061$, $t_2(14) = 1.74$, $p = .11$, the 800-ms to 899-ms time region (.58), $t_1(59) = 1.75$, $p = .09$, $t_2(14) = 1.56$, $p = .14$, and the 900-ms to 999-ms time region (.57), $t_1(59) = 2.06$, $p = .044$, $t_2(14) = 1.34$, $p = .20$, approached significance by participants (but not by items).[1] In the category control condition the target–distractor ratios during the baseline region did not diverge significantly from the target–distractor ratios during the subsequent time windows. Thus, on hearing the competitor word (e.g., "s*pinach*") in the competitor condition there was a small shift in overt attention to the target picture (e.g., the frog) but this trend was not statistically robust. There was no change in overt attention in the category control condition.

Discussion

On the competitor trials, the concepts accessed by the spoken words were associated with a prototypical colour. In these competitor trials one of the pictured objects was also associated with the same colour. Importantly, none of the pictures was presented in colour. In this regard any contingent effect would have been due to stored colour knowledge and not surface colour. However, we did not observe any statistically reliable effect of stored colour knowledge.

Though the absence of a reliable effect of stored colour knowledge in Experiment 1 is striking, it does not rule out that stored colour knowledge influences language-mediated eye movements. It is conceivable that the use of black-and-white line drawings induced an attentional bias that resulted in the null effect of colour properties. Indeed it has been argued in the visual search literature that effects of a particular visual feature may be more or less likely to draw attention depending on the "attentional control setting" adopted by the participant (cf. Folk, Remington, & Johnston, 1992, 1993; see also Pratt & Hommel, 2003). Similarly, according to dimension-based theories of selective attention (e.g., Allport, 1971; Müller, Heller, & Ziegler, 1995) selection is based on the dimensional properties of objects in the visual field. Top-down information of target-defining features is assumed to influence feature-processing stages (Found & Müller, 1996; Treisman, 1988; see also Kumada, 2001). Thus it is possible that in the present task the black-and-white pictures may have reduced attention to the colour dimension because it was of little task relevance in the particular context. The absence of a stored colour knowledge effect thus may simply reflect that the black-and-white line drawings induced an attentional setting that made retrieval (and attentional use) of colour information less likely. Indeed the slight trend in shifts in eye gaze in the competitor condition (Figure 2) could be viewed as support for this interpretation of the results. Experiment 2 was conducted to investigate whether stored colour knowledge associated with

[1]A difference score (DS) between the forward strength (FSG) values of target and colour competitor was computed (mean DS $= -.022$, $SD = .11$) to assess whether a large difference in FSG values between target and competitor would affect the result. Only one item pair (lettuce/emerald) had a difference score greater then 2.5 standard deviations. The statistical analyses were repeated with this outlier removed but showed a similar pattern to the results reported.

visual objects can ever be used during language-mediated eye movements.

EXPERIMENT 2

We presented the experimental sentences and visual stimuli (in colour and in black and white) of the competitor condition of Experiment 1 to a new set of participants. However, rather than presenting line drawings we presented photographs of the objects in Experiment 2 in a colour photographs condition and in a black-and-white photographs condition. Critically, in the colour condition, the target objects (e.g., frog) were presented in an appropriate but atypical colour (e.g., a yellow frog, see Figure 3b). South American frogs, for instance, can be any colour, ranging from yellow through red to blue; around the Netherlands, however (our participant population), frogs are typically green. We used atypically coloured objects in order to be able to use the same competitor items as those in Experiment 1 (e.g., frog) but presented in colour (e.g., a yellow frog) and thereby also avoided presenting the prototypical surface colour (e.g., green) of the competitor objects. If the absence of a robust stored colour effect in the previous experiment was due to the absence of colour in the black-and-white trials but participants routinely use stored colour knowledge in colour trials then we should be able to observe a stored colour effect with the present manipulation. Huettig and Altmann (2004) showed that participants access colour knowledge when hearing spoken words that are high in colour diagnosticity (e.g., "*spinach*"). If language-mediated eye movements can be influenced by stored object colour knowledge then on hearing "*spinach*" participants should shift overt attention to the yellow frog because they have stored colour knowledge that frogs in the Netherlands are typically green. To increase the ecological validity of the present experiment we chose a between-participants design. One set of participants were presented with colour photographs only. A second set of participants were presented with black-and-white versions of the same photographs. Thus for the participants who received the colour version of the experiment colour was present on all trials but the experimental design did not create an artificial focus on the presence or absence of colour. The use of black-and-white photographs for the second set of participants served as an additional control. Black-and-white photographs contain more useful visual information, such as light and dark contrast, than do the line drawings used in the previous experiment. The use of black-and-white photographs therefore represents another test of whether language comprehenders can ever use stored colour knowledge of visual objects with black-and-white stimuli for language-mediated overt attention.

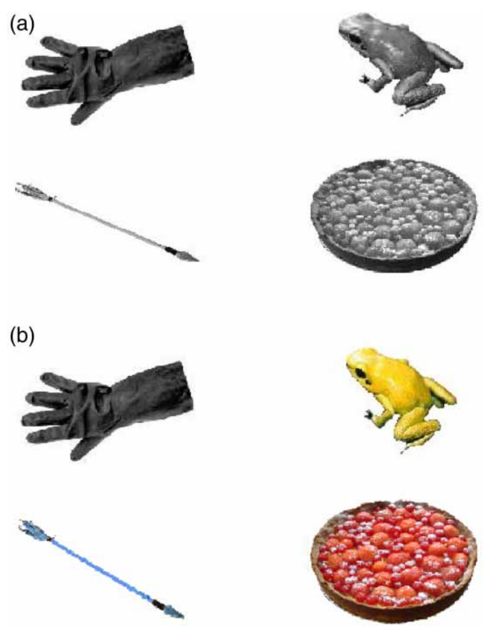

Figure 3. *Example of a visual stimulus in (a) the black-and-white condition (depicting: target frog and three unrelated distractors) and (b) the colour condition of Experiment 2. To view a colour version of this figure, please see the online issue of the Journal.*

Method

Participants

A total of 70 members of the participant panel of the Max Planck Institute for Psycholinguistics were paid for participation. All were native speakers of Dutch and either had uncorrected vision or wore soft contact lenses or glasses.

Materials

A total of 14 of the 15 experimental items of Experiment 1 were used but instead of line-drawings we used photographs of real objects. There were two lists. In List A the visual displays consisted of black-and-white photographs (Figure 3a); in List B the colour versions of the same photographs were used (Figure 3b). Photographs of the same 30 filler displays as those in Experiment 1 were also presented. For List A all of the filler photographs were in black and white; for List B all of the fillers were colour photographs,

The English spoken sentences of the competitor condition of Experiment 1 (e.g., "*The man thought about it for a while and then he looked at the spinach and decided to try out the recipe*") were translated into Dutch (e.g., "*De man dacht er even over na, waarna hij naar de spinazie keek en besloot het recept uit te proberen*") by a Dutch native speaker. Dutch and English are closely related languages, and thus (as in the English sentences) the critical word occurred on average 3 seconds after the onset of the spoken sentences (i.e., 4 seconds after display onset). The 14 experimental trials and 30 filler trials were read aloud by a female native speaker of Dutch in a sound-damped booth. The digital recordings of these utterances (at a sample rate of 44.1 kHz with 16-bit resolution) were stored directly on a computer.

Design

A between-participants design was used. This ensured that every participant saw all 14 experimental items. The spoken sentences were identical across lists. Participants in List A saw only black-and-white photographs. Participants in List B saw the colour versions of the same photographs. Within each list the trials were presented in fixed random order.

Procedure

The procedure was the same as that in the previous experiment.

Results

Figure 4 shows the time-course graph. The graph shows that for the group of participants who received black-and-white photographs fixations to the different types of picture never diverged. The figure also suggests that there was a small shift in overt attention towards the target pictures for the group of participants who received colour photographs. The same statistical analyses as those in the previous experiment were conducted.

In the group of participants who received colour photographs, there were no reliable differences between the mean target–distractor ratios during the baseline region and the mean target–distractor ratios until 1,299 ms after the acoustic onset of the critical word. Figure 4 suggests that at about 1,300 ms after critical word onset there was a small shift in overt attention towards the target object. The small difference between the mean target–distractor ratios during the baseline region (.53) and the mean target–distractor ratios during the 1,300-ms to 1,399-ms time region (.56), $t_1(34) = 0.76$, $p = .454$, $t_2(13) = 1.15$, $p = .27$, however, was not statistically reliable. The difference became more robust during the 1,400-ms to 1,499-ms time region (.59), $t_1(34) = 1.58$, $p = .123$, $t_2(13) = 2.20$, $p = .047$, and the 1,500-ms to 1,599-ms time region (.59), $t_1(34) = 1.62$, $p = .116$, $t_2(13) = 2.36$, $p = .035$.

In the group of participants who received black-and-white photographs mean target–distractor ratios never diverged significantly.

Discussion

In Experiment 2 we observed no hint of an influence of stored object colour knowledge when black-and-white stimuli were presented. These data therefore replicate the results of Experiment 1 with black-and-white photographs containing more useful visual information than simple line drawings. Experiment 2 also revealed a weak effect of stored object colour knowledge during language-mediated overt attention when colour photographs were used. This suggests that participants retrieved prototypical colour information

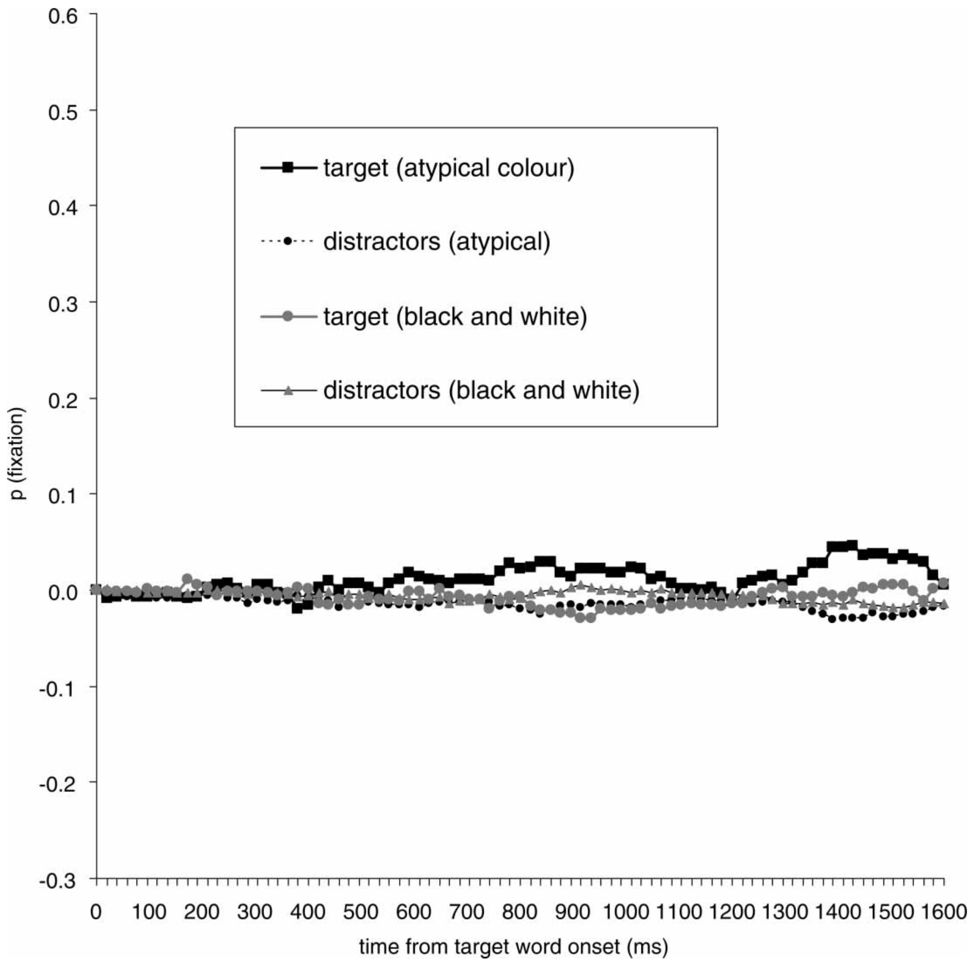

Figure 4. *Time-course graph plotting the change in fixation proportions directed to the target pictures (e.g., frog) and the unrelated distractors in Experiment 2.*

(i.e., green) on hearing "*spinach*", which matched with the stored prototypical colour information that they retrieved from seeing the yellow frog (i.e., that frogs are typically green). The effect, however, was marginal and occurred rather late (more than 1 s after information from the acoustic target word started to become available). These results suggest that stored object colour knowledge can have an influence on language-mediated eye movements but this influence appears to be small. Before we discuss the implications of these findings we present a third experiment investigating whether a colour effect can be observed for objects whose surface colour, but not intrinsic (i.e., prototypical) colour, is similar to the prototypical colour of the concept activated by the spoken word.

EXPERIMENT 3

An object such as a blouse can, in general, be any colour (or at least, many colours), and thus the conceptual make-up (i.e., the stored colour knowledge) of blouses should not include prototypical colour information (at least not the specific colour green that is part of the conceptual make-up of pea). The aim of Experiment 3 was to investigate whether hearing words that are associated with a prototypical colour (e.g., "*pea*") will direct overt attention to objects that are depicted in the same prototypical colour as that of the target word but that are not themselves associated with any prototypical colour (such as a green blouse). Shifts in overt attention to a blouse on hearing "*pea*" can only be due to the perceived surface colour in the visual display.

The second aim of Experiment 3 was to contrast the time-course of language-mediated attention shifts contingent upon surface colour information with shifts in attention contingent upon semantic/conceptual information. The intention is to attempt to draw direct comparisons between semantic/conceptual competitor effects and the predicted surface colour competitor effect. In order to interpret time-course effects appropriately it is important to note that participants in visual-world studies do not comprehend spoken sentences just as they do when comprehending sentences in isolation, with looks to the objects simply being means of indexing aspects of this process. When exposed to visual objects and concurrent spoken language participants actively combine information from the utterance and the visual stimuli to best interpret the situation at hand (cf. Altmann & Kamide, 2007; Huettig & Altmann, 2007; Huettig & McQueen, 2007; Knoeferle & Crocker, 2006). A situation in which participants are exposed to only one competitor (e.g., a surface colour competitor) in the visual display among unrelated distractors is therefore very different from a situation in which participants are exposed to two types of competitor (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 the participant processes two different types of competitor. If only one type of competitor is present in the display 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 competitor are copresent one type of information may be weighed more strongly than the other. Categorical/functional knowledge, for instance, is a particularly salient aspect of lexical knowledge (e.g., Moss et al., 1997); it is not difficult to imagine a model in which the strength of the activation of particular features in a given situation translates into the probability of attending towards whatever shares those features.

Experiment 3 therefore consisted of three conditions. In one condition a surface colour competitor (e.g., a green blouse) was presented among three unrelated distractors; in a second condition a conceptual competitor (e.g., a mushroom) was presented among the same unrelated distractors. A third condition investigated shifts in overt attention (on hearing "*pea*") in the presence of both the surface colour competitor (e.g., a green blouse) and the conceptual category competitor (e.g., a mushroom) contained within the same visual display.

Method

Participants

A total of 60 participants from the University of York student community took part in this study. All were native speakers of British English and either had uncorrected vision or wore soft contact lenses or glasses.

Materials

Three experimental conditions were generated: a perceptual condition, a conceptual condition, and a "both" condition. The same sentences were used in all three conditions (e.g., "*The boy thought about it carefully and then he spotted the pea and asked whether it was a vegetable too*"), but now the visual displays were rotated across the conditions.

The nature of these conditions is described by way of the following examples. In the perceptual condition (Figure 5a), the four pictures included a surface colour competitor (e.g., a green blouse) to the target concept activated by the word ("*pea*") and three distractors that were unrelated conceptually to the target. The colour competitor was presented in the prototypical colour of the target concept (green, in this case). Each distractor

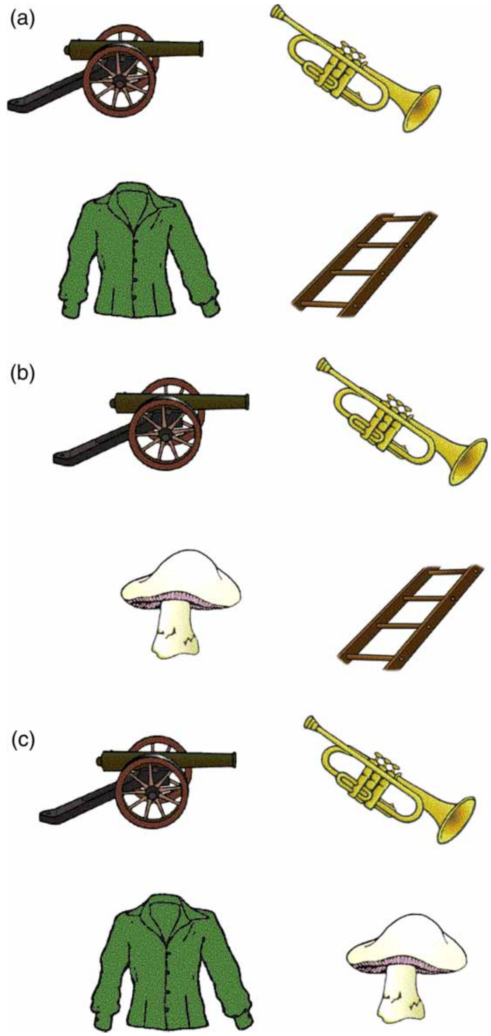

Figure 5. *Example of a visual stimulus in (a) the perceptual competitor condition (depicting surface colour competitor: green blouse and three unrelated distractors), (b) the conceptual competitor condition (depicting conceptual category competitor: mushroom and three unrelated distractors), and (c) the both condition (depicting surface colour competitor: green blouse, conceptual category competitor: mushroom, and two unrelated distractors) in Experiment 3. To view a colour version of this figure, please see the online issue of the Journal.*

was coloured uniquely but in an appropriate manner.

In the conceptual condition (Figure 5b), the displays contained the same distractors as those in the perceptual condition (in the same positions) together with a conceptual competitor (e.g., a mushroom) to the referent of the target word (e.g., "*pea*"). The conceptual competitor occupied the position that in the perceptual condition had been occupied by the surface colour competitor. Conceptual categories were taken from the Battig and Montague (1969) inventory.

In the "both" condition (Figure 5c), the displays contained the perceptual competitor (e.g., the green blouse) from the perceptual condition, the conceptual competitor (e.g., the mushroom) from the conceptual condition, and two of the remaining distractors from the other two conditions.

The following constraints were also applied in selecting the stimulus materials. Primarily items had to be included in the Snodgrass and Vanderwart (1980) set. In addition, in order to ensure that each target word referent (e.g., "*pea*") was associated with a prototypical colour, target words were selected only if they elicited a colour response as indexed by the Nelson word association norms (Nelson et al., 1998; see Table 3). Only words with a FSG for a colour according to the norms were included. The FSGs for the target words ranged from .010 (pig–pink,

Table 3. *Acoustic target words and competitors in Experiment 3*

Acoustic target word			*Competitor picture*	
Word	*Colour*	*FSG*	*Surface colour*	*Conceptual category*
ballerina	pink	.030	pink whistle	clown
bean	green	.140	green sweater	peanut
cabbage	green	.070	green sock	pumpkin
courgette	green	.140	green vase	potato
cucumber	green	.080	green glove	onion
olive	green	.100	green bicycle	carrot
parsley	green	.140	green stool	tomato
pea	green	.260	green blouse	mushroom
pear	green	.030	green lamp	strawberry
pig	pink	.010	pink telephone	sheep
plum	purple	.060	purple ashtray	orange
ruby	red	.270	red kettle	necklace
tongue	pink	.020	pink kite	hair
wagon	red	.190	red clock	plane
walnut	brown	.010	brown cap	banana

Note: FSG: forward strength values for colour associations of the acoustic target words.

walnut–brown) to .270 (ruby–red) with a mean of .107 ($SD = .082$). In order to ensure that the shape similarity between target and surface colour competitors and target and distractors did not differ, a norming study was conducted. A total of 12 participants provided relevant ratings. Participants were presented with the written target word (e.g., pea) together with pictures used in the later eye-tracking study, and they were asked to judge how similar the typical physical shape of the target referent was with the physical shape of the referents of the depicted objects. Participants were asked to judge the shape similarity on a scale from 0 to 10 (0 representing "absolutely no similarity in physical shape"; 10 representing "identical in physical shape"). The mean rating for the surface colour competitors was 1.3 ($SD = 0.96$), for Distractor 1 was 0.8 ($SD = 0.8$), for Distractor 2 was 0.6 ($SD = 0.8$), and for Distractor 3 was 1.3 ($SD = 1.3$). There were no statistically significant differences between these ratings, $F(3, 42) = 2.56$, $p > .05$.

Using the Battig and Montague (1969) category norms as a guide, distractor items were selected from different conceptual categories to that of the target word referent. For example, the surface colour competitor, the (green) blouse, belongs to the Battig and Montague category "clothing", whereas the distractors belonged to the categories "musical instrument", "weapon", and "a carpenter's tool". The selected items were taken from the following categories: four-footed animal, furniture, human body part, kitchen utensil, musical instrument, clothing, type of vehicle, part of building, weapon, fruit, carpenter's tool, bird, toy, insect, and vegetable. Table 3 lists the target words, their forward strength values for colour associations, their perceptual competitors, and their conceptual competitors.

The pictures were matched for picture-naming agreement, image agreement, familiarity, visual complexity, and word frequency of the corresponding name, and the means and standard deviations are shown in Table 4. A repeated measures analysis of variance (ANOVA) revealed that there was no statistically significant difference for the pictures on picture-naming agreement, $F(4, 56) = 0.28$, $MSE = 0.26$, $p > .1$, image agreement, $F(4, 56) = 1.86$, $MSE = 0.30$, $p > .1$, familiarity, $F(4, 56) = 1.94$, $MSE = 0.74$, $p > .1$, visual complexity, $F(4, 56) = 1.70$, $MSE = 0.80$, $p > .1$, and word frequency of the corresponding name, $F(1.656, 21.53) = 0.81$, $MSE = 5{,}816.26$, $p > .1$; Mauchly's test of sphericity showed a significant result ($p < .05$), and thus the conservative Greenhouse–Geisser adjustment was used.

Design

Each participant was presented with a total of 37 trials; 15 were designated experimental trials and 22 as filler trials. On each filler trial the spoken sentence contained the name of one of the pictures. A within-participants counterbalanced design was used across the three conditions. The spoken

Table 4. *Proportions of picture-naming agreement and means for image agreement, familiarity, visual complexity, and word frequency of the visual stimuli in Experiment 3*

Picture	*Picture-naming agreement (proportion)*	*Image agreement (mean)*	*Familiarity (mean)*	*Visual complexity (mean)*	*Word frequency*
Perceptual competitor	.87 (.19)	3.38 (0.68)	3.67 (0.76)	2.68 (0.61)	12.93 (19.17)
Conceptual competitor	.89 (.13)	3.83 (0.50)	3.43 (0.59)	2.74 (0.79)	17.60 (37.01)
Distractor 1	.91 (.11)	3.83 (0.99)	2.89 (0.99)	3.14 (0.97)	11.14 (17.05)
Distractor 2	.88 (.18)	3.81 (0.60)	3.04 (1.01)	3.06 (0.98)	37.87 (88.83)
Distractor 3	.85 (.16)	3.66 (0.63)	3.33 (0.97)	3.41 (0.83)	31.27 (41.07)

Note: Standard deviations in parentheses. Image agreement, familiarity, and visual complexity: scale from 1 to 5. Word frequency from the Kucera–Francis corpus.

sentences were identical across the groups, and the visual displays were rotated across these groups. In each of groups, 5 experimental trials were taken from the perceptual condition, conceptual condition, and the both condition. The same 22 fillers were used for all of the three groups, and within each group the trials were presented in fixed random order.

Procedure
The procedure was the same as that in the previous experiments.

Results

Figure 6 shows the time-course graph in the perceptual condition and the conceptual condition. Participants shifted overt attention towards the surface colour competitor in the perceptual condition and towards the conceptual competitor in the conceptual condition as soon as information from the acoustic target word became available.

Paired t tests showed that in the perceptual condition the mean competitor–distractor ratios during the baseline region (.33) first differed significantly from the mean ratios during the 300-

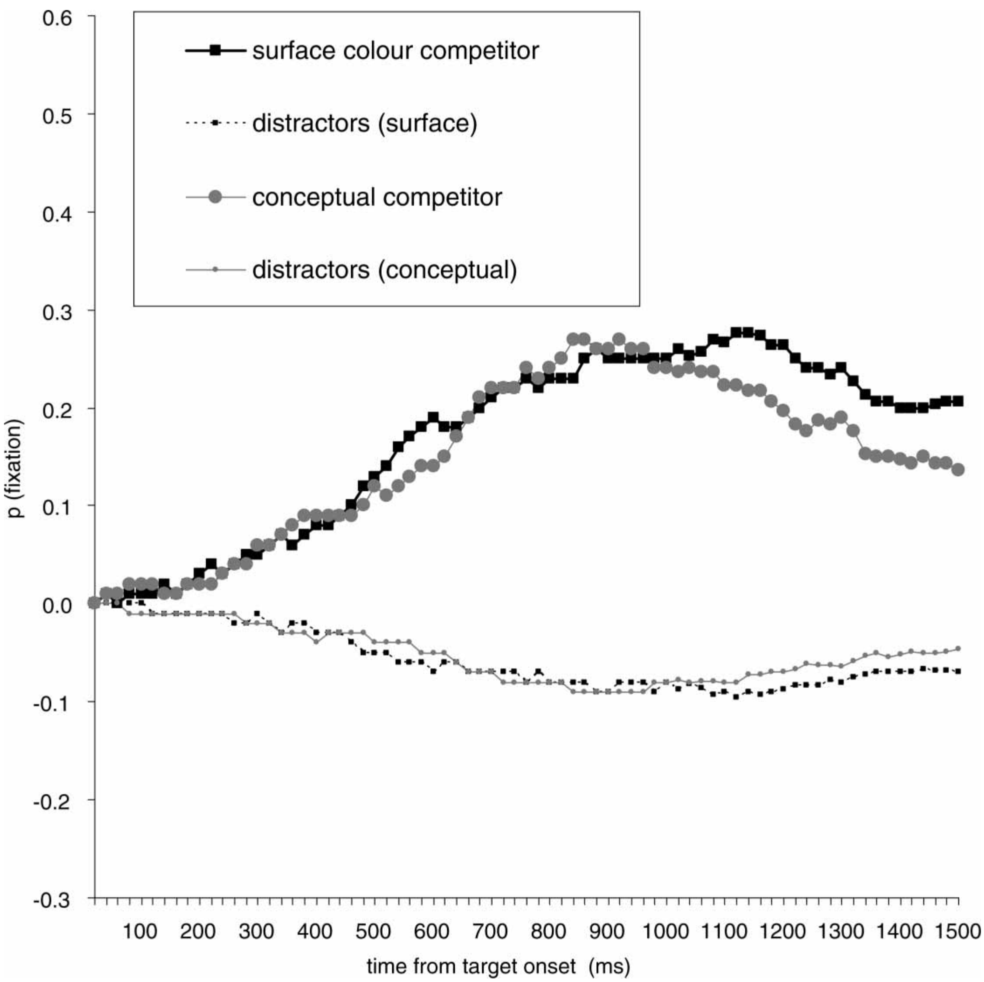

Figure 6. *Time-course graph plotting the change in fixation proportions directed to the competitors and the unrelated distractors in the perceptual and conceptual conditions of Experiment 3.*

ms to 399-ms time region (.44), $t_1(59) = 3.18, p = .002$, $t_2(14) = 2.82$, $p = .014$. The shift in overt attention towards the surface colour competitors remained statistically significant in all subsequent time regions. Thus, despite an initial bias away from the competitors, on hearing the spoken target word in the perceptual condition, participants shifted their attention towards a picture of a conceptually unrelated object that shared a common colour with the named object.

In the conceptual condition the mean competitor–distractor ratios during the baseline region (.42) first differed significantly from the mean ratios during the 300-ms to 399-ms time region (.50), $t_1(59) = 2.69$, $p = .009$, $t_2(14) = 1.88$, $p = .08$, and the 400-ms to 499-ms time region (.53), $t_1(59) = 3.18$, $p = .002$, $t_2(14) = 2.75$, $p = .016$. The shift in overt attention towards the conceptual category competitors remained statistically significant in the subsequent time regions. Thus on hearing the spoken target word in the conceptual condition, participants shifted their attention towards the picture of an object that shared the same conceptual category as the target word.

Figure 7 shows the time-course graph in the both condition. Participants shifted overt attention towards the conceptual competitor as soon as information from the acoustic target word became available. Fixations towards the surface colour competitor started to diverge from the unrelated distractors only at about 700 ms after the acoustic onset of the critical word. Note that the reduction in magnitude of competitor fixations in the both condition is due to two competitors competing for overt attention rather than just one as in the other two conditions.

In the both condition the mean conceptual competitor–distractor ratios during the baseline region (.48) first differed significantly from the mean ratios during the 500-ms to 599-ms time region (.55), $t_1(59) = 2.51$, $p = .015$, $t_2(14) = 1.95$, $p = .07$, and the 600-ms to 699-ms time region (.62), $t_1(59) = 3.63$, $p = .001$, $t_2(14) = 2.39$, $p = .032$. The shift in overt attention towards the conceptual category competitors remained statistically significant in the subsequent time regions. The mean perceptual competitor–distractor ratios during the baseline region (.48) first differed significantly (by items only) from the mean ratios during the 800-ms to 899-ms time region (.66), $t_1(59) = 1.31$, $p = .19$, $t_2(14) = 2.81$, $p = .014$, and more reliably so during the 900-ms to 999-ms time region (.69), $t_1(59) = 1.97$, $p = .054$, $t_2(14) = 3.38$, $p = .004$. The shift in overt attention towards the surface colour competitors remained statistically significant in the subsequent time regions. Thus when both types of competitor were copresent in the visual display shifts to conceptual category competitors preceded shifts towards the surface colour competitors by about 300 ms.

Discussion

On hearing a word that is associated with a prototypical colour (such as "*pea*") participants directed overt attention towards a picture of an object that was portrayed in the relevant colour even though the object is not associated with any diagnostic colour (such as a green blouse). This suggests that hearing "*pea*" activated prototypical colour information, which overlapped with the perceived surface colour of the visually concurrent blouse. Experiment 3 thus shows that language-driven eye movements can be mediated by the surface colour of the perceived objects. In addition, shifts in attention towards the picture of a conceptual competitor took place quicker than did shifts in attention to the picture of a colour competitor when pictures of both were contained in the same visual display.

GENERAL DISCUSSION

Colour can be an important factor that codetermines how we interact with objects in the visual environment. Here, we have shown that for objects that are relatively high in colour diagnosticity, such as a frog, both surface colour and stored colour knowledge can in principle be a factor that mediates language-driven overt attention around a display of visual objects. In the current experiments and on a given trial participants listened

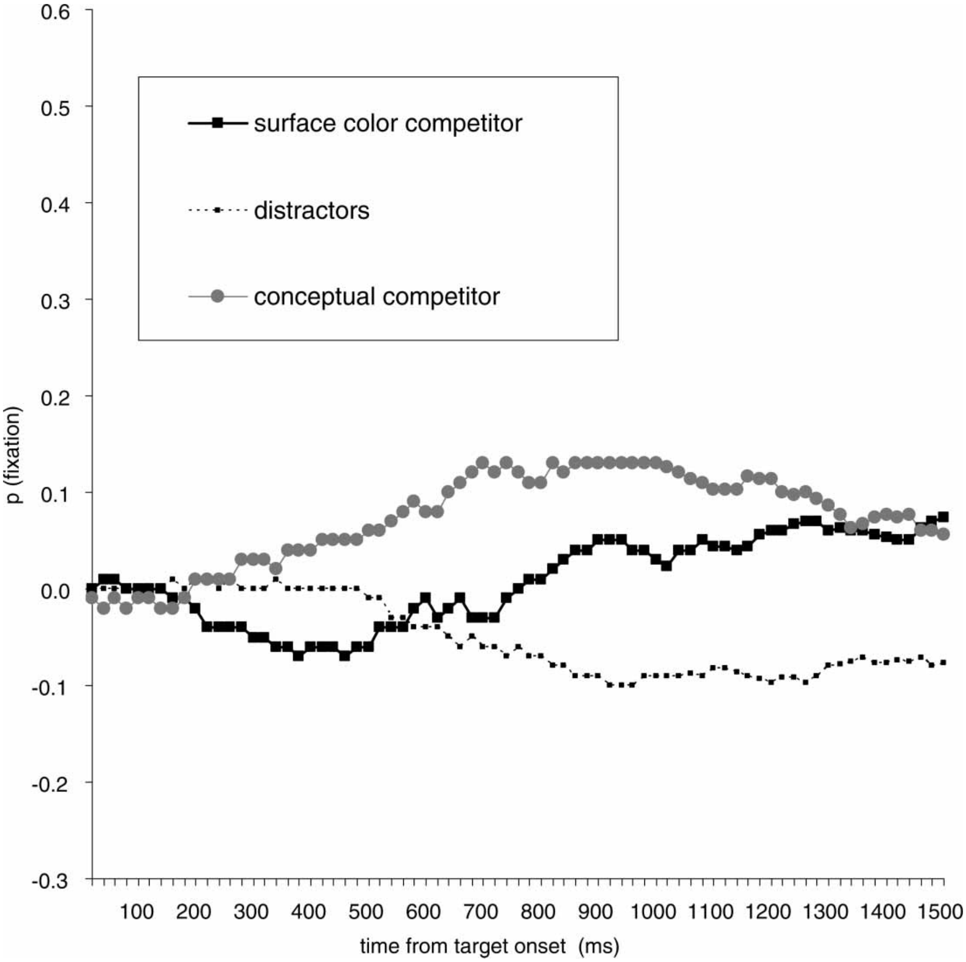

Figure 7. *Time-course graph of the both condition of Experiment 3.*

to a spoken sentence and viewed a visual display containing four pictures of common objects. Their eye movements were measured throughout but they were given no other task instructions. Under very similar experimental conditions previous research has revealed many systematic relationships between the nature of the speech signal and the direction of gaze. Of particular interest is the manner in which shifts in gaze are tightly time locked to unfolding properties of speech and the properties of the objects in the visual display. The basic aim is to examine how shifts in gaze are contingent upon the presentation of a critical target word in the speech signal. In this regard the main index of interest is where the eyes are looking as the target word unfolds. In the present studies an additional issue has been to examine the temporal aspects of these shifts in attention and to ask questions about when such shifts take place relative to the acoustic onset of the target word.

Prior studies have taken such time-locked shifts in attention as revealing the time-course of the access of various types of lexical knowledge from spoken words and its integration with concurrent visual input. For instance, the argument is that very rapid shifts in attention towards semantically related objects reflect the very rapid integration of semantic knowledge. Prior to the current experiments, therefore, the evidence has been taken to reveal, for instance, that eye movements to objects in the concurrent visual environment are driven by semantic similarity rather than all-or-none categorical or associative knowledge (Huettig & Altmann, 2005; Huettig et al., 2006; Yee, Overton, & Thompson-Schill, 2009). In addition Dahan and Tanenhaus (2005) and Huettig and Altmann (2004, 2007) presented evidence that the mapping between spoken words and visual objects can also be mediated by the global visual shape of the referents. However, it remained unclear from these studies whether these visual effects were contingent upon perceived information (e.g., the perceived surface colour of the visual object) or stored conceptual information (e.g., the stored knowledge about the typical colour of the displayed object). The present experiments manipulated colour relations between the lexical concepts accessed by the spoken words and the visual objects in order to dissociate conceptual attributes (the stored colour knowledge about an object) and perceptual attributes (the perceived but nondiagnostic colour of an object).

Experiment 1 was designed to investigate the effect of stored colour knowledge of visual objects in the absence of surface colour. We found no robust evidence, when participants heard spoken words such as "*spinach*", of a bias to look towards the black-and-white line drawings of an object (e.g., a frog) that is typically of the same colour as that associated with the target spoken word. A caveat with this finding was that the black-and-white line drawings may have induced an attentional control setting that reduced attention to colour features because they were of little task relevance. In Experiment 2 thus we further investigated whether stored object colour can ever be used to direct language-mediated overt attention. We presented colour photographs (with the target objects shown in an atypical but appropriate colour) and black-and-white versions of the same photographs. With the colour photographs we observed an influence of stored object colour knowledge but the effect was small and occurred rather late (more than one second after information from the acoustic target word became available). With the black-and-white photos we replicated the results from Experiment 1 that black-and-white stimuli do not give rise to a significant stored object colour effect. Finally, Experiment 3 investigated the effect of the perceived surface colour in the absence of stored colour knowledge. Participants, on hearing target words that are associated with a prototypical colour, such as "*pea*", looked towards a picture displayed in that colour even though the referent of the picture was not itself associated with that colour. On accessing the prototypical colour information of the target referent, participants shifted overt attention to anything with the same surface colour. Therefore Experiment 3 was a clear demonstration of a pure surface colour effect. Experiment 3 also revealed a time-course difference between shifts in overt attention to (surface) colour competitors and shifts in overt attention to conceptual competitors (nonassociated category coordinates such as pea/mushroom) when both types of competitor were present in the same visual display.

What are the implications of our findings for the study of language–vision interactions?

Our research suggests that it is not only stored knowledge but also the perceived surface attributes of visual objects that are a principal determinant of language-mediated visual attention. Interestingly, at least for colour, the perceived surface attributes of visual objects seem a more important determinant of language-mediated overt attention than stored knowledge about the typical visual form of the referent.[2] This suggests that the visual shape effects (snake/cable) reported previously both by

[2]Although compelling, this result does not invalidate the findings of a strong influence of stored colour knowledge in other tasks. Indeed the task constraints in previous studies on this issue seem very different from the ones in our research. For instance, in the task employed by Joseph and Proffitt (1996), participants had to decide whether a briefly presented and masked picture matched the subsequently presented object name. Therefore participants were explicitly engaged in an object verification task. Moreover, there are

ourselves (Huettig & Altmann, 2004, 2007) and by Dahan and Tanenhaus (2005) may also been driven primarily by the perceived surface attributes of the visual objects rather than the stored knowledge about the typical shape of the objects (though this is an open empirical question).[3]

The present research highlights interesting differences between the information that is accessed via linguistic modes versus that via visual modes. The colour competitor effects obtained in Experiment 2 and Experiment 3 are contingent upon access of stored colour knowledge on hearing the spoken critical words (such as "*spinach*" or "*pea*"). Moreover the conceptual make-up of the critical visual stimuli in Experiment 3 did not contain stored colour knowledge (e.g., the green blouse). Thus the only reason why participants looked at the green blouse in Experiment 3 is that they accessed the colour knowledge (e.g., the greenness of a pea) from the spoken word. Interestingly, stored colour knowledge seems not to have been accessed in Experiments 1 and 2 on seeing black-and-white line drawings and black-and-white photographs. This seems to have been due to participants adopting an attentional set in which colour was not relevant.[4]

Differential effects contingent upon verbal versus visual representation of the stimuli have also been reported by Naor-Raz et al. (2003). These authors observed an effect of stored colour knowledge in a Stroop colour-naming task but found that the direction of the effect differed for verbal and visual stimuli. In addition, they followed up these colour-naming tasks with a lexical decision task in which some items were conceptually associated to items they had previously presented in the colour-naming task. It was found that verbal but not visual presentation during colour naming resulted in subsequent facilitation of responses (banana priming monkey). Naor-Raz et al. concluded that verbal stimuli resulted in a more ready access of stored conceptual knowledge than visual stimuli in the colour-naming task. Our data support this account.

Our previous research (Huettig & Altmann, 2005; Huettig et al., 2006) and the present Experiment 3 of course show that categorical/functional conceptual knowledge is readily accessed from seeing black-and-white pictures of objects. Our findings could thus be taken as support for a visual form versus semantic/conceptual dichotomy of stored conceptual knowledge. Our stored knowledge of the concept banana evidently includes knowledge of physical properties (e.g., visual attributes; we know bananas are yellow) and knowledge of nonphysical properties (e.g., functional attributes; we know bananas are eatable). Miceli et al. (2001) recently reported a brain-damaged patient (I.O.C.) who is severely impaired in accessing stored colour knowledge but shows no impairments for functional knowledge. The current study provides further evidence that storage and/or retrieval of colour properties of concepts is independent of storage/retrieval of nonphysical properties (e.g., functional knowledge). Interestingly, Miceli et al. also found a dissociation of object colour knowledge from object shape knowledge in patient I.O.C. This suggests that storage and/or retrieval of colour properties of concepts is also independent of storage/retrieval of object shape knowledge.

In Experiment 2 we observed a small (and delayed) effect of stored object colour on language-mediated overt attention with colour photographs. Why did people attend to a yellow frog more than one second after hearing "*spinach*", when they attended to a green blouse after only 300 ms (Experiment 3, perceptual condition)? The fact that they can attend to a surface colour competitor after only 300 ms

important differences with respect to the stimuli. Our displays contained four objects rather than just one. In addition, the Joseph and Proffitt stimuli were presented for 50 ms followed by a 50-ms mask, whereas in our experiments the visual objects are presented for a much longer duration.

[3]It may be possible to explore this issue by presenting visual objects in an atypical visual form (e.g., an open vs. a closed umbrella).

[4]How attentional sets are established in tasks such as the one used here is beyond the remit of the present research. Future research, however, could usefully be directed at this issue.

shows that there is fast access of colour knowledge from spoken words that are associated with a prototypical colour. The difference in the time-course of attentional shifts due to surface colour and those due to stored colour knowledge therefore appears to be due to the access and/or use of stored colour knowledge from the visual objects during language–vision interactions. The data of all three experiments reported here fit best with the notion that language-mediated attention based on colour information is determined primarily by the surface colour of the visual objects. The delay and small magnitude of the stored object colour effect in Experiment 2 suggest that stored object colour is only used when there is no match on surface colour (or other dimensions such as shape or conceptual category).

Our data, however, do not conclusively rule out that stored colour knowledge was accessed from the visual objects considerably earlier than 1 second after the onset of the target word but that such access was masked by the presence of the nonmatching surface colour of the competitor. Note in this regard that our past research (Huettig & Altmann, 2004, 2005; Huettig & McQueen, 2007; Huettig et al., 2006) has shown that participants shift their attention towards the visual object in the display that best matches the conceptual and perceptual specification of the concept activated by the spoken word. In other words the best matching object in the visual display is fixated even if this object has little in common with the target concept activated by the spoken target word. The atypical colour competitors in the displays in Experiment 2 (e.g., the yellow frog) were the only objects in the display with any match (a match in stored colour) between spoken target concept and the visual display. Of course, the atypical colour competitors also mismatched on surface colours but so did all the unrelated distractors in the displays. The atypical colour competitors were thus still the best matching objects in the display.

The incompatible surface colour, however, may have inhibited attentional shifts towards the otherwise best matching object (the stored colour knowledge competitors) in the display. It is at least conceivable that the stored conceptual colour information associated with the picture of a frog (i.e., the colour green), which is compatible with the target word ("spinach"), competed with the surface colour information (i.e., the colour yellow), which is incompatible with the target word. This may have greatly reduced the likelihood of overt attention to the competitor. In other words, there was one object in the display that matched but also mismatched with the target word on one dimension—that is, it is possible that the perceptual mismatch inhibited language-mediated eye movements based on a conceptual match. Such possible effects of a conceptual match versus a perceptual mismatch on a single object dimension have to our knowledge not been investigated during language–vision interactions. We believe that future research could usefully be directed at this issue.

Why did looks to the surface colour competitors occur later than looks to the conceptual competitors in the both condition of Experiment 3 but there was no delay in attentional shifts to surface colour competitors when there were no conceptual competitors copresent as in the perceptual condition of Experiment 3? As discussed above, the only reason why participants looked at the surface competitor is that they accessed the stored colour knowledge from the spoken word. We believe that the time-course differences observed in the both condition are due to categorical/functional knowledge being a particularly salient aspect of word meaning (e.g., Moss et al., 1997). Our findings suggest that the strength of the activation of particular features translates into the probability of attending towards whatever shares those features. Our research reveals that even for natural objects that are relatively high in colour diagnosticity categorical/functional knowledge is a more salient aspect of knowledge than colour knowledge. Stored colour knowledge is a quite different kind of conceptual feature to category membership—a frog is an animal whatever its colour. It is thus unclear what manner of abstraction results from the experience of different colours of different things (even fruits and vegetables). Colour appears to be largely context dependent in a way that category membership is not, which is why for the majority of their lifetimes bananas are not yellow, and yet they are still fruit. We suggest that this distinction between category membership and prototypical colour is the cause of the time-course difference in the both condition of Experiment 3. The situation is different if a colour match is the only match possible between spoken word and visual object as in the perceptual condition of Experiment 3. In such circumstances language comprehenders will tend to use colour information as soon as it becomes available and attend towards objects that share those features. Thus, there was no delay in shifts in overt attention to surface colour competitors when no conceptual category competitors were copresent in the displays.

We conclude that the mapping between language and concurrent visual objects can be mediated by both perceived surface attributes and conceptual knowledge of visual objects. If conceptual category competitors and surface colour competitors are copresent in the visual environment overt attention tends to be preferentially influenced by conceptual category information. If only colour matches are possible between spoken word and a visual referent the attentional system appears to primarily rely on the perceived surface attributes of the visual objects rather than stored knowledge about the typical colour of the object. More generally, our data provide converging evidence for the notion that spoken words result in a more ready access of stored colour knowledge than do visual objects (cf. Naor-Raz et al., 2003) and for a colour versus semantic/functional distinction of stored knowledge (cf. Miceli et al., 2001).

Original manuscript received 5 August 2009
Accepted revision received 3 February 2010
First published online day month year

REFERENCES

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*, 419–439.

Allport, D. A. (1971). Parallel encoding within and between elementary stimulus dimensions. *Perception & Psychophysics*, *10*, 104–108.

Altmann, G. T. M., & Kamide, Y. (2004). Now you see it, now you don't: Mediating the mapping between language and the visual world. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision and action* (pp. 347–386). Hove, UK: Psychology Press.

Altmann, G. T. M., & 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.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monographs*, *80*(3), 1–46.

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

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.

Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, *132*, 163–201.

Dahan, D., & Tanenhaus, M. (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*, *40A*, 533–544.

Flores d'Arcais, G. B., Schreuder, R., & Glazenborg, G. (1985). Semantic activation during recognition of referential words. *Psychological Research*, *47*, 39–49.

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). *Journal 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.

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: 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., & 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,* 341–360.

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., 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.

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.

Kumada, T. (2001). Feature-based control of attention: Evidence for two forms of dimension weighting. *Perception & Psychophysics*, *63,* 698–708.

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.

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

Miceli, G., Fouch, E., Capasso, R., Shelton, J. R., Tomaiuolo, F., & Caramazza, A. (2001). The dissociation of colour from form and function knowledge. *Nature Neuroscience*, *4,* 662–667.

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,* 667–680.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). *The University of South Florida word association, rhyme, and word fragment norms*, Retrieved December 10, 2004, from http://www.usf.edu/FreeAssociation/.

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.

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.

Pratt, J., & Hommel, B. (2003). Symbolic control of visual attention: The role of working memory and attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *29,* 835–845.

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

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

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity.

Journal of Experimental Psychology: Human Learning and Memory, *6,* 174–215.
Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology*, *40A,* 201–237.
Yee, E., Overton, E., & Thompson-Schill, S. L. (2009). Looking for meaning: Eye movements are sensitive to overlapping semantic features, not association. *Psychonomic Bulletin & Review*, *16,* 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.