Language-induced visual and semantic biases in visual search are subject to task requirements

Floor de Groot, Falk Huettig & Christian N. L. Olivers

Published online: 25 May 2017.

Submit your article to this journal ↗

Article views: 190

View related articles ↗

View Crossmark data ↗

Routledge
Taylor & Francis Group

Language-induced visual and semantic biases in visual search are subject to task requirements

Floor de Groot[a], Falk Huettig[b,c] and Christian N. L. Olivers[a,d]

[a]Department of Experimental and Applied Psychology, Vrije Universiteit, Amsterdam, The Netherlands; [b]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; [c]Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, The Netherlands; [d]Institute for Brain and Behaviour, Vrije Universiteit, Amsterdam, The Netherlands

ABSTRACT

Visual attention is biased by both visual and semantic representations activated by words. We investigated to what extent language-induced visual and semantic biases are subject to task demands. Participants memorized a spoken word for a verbal recognition task, and performed a visual search task during the retention period. Crucially, while the word had to be remembered in all conditions, it was either relevant for the search (as it also indicated the target) or irrelevant (as it only served the memory test afterwards). On critical trials, displays contained objects that were visually or semantically related to the memorized word. When the word was relevant for the search, eye movement biases towards visually related objects arose earlier and more strongly than biases towards semantically related objects. When the word was irrelevant there was still evidence for visual and semantic biases, but these biases were substantially weaker and similar in strength and temporal dynamics without a visual advantage. We conclude that language-induced attentional biases are subject to task requirements.

When hearing or reading a word, people direct their gaze to the object in the visual environment that the word refers to. Various studies, including by Glyn Humphreys and colleagues, have shown that such language-induced orienting occurs for different levels of representation, resulting in either visually or semantically driven biases in visual selection (e.g., Belke, Humphreys, Watson, Meyer, & Telling, 2008; Dahan & Tanenhaus, 2005; Huettig & Altmann, 2005, 2007; Huettig & McQueen, 2007; Moores, Laiti, & Chelazzi, 2003; Rommers, Meyer, Praamstra, & Huettig, 2013; Soto & Humphreys, 2007; Telling, Kumar, Meyer, & Humphreys, 2010). The present study combines two of Glyn Humphreys' research legacies: the type of representation that is available for attentional selection (visual or semantic), and the extent to which observers' task requirements determine selection.

The first and by far strongest demonstrations of language-induced visual orienting come from the field of psycholinguistics. Studies have made use of the visual world paradigm (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; for a recent review, see Huettig, Rommers, & Meyer, 2011), in which observers are first presented with a number of pictures of objects, and then hear a word or a phrase. Observers have been found to spend more time fixating on objects that are visually, semantically, or phonologically related to a critical word. For example, when hearing the word banana, observers fixate more on a picture of a canoe (visually related) and a monkey (semantically related) relative to unrelated objects (e.g., a hat or tambourine). Because of the reliance on world knowledge, such language-mediated attentional biases are typically driven by overlearned associations and can occur relatively rapidly (within a few hundred milliseconds from word onset, although exact timing is difficult as the linguistic input develops across time). Moreover, such biases have been found in studies using a passive viewing paradigm where language was not directly relevant (e.g., Huettig & Altmann, 2005; Rommers et al., 2013; see Huettig, Rommers, et al., 2011, for a review), or where it was even irrelevant for the task (e.g., Salverda & Altmann, 2011). As such,
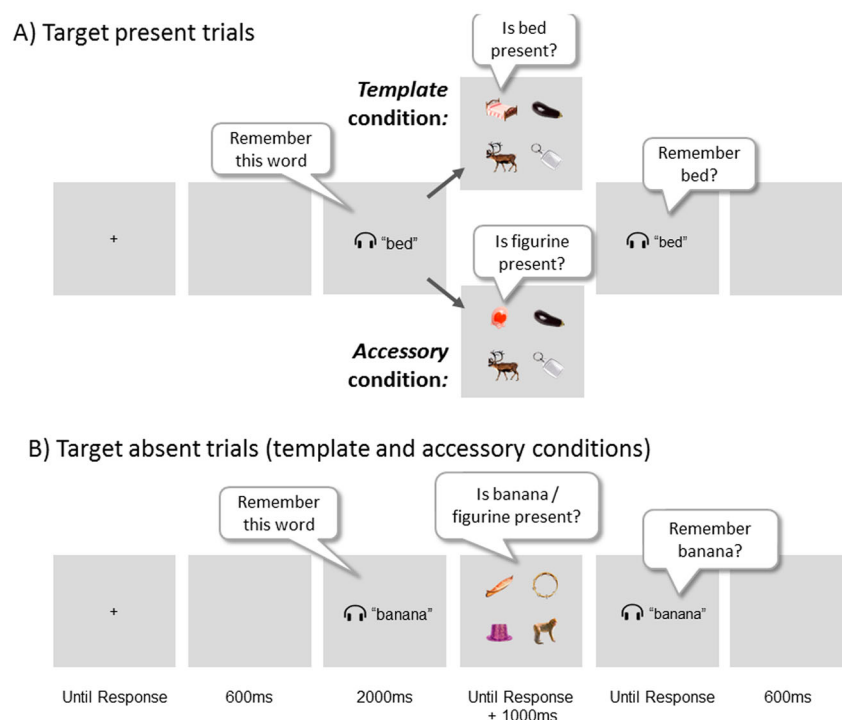
these biases fit most of the criteria of an automatic process (cf. Logan, 1988; Mishra, Olivers, & Huettig, 2013; Moors & De Houwer, 2006). Indeed, existing models of language-induced orienting either explicitly or implicitly assume that activation automatically spreads to associated representations, thus resulting in automatic orienting biases (Huettig & McQueen, 2007; Mayberry, Crocker, & Knoeferle, 2009; Mirman & Magnuson, 2009; Smith, Monaghan, & Huettig, 2013; Spivey, 2008). For example, the *cascaded activation model* of visual–linguistic interactions (Huettig & McQueen, 2007) assumes that incoming words first activate phonological representations, after which activation spreads to both visual and semantic representations. Incoming pictures first activate visual representations, after which activation spreads to associated semantic and phonological representations. Consistent with this, we have recently shown that the relative timing of visual and semantic biases is strongly affected by whether the linguistic or the visual stimulus is presented first (de Groot, Huettig, & Olivers, 2016a). However, in that study we did not manipulate task requirements, as observers were required to look for the referred-to object in all conditions.

Language-induced semantic biases in visual orienting have also been found in the field of visual attention research (Belke et al., 2008; Moores et al., 2003; Telling, Kumar, et al., 2010; see also Soto & Humphreys, 2007, for a related finding). These researchers have made use of the visual search task, where observers receive, prior to object presentation, a linguistic instruction as to which target to look for among distractors. As in the visual world paradigm, one of the distractors could be semantically related to the sought-for target. Belke et al. (2008) as well as Moores et al. (2003) reported biases in overt attention towards such semantically related distractors. However, whereas psycholinguistic models of language–vision interactions allow for orienting on the basis of semantic representations, and assume these representations to become automatically available for guiding attention, current models of visual search postulate the exact opposite. To date, semantic information associated with individual objects has not been incorporated in any formal models of visual search (e.g., Navalpakkam & Itti, 2005; Zelinsky, 2008), or has even been explicitly excluded from being available for attentional guidance (Wolfe, 2007).[1] At the same time, these models do not

assume attentional orienting to be driven by automatically activated representations. Instead, they assume orienting to be based on the activation of task-relevant representations, leading to top-down guidance towards likely target candidates, while task-irrelevant objects are ignored.

In the present study, we show evidence that further bridges the fields of psycholinguistics and visual attention and argue that crucial aspects of psycholinguistic and visual attention models are best combined. Orienting can be driven by semantic as well as visual properties, but the absolute and relative strength of such biases is to a large extent subject to task relevance. We adopted the combined memory and visual search task illustrated in Figure 1. Participants were instructed to memorize a spoken word for a subsequent verbal recognition task at the end of each trial. During the retention period, they performed a search task. Both the memory task and the search task had to be completed in all conditions. Importantly, participants either searched for the object referred to by the memorized word, or they searched for a different target object. In the first case, the word would have to be transformed into a *template* for search; in the latter case, it had to be remembered but shielded from the search task in what we will refer to as the *accessory* state (Olivers, Peters, Houtkamp, & Roelfsema, 2011). Critical were the target absent trials, where search displays contained an object that was semantically (but not visually) related, an object that was visually (but not semantically) related to the memorized word, and two unrelated objects. These target absent displays were the same for the template condition and the accessory condition; all that differed was the task structure. We measured overt orienting biases towards the visually and semantically related objects (relative to unrelated objects) as a proxy for which type of representation was activated by the word.

Our predictions were as follows. If language-induced attentional biases are subject to task requirements, then stronger biases should occur when the word is relevant for the search (in the *template* condition) for both visual and semantic representations. Moreover, following de Groot et al. (2016a), observers are expected to rely more on visual representations than on semantic representations when the word is relevant for the visual search task (again in the *template* condition). The specific predictions for the

Figure 1. Illustration of the procedure and different trial types, with various stimulus examples. On all trials, participants memorized a spoken word for a verbal recognition test at the end of the trial. During the retention period, they performed a visual search task. Panel A) illustrates target present trials, where in the *template* condition, people searched for the object referred to by the word, while in the *accessory* condition, the word was not relevant for the search, but was still needed for the memory test. Here, observers searched for a member of a set of plastic figurines instead. The crucial trials were the target absent trials, illustrated in Panel B). These contained an object that was *semantically* related (in this example the monkey), an object that was *visually* related (here the canoe) and two objects that were unrelated (here the hat and the tambourine) to the memorized word. In terms of stimulus sequence, these trials were identical for both template and accessory conditions. However, as for the present trials, the task set differed per block, as in the template condition observers were searching for the spoken word, while in the accessory condition they were searching for a figurine, while still remembering the word for the memory test.

accessory condition were a priori more difficult to make. Although we expected overall weaker biases, how one type of bias would relate to the other in terms of timing or strength was less obvious. Accessory memories must be prevented from interfering with the current visual task, and thus we might expect them to be especially less strongly instantiated at the visual level than at a more semantic level (Bae & Luck, 2016; Christophel, Klink, Spitzer, Roelfsema, & Haynes, 2017), which would predict a relative dominance of semantic over visual biases.

Method

Participants

A planned number of 24 Dutch native speakers (four male, aged 17–37, average 21.6 years) participated for course credits. Sample size was based on our previous studies (de Groot et al., 2016a; de Groot, Huettig, & Olivers, 2016b). No participant was

replaced or left out of the analyses. None reported to suffer from colour blindness and/or language disorders or had participated in our earlier studies. The study was conducted in accordance with the Declaration of Helsinki and approved by the Scientific and Ethical Review Board of the Faculty of Behaviour and Movement Sciences at the Vrije Universiteit Amsterdam.

Stimuli, apparatus, design, and procedure

Participants were tested on a HP ProDesk 600 G1 CMT computer with a Samsung Syncmaster 2233RZ monitor (refresh rate of 120 Hz, resolution 1680 × 1050). The distance between the monitor and the chin rest was 70 cm. Stimuli were presented using OpenSesame (Mathôt, Schreij, & Theeuwes, 2012) version 2.9.7. An Eyelink 1000 Desktop Mount with a temporal and spatial resolution of respectively 1000 Hz and 0.01° was used to track the left eye.

Words were presented through headphones (Sennheiser HD202) connected via a USB Speedlink soundcard.

The stimuli were based on Gunseli, Olivers, and Meeter (2016) and de Groot et al. (2016a), and we refer to those studies for full details. There were 240 search displays, each consisting of four objects. The crucial trials were the 120 target absent trials, which contained one object that was *semantically* (but not visually) related, one object that was *visually* (but not semantically) related, and two objects that were *unrelated* to a spoken word as independently rated by 61 Dutch natives. In the rating study participants indicated on a 11-point scale how much the depicted object and the object that the spoken word was referring to "had something to do with each other, i.e., shared something in meaning or function" (semantic rating study) or how much they "looked alike" (visual rating study). Raters were instructed to focus only on the relationship of interest, and to ignore any other relationships as much as possible. The mean semantic word–picture similarity rating for the semantically related picture was 6.74 (SD = 1.29), and 0.98 (0.58) for the visually related picture (a reliable difference at $p < .001$). Conversely, the mean visual word–picture similarity rating for the semantically related picture was 1.17 (1.02), and 5.01 (1.65) for the visually related picture (again reliable at $p < .001$). The average similarity word–picture ratings for the neutral distractors were 0.40 and 0.71 for semantic and visual similarity, respectively. In addition, the different object categories were controlled for several visual and psycholinguistic factors, including luminance, visual complexity, and naming agreement. An overview of all target absent trials can be found in Appendix A. Objects on the target absent trials had an average overall surface size of 39,292 pixels (SD = 22302 pixels) and the average radius of the smallest fitting circle around the object was 181 pixels (4.17°; SD = 15 pixels, 0.35°). In the other 120 search displays, the target was present. All pictures were scaled to 346 by 346 pixels (8°) and were presented on a grey background (RGB: 230,230,230), one in each quadrant (randomized per trial and per participant; 8° horizontal distance, 6° vertical distance).

Apart from the visual and semantic word–picture relationships on target absent trials, the study employed a 2 × 2 within-participant design with Trial Type (target absent/target present) and Task Relevance (template/accessory) as factors. Participants were instructed to remember a word for a verbal recognition task at the end of each trial. In the template condition, this word was also relevant for the search as it directly indicated the target object. In the accessory condition, the word was not relevant for the search but only for the subsequent memory test. The search target was a picture of a plastic figurine, randomly drawn from a set of nine instances. The figurines differed in shape and colour, but were all recognizable as belonging to the same target category. These plastic figurines were chosen because, although a distinct category, they varied in colour and shape similar to the other objects in the stimulus set and did not pop out as such. Appendix B shows the set of figurines used as the targets in the accessory condition. By using the same target category throughout, plus the fact that the figurines were repeated across trials, we minimized the working memory load for the target, leaving sufficient working memory capacity for the to-be-remembered, but accessory, word (Gunseli et al., 2016; Olivers, 2009). Except for the plastic figurines, each picture and spoken word was only presented once. Trial Type was mixed within blocks, whereas Task Relevance was blocked in a counterbalanced ABABAB design, with six blocks × 40 trials per block. Target presence was randomly mixed within blocks (50% present and 50% absent). The memory test at the end of the trial showed the memory item on 50% of the trials (same response), and a different item on the other 50%. Written instructions before each block indicated whether the search target was the spoken word or the plastic figurine. Displays were counterbalanced such that if for participant A a random set of stimuli was shown in the template condition, the same set was shown in the accessory condition for participant B. Participants received one practice trial of each condition. Feedback was given only during the practice trials.

The trial started with a drift correction triggered by the space bar, and followed by a blank screen for 600 ms. Then the to-be-memorized word was presented through headphones. The search display followed 2000 ms after word onset. Participants used the keyboard to indicate whether the target was present ("J") or absent ("N"). After the response they heard a click, and the search display stayed on the screen for another 1000 ms. Then, as a memory test, a spoken word was presented again. Participants had to indicate whether this word was similar ("S")

or different ("D") than the word they had heard before the search display. After a blank screen for 600 ms, a new trial started (see Figure 1).

Results

Manual responses

Average search reaction times (RTs) of the correct trials (i.e., correct responses on both the search and the memory task) were entered in a repeated measures analysis of variance (ANOVA) with Trial Type (target absent/present) and Task Relevance (template/accessory) as factors. Search was faster on target present ($M = 1014$ ms, $SD = 249$) than on target absent trials ($M = 1212$ ms, $SD = 301$), Trial Type, $F(1,23) = 28.84$, $p < .001$, $\eta_p^2 = 0.556$, and was also faster in the accessory ($M = 1035$, $SD = 257$) than in the template condition ($M = 1200$ ms, $SD = 306$), Task Relevance, $F(1,23) = 90.26$, $p < .001$, $\eta_p^2 = 0.797$. There was no interaction, $F(1,23) = 0.52$, $p = .480$. The pattern of search errors followed that of the RTs. Participants made more errors on target present ($M = 8\%$, $SD = 4\%$) than on target absent trials ($M = 2\%$, $SD = 2\%$), Trial Type, $F(1,23) = 10.46$, $p < .01$, $\eta_p^2 = 0.313$. More errors were made in the template ($M = 6\%$, $SD = 3\%$) than in the accessory condition ($M = 4\%$, $SD = 3\%$), Task Relevance, $F(1,23) = 58.51$, $p < .001$, $\eta_p^2 = 0.718$. There was no interaction, $F(1,23) = 0.31$, $p = .584$. The proportion of memory errors was low on both target absent (both accessory and template condition: $M = 2\%$, $SD = 2\%$) and target present trials (accessory condition: $M = 2\%$, $SD = 2\%$; template condition: $M = 2\%$, $SD = 3\%$). There were no effects on memory errors, $Fs \leq 1.10$, $ps > .305$.
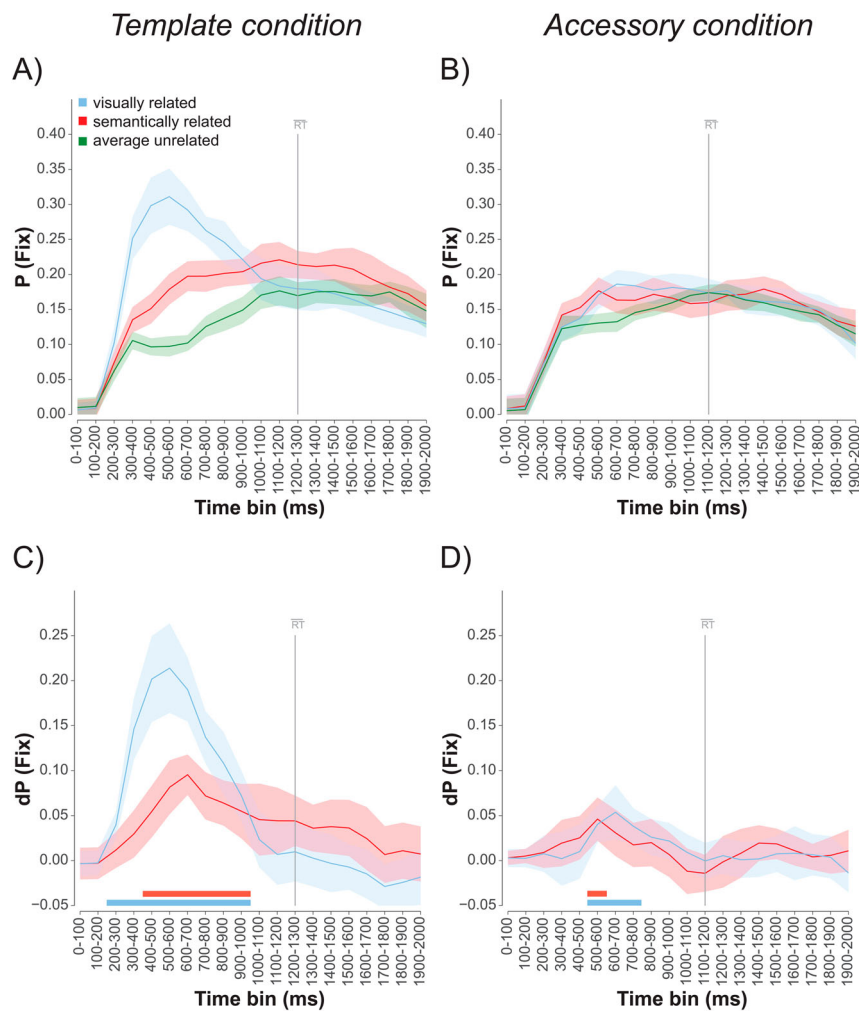
Eye movement data: proportion fixation time

The primary measure of interest was which objects were being fixated on target absent trials as a function of their visual or semantic relatedness and task relevance. Only trials with a correct response on both the search and the memory task were included. Regions of interest (ROIs) were defined as 8° squared areas, centred on the middle of each picture completely covering all objects.

Overall biases. We conducted a repeated measures ANOVA on the overall mean absolute proportion fixation time: the proportion of time spent fixating

objects, from 150 ms after search display onset until the response time of each trial) with Task Relevance (template/accessory) and Object Type (semantically related/visually related/unrelated) as factors. Greenhouse-Geisser corrected values are reported when sphericity was violated. All subsequent pairwise comparisons were planned. Nevertheless, we will indicate when p-values do not hold under a Bonferroni correction. The ANOVA revealed main effects for Task Relevance, $F(1,23) = 65.53$, $p < .001$, $\eta_p^2 = 0.740$, Object Type, $F(1.518,34.903) = 66.04$, $p < .001$, $\eta_p^2 = 0.742$, and, importantly, a highly significant interaction, $F(2,46) = 77.63$, $p < .001$, $\eta_p^2 = 0.771$. To trace the source of this interaction, we conducted follow-up t-tests, which showed that in the template condition participants spent more time fixating the visually ($M = 0.22$, $SD = 0.05$) and semantically ($M = 0.16$, $SD = 0.04$) related objects than the unrelated objects ($M = 0.11$, $SD = 0.03$), $t(23) = 12.74$, $p < .001$, $r = 0.936$ and $t(23) = 8.77$, $p < .001$, $r = 0.877$ respectively. Moreover, people spent overall more time fixating visually related than semantically related objects, $t(23) = 6.37$, $p < .001$, $r = 0.799$. In the accessory condition, effects were much reduced. The overall time spent on visually related objects ($M = 0.14$, $SD = 0.05$) was reduced to such an extent that it did not differ significantly anymore from the time spent on unrelated objects ($M = 0.13$, $SD = 0.04$), $t(23) = 1.66$, $p = .11$). Observers still spent overall more time fixating semantically related ($M = 0,14$, $SD = 0.04$) than unrelated objects ($M = 0.13$, $SD = 0.04$), $t(23) = 2.15$, $p < .05$, $r = 0.409$, although this effect does not hold under Bonferroni correction. In addition, there was no longer a difference between time spent on visually and semantically related objects ($M = 0.14$, $SD = 0.04$), $t(23) = 0.053$, $p = .96$.

Time course analyses. A drawback of looking only at overall biases is that it may hide more complex underlying dynamics of such biases. Following de Groot et al. (2016a), we therefore investigated the time course of the visual and semantic biases more closely. A period of 2000 ms from search display onset was divided into twenty 100 ms bins. For each time bin, we computed the proportion of time that people spent fixating a particular object in a bin (proportion fixation time or "P(fix)"), together with within-subjects 95% confidence intervals (Figure 2A and 2B). Proportion fixation time for the two unrelated objects was averaged. Figure 2C and 2D then show the difference in proportion fixation time between the related

Figure 2. Target absent trials. The proportion fixation time, P(fix), towards visually related, semantically related, and unrelated objects for every 100 ms time bin for the template (A) and the accessory (B) condition and the difference in proportion fixation time, dP(fix), for the semantically and visually related objects relative to the average of the unrelated objects (C for template condition and D for the accessory condition). The grey vertical lines mark the average search times for each condition. Error bars reflect 95% confidence intervals (two-tailed) for within-participants designs (Cousineau, 2005; Morey, 2008). The red and blue horizontal bars mark the bins for which there was a significant semantic (red) and visual (blue) bias as indicated by t-tests for each bin, corrected for multiple comparisons using Bonferroni correction (correcting for the number of bins up to the one including the response, i.e., α = 0.05/13).

and the average of the unrelated objects ("dP(fix)"), together with 95% confidence intervals (which we take as the indicator of reliability). These graphs show that, in the template condition, the visual bias is stronger and arises earlier than the semantic bias. The visual bias reliably starts around 200–300 ms and peaks around 500–600 ms, whereas the semantic bias emerges from 300–400 ms and peaks around 600–700 ms. Individual t-tests comparing the bias for each bin against zero confirmed this pattern, showing significant visual biases from bin 200–300 to bin 900–1000, and significant semantic biases from bin 400–500 to bin 900–1000 (all Bonferroni corrected for the number of bins up to the one including

the average response time, i.e., α = 0.05/13 = 0.0038). The graphs also show clearly that both the visual and the semantic bias were weaker in the accessory condition. Interestingly, the strength and time course of the semantic and the visual bias now appeared quite similar. Importantly, the biases in this condition were within the same time window as in the template condition (and as in similar conditions in our previous work; de Groot et al., 2016a): both visual and semantic biases emerged from 400–600 ms, and peaked at 500–700 ms. Individual t-tests comparing the bias for each bin against zero revealed significant visual biases from bin 500–600 to bin 700–800, and significant semantic biases in bin 500–600 (again Bonferroni corrected for

the number of bins up to the one including the average response time, $a = 0.05/13 = 0.0038$).

To assess these patterns more formally, we applied a cubic spline interpolation (Matlab R2014a) to fit a curve through the data points of the semantic and visual difference scores, for each participant separately. This allowed us to estimate the average time point at which the biases were the strongest (t_{peak}) and their strength at that time point (peak amplitude), within a period of 150 ms after visual display onset (as eye movements below 150 ms were expected not to be meaningful) and the average response time for that particular participant in each condition separately. A repeated measures ANOVA on t_{peak} with Task Relevance (template/accessory) and Object Type (semantically/visually related) as factors revealed a main effect of Object Type, $F(1,23) = 6.17$, $p < .05$, $\eta_p^2 = 0.212$, which was modulated by a significant interaction, $F(1,23) = 14.21$, $p < .01$, $\eta_p^2 = 0.382$. Follow-up t-tests showed that the visual bias peaked significantly earlier than the semantic bias in the template condition, 556 ms ($SD = 149$) vs. 784 ms ($SD = 278$),[2] $t(23) = 4.89$, $p < .001$, $r = 0.714$, but not so in the accessory condition, 647 ms ($SD = 240$) vs. 647 ms ($SD = 286$), $t(23) = 0.007$, $p = .994$. Also, semantic biases peaked earlier in the accessory than in the template condition, $t(23) = 2.752$, $p < .05$, $r = 0.498$, where the t_{peak} of the visual biases did not differ, $t(23) = 1.58$, $p = .128$. The same ANOVA for peak amplitude revealed effects of Task Relevance, $F(1,23) = 76.38$, $p < .001$, $\eta_p^2 = 0.769$, Object Type, $F(1,23) = 26.93$, $p < .001$, $\eta_p^2 = 0.539$, and again a highly reliable interaction, $F(1,23) = 18.94$, $p < .001$, $\eta_p^2 = 0.452$. Subsequent paired t-tests confirmed that both the semantic and visual biases were weaker in the accessory than in the template condition, $t(23) = 4.52$, $p < .001$ and $t(23) = 8.07$, $p < .001$, respectively. More importantly, they revealed that in the template condition the visual bias was reliably stronger than the semantic bias, 0.263 ($SD = 0.086$) vs. 0.158 ($SD = 0.057$), $t(23) = 5.21$, $p < .001$, $r = 0.736$, but that this was not the case in the accessory condition, 0.102 ($SD = 0.045$) vs. 0.096 ($SD = 0.047$), $t(23) = 0.68$, $p = .503$.
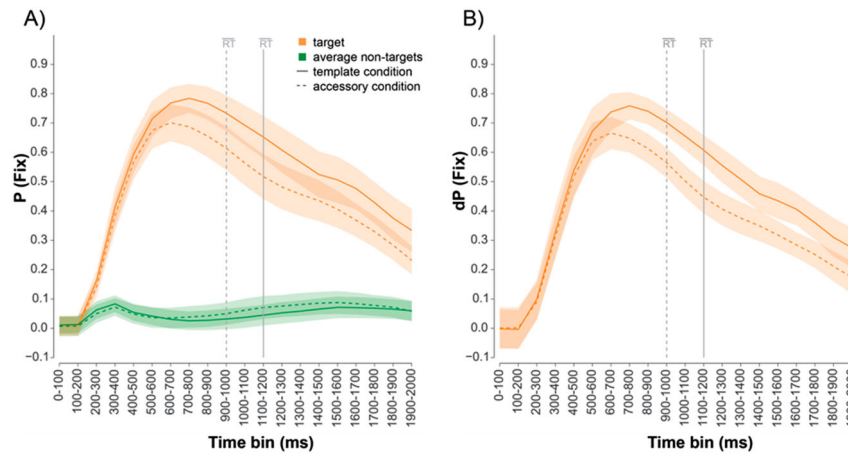
Thus, biases were reduced when the word was irrelevant for the search compared to a condition where the word was relevant. At the same time, there was still evidence, albeit relatively weak, for visual and semantic biases even in the accessory condition. To provide further evidence that the biases in this condition were not just random fluctuations, we conducted an additional correlation analysis (see also de Groot et al., 2016b). If the observed biases are non-random, they are likely to be driven by partly the same mechanisms in the template and accessory conditions. If so, the dynamics of the biases in the template condition should be predictive of the biases in the accessory condition. To test this, we correlated the biases between the template and the accessory conditions across time bins for the two types of bias. To test the reliability of these correlations, we estimated a distribution using a bootstrapping procedure, which randomly resampled the time bin data 10,000 times, with participant as index. The analyses were confined to the time period from 0 to 1300 ms after picture onset (bins 1–13), as the average response for the slowest condition (i.e., the template condition) fell in bin 13, and we did not consider any post-response biases as a priori meaningful. For each bootstrap sample, the correlation across time bins between the sample average time series data for the template and accessory conditions was computed, using Pearson correlations (r). The ensuing distributions of r-values were Fisher Z-transformed (using hyperbolic arctangent transformation) to correct for skewedness. From this transformed distribution, the two-tailed 95% confidence intervals were computed, which were inverse transformed back to the original r-space (-1 to 1). We report the median r together with these confidence intervals. These analyses showed that the time course of the visual bias in the template condition was indeed predictive of the visual bias in the accessory condition, $r = 0.66$ (CI: 0.20; 0.89). Similarly, the time course of the semantic bias in the template condition was moderately predictive of the same bias in the accessory condition, $r = 0.50$ (CI: -0.03; 0.81), although here confidence intervals just failed to exclude 0. Furthermore, within the accessory condition itself, the semantic bias was predicted by the visual bias, $r = 0.66$ (CI: 0.08; 0.91). These results suggest that the time course of modulations in accessory condition carried non-coincidental information on the visual and semantic relationships between the word and the pictures.

Target present trials

For the sake of completeness, Figure 3A shows the proportion fixation time towards the target and the

Figure 3. Target present trials. A) The proportion fixation time, P(fix), towards the target and the average of the non-targets for every 100 ms time bin. B) The difference scores, dP(fix), for the target relative to the average of the non-targets. The template condition is indicated with a solid line, whereas the accessory condition is indicated with a dotted line. The grey vertical lines mark the average search times for each condition. Error bars reflect 95% confidence intervals (two-tailed) for within-participants designs (Cousineau, 2005; Morey, 2008).

average of the non-targets as a function of time for the template (solid lines) and accessory (dotted lines) condition. Figure 3B depicts the difference in proportion fixation time between the target and the average of the non-targets. In both conditions there are clear orienting biases towards the target, although somewhat less strong for the accessory than for the template condition.

Additional eye movement measures: attentional capture and delayed disengagement

Table 1 lists, for the target absent trials for the time window from 150 ms after search display onset until the response time of each trial, the proportion of attracted fixations (i.e., the proportion of fixations that went towards a specific object from elsewhere at any time during this window), average fixation duration (i.e., how long did each individual fixation on an object last), and the average length of the fixation sequence (i.e., on average how many fixations in a row were made on an object). Whereas the first can be considered as a measure of attentional capture,

the latter two can be seen as measures of delayed disengagement. A repeated measures ANOVA on proportion of attracted fixations, with Task Relevance (template/accessory) and Object Type (semantically related/visually related/average unrelated) as factors revealed an effect of Task Relevance, $F(1,23) = 21.29$, $p < .001$, $\eta_p^2 = 0.48$, Object Type, $F(2,46) = 33.38$, $p < .001$, $\eta_p^2 = 0.592$, and a significant interaction, $F(2,46) = 32.59$. $p < .001$, $\eta_p^2 = 0.586$. Subsequent t-tests showed that in the template condition the proportion of attracted fixations was higher for the visually and semantically related objects than for the average of the unrelated objects, $t(23) = 8.35$, $p < .001$, $r = 0.867$ and $t(23) = 3.45$, $p < .01$, $r = 0.584$, respectively. Additionally, the proportion of attracted fixations was also higher for visually than for semantically related objects, $t(23) = 8.99$, $p < .001$, $r = 0.882$. For the accessory condition there was no significant difference between the different object types, ts ≤ 1.608, ps $\geq .121$. The same analysis performed on fixation duration revealed no reliable effects, Fs ≤ 0.802, ps ≥ 0.400. Finally, the same analysis on the average length of the fixation sequence showed main effects

Table 1. Averages and standard deviations (in parentheses) for several eye movements measures of attentional capture and delayed disengagement on target absent trials for each condition and object type separately.

Eye movement measures	Template condition			Accessory condition		
	Visually related	Semantically related	Average unrelated	Visually related	Semantically related	Average unrelated
Proportion attracted fixations	0.30 (0.04)	0.27 (0.02)	0.21 (0.03)	0.26 (0.03)	0.26 (0.03)	0.24 (0.01)
Average fixation duration (ms)	226 (45)	232 (51)	226 (41)	229 (49)	238 (60)	231 (47)
Average fixation sequence length	1.39 (0.16)	1.20 (0.12)	1.10 (0.09)	1.15 (0.13)	1.14 (0.10)	1.10 (0.08)

of Task Relevance, $F(1,23) = 64.18$, $p < .001$, $\eta_p^2 = 0.736$, and Object Type, $F(1.602,36.845) = 64.72$, $p < .001$, $\eta_p^2 = 0.738$, plus an interaction, $F(2,46) = 41.34$, $p < .001$, $\eta_p^2 = 0.643$. Subsequent paired t-tests showed that for both conditions the fixation sequence was longer for semantically and visually related objects than for the unrelated objects (template condition: $t(23) = 5.41$, $p < .001$, $r = 0.748$ and $t(23) = 11.61$, $p < .001$, $r = 0.924$ and accessory condition: $t(23) = 2.69$, $p < .05$, $r = 0.489$ and $t(23) = 2.86$, $p < .01$, $r = 0.512$, respectively, with the last two t-tests arriving at respectively $p = .08$ and $p = .05$ under Bonferroni correction. But only in the template condition was the fixation sequence longer for the visually than for the semantically related picture, $t(23) = 7.63$, $p < .001$, $r = 0.847$. Thus, in the template condition the visual and the semantic biases were caused by both attentional capture and delayed disengagement, while the accessory condition there was, if anything, predominantly an effect on disengagement.

Discussion

Previous research has shown that visual attention can be driven by both visual and semantic representations activated by spoken words, resulting in orienting biases. Here we assessed to what extent such biases are subject to task requirements, or are driven automatically by the maintained memory representations.

In the condition where the spoken word was relevant for search (the *template* condition) we observed strong visual and semantic biases, consistent with our earlier work (de Groot et al., 2016a). Consistent with our earlier work, we found that the visual biases arose earlier, and more strongly, than the semantic biases, indicating that the top-down guidance of attention occurred foremost, though not solely, on the basis of visual representations. The fact that semantic biases occurred later is furthermore predicted by the cascaded activation model of Huettig and McQueen (2007). First, at the auditory input end, the word activates both associated visual and semantic representations – here before the appearance of the visual objects. Then, at the visual input end, representations of visual properties such as shape and colour will be activated first, before any associated semantic representations. However, the cascaded activation model does not specify whether or how task

demands affect such biases and whether observers can put more weight on one type of information than the other.

The condition where the word was irrelevant (the *accessory* condition) becomes informative. We found that in this condition both the visual and the semantic biases were substantially reduced. Moreover, visual biases no longer prevailed over semantic biases. This implies that the larger biases in the template condition, plus the dominance of the visual over the semantic bias, were largely caused by task requirements: the word was directly relevant for the search. Thus, when observers actively look for an object, they rely relatively more on visual than on semantic representations.

Although, in the accessory condition, the word was irrelevant for the search, it was still relevant for the task as a whole, since it still needed to be remembered for the memory test at the end of the trial. As such, the results are compatible with the idea that working memory can be divided into representations that are relevant for the current task versus representations that are still maintained, but for a prospective task, in a state detached from the current perceptual task (Olivers et al., 2011). A number of authors recently proposed that memories that need to be carried across an interfering stimulus or task (as in our current accessory condition) will be represented more categorically/semantically, rather than at a detailed sensory level (Bae & Luck, 2016; Christophel et al., 2017). If this was the case here too, we might have expected semantic biases to become stronger than visual biases in the accessory condition. However, even though, relative to the visual bias, the semantic bias showed a smaller reduction when becoming temporarily task-irrelevant, it turned out no stronger than the visual bias. One question for the future is therefore whether the recruitment of semantic representations over visual representations can be boosted by changing the task requirements further, for example by asking observers to make a categorical decision on the memorized word. In any case, the reduction in both the absolute strength and relative difference when the word is irrelevant for search demonstrates that the weight with which an active memory representation drives attention varies with task demands, pointing to an active role for working memory in language-induced attentional biases (Huettig, Olivers, & Hartsuiker, 2011).

The conclusion that visual and semantic biases in visual search are subject to task demands is consistent with an earlier study from Glyn Humphreys' lab (Belke et al., 2008). They found increased interference from semantically related distractors during visual search under conditions where a high working memory load was imposed compared to conditions with a low working memory load (i.e., additionally remember a string of five numbers vs. one number). Belke et al. argued that increased cognitive load leads to a reduced target weighting, and hence longer lingering on the related distractor. A similar pattern was reported recently by Walenchok, Hout, and Goldinger (2016), who found that search suffered from interference from distractors that were phonologically similar to the verbally instructed targets (e.g., a picture of a beaver when observers were instructed to look for a "beaker"), but only when observers had to search for multiple targets at the same time (not when they looked for just a single target). No interference was found for when instructions were pictorial rather than verbal. Walenchok et al. argued that, with increased target sets and verbal instructions, observers may revert to broader levels of representation, rather than deploying a specific visual template. Furthermore, Telling, Meyer, and Humphreys (2010) found that frontal lobe patients were more disrupted by semantically related distractors, to the extent that they sometimes even responded to these objects rather than the search target. This is consistent with the strong role of the frontal lobes in cognitive control. Taken together, these studies thus support the idea that not only visual but also semantic biases in attention are subject to task control.

We believe the results are important for computational models of top-down driven visual orienting. On the one hand, within the field of psycholinguistics, a number of models have been developed to simulate language-induced phonological, semantic, and/or visual orienting. Most of these use a connectionist architecture, and involve automatic activation of associated representations (Mayberry et al., 2009; Mirman & Magnuson, 2009; Smith et al., 2013; Spivey, 2008). Although these models account for the influence of different levels of representation on visual orienting, in their current form they appear unable to accommodate task demands. On the other hand, within the field of visual attention, there are models of visual search that assume that top-down

guidance of attention changes with task demands (e.g., Navalpakkam & Itti, 2005; Wolfe, 2007; Zelinsky, 2008). However, these models in turn do not include object semantics as information that is available for attentional guidance, nor do they have the connectionist network architecture that would facilitate the spread of activation between associated representations. What is needed is a type of model that can use multiple levels of representation for guiding attention (including semantic), and also selectively weigh those levels according to task demands.

Importantly, task demands also influenced which type of associated representation was being activated the strongest. The visual bias arose earlier in time and was stronger than the semantic bias when the word was relevant for search, but this visual dominance no longer held when the word was irrelevant for search. It makes sense to have a stronger visual representation when the word is relevant for search, as visual search has been shown to be more efficient on the basis of visual representations compared to semantic representations (e.g., Schmidt & Zelinsky, 2009, 2011; Wolfe, Horowitz, Kenner, Hyle, & Vasan, 2004). A remaining question for future studies is whether working memory enhances such language-induced visual representations when task-relevant, or suppresses such activation when irrelevant, or both (McQueen & Huettig, 2014).

While the data of the current study indicate a strong influence of task requirements on the occurrence of word-driven visual and semantic biases, they do not exclude an additional automatic component, as we also found evidence for such biases when the word was irrelevant for the search (in the accessory condition). Observers showed a small but reliable preference for visual and semantic matches, especially around the time when in the task-relevant (template) condition such biases were very strong, and the dynamics of the modulation across time in the template condition was predictive of the dynamics in the accessory condition. Although we believe this provides evidence for a bias also in the accessory condition, a number of issues preclude too strong conclusions. First, one might argue that even these biases could have been driven by an active task set, if observers accidentally confused the tasks and occasionally searched for the referred to object even in the accessory condition. Thus, the difference between template and accessory conditions could

then reflect a difference in degree of template activation, or a handful of trials in which the template is fully activated, rather than a qualitative difference. Although we do not exclude a difference in terms of degree of activation, we think that it is unlikely that the bias in the accessory condition is caused by observers actively searching for the mentioned object. If so, we would have expected the visual bias to exceed the semantic bias (resembling the pattern in the template condition), which was not the case. Second, one might argue that biases had a reduced chance of developing because responses were overall faster in the accessory condition than in the template condition, due to the consistent and repeated target category (figurines) in the search task. However, although search RTs on target absent trials were indeed about 150 ms faster in the accessory condition than in the template condition, any modulation of visual and semantic biases clearly occurred well before the response (see Figure 2D). Finally, one potential reason why in the accessory condition the biases were small may be that the word onset preceded the visual search display by 2000 ms. Hearing the word may initially send an automatic surge of activation through associated visual and semantic representations even in the accessory condition, but by the time the visual search display appears the activity is reduced and maintained at levels sufficient for the memory task. Under this scenario, both automatic and controlled processes could contribute, but at different time points. To investigate this further, future studies would have to systematically vary the SOA between the word and the visual display.

Note that visual and semantic biases have also been found in passive viewing studies where observers have no task other than to look and listen (Huettig, Rommers, et al., 2011), further supporting some automaticity to the process. A study by Soto and Humphreys (2007) also indicates an automatic component to attentional guidance by active working memory representations. In their study, observers memorized either a coloured shape or a word describing that coloured shape (e.g., "red square"). When the actual coloured shape returned as a distractor in the search display that was presented during retention, search times increased, suggesting that irrelevant items in working memory can automatically attract attention. We note though that in their study the memory item and the distractor in

the search display matched on *both* visual and semantic levels of representation, and it therefore remains unclear which type of match is responsible for the bias. Second, the relevance of the memory for the search was not manipulated (i.e., the memorized item was always irrelevant). Our study allowed us to look at the influence of visual and semantic representations separately and directly compared conditions where the memorized word was either task relevant or task irrelevant.

In sum, task demands clearly modulated both the strength and the relative balance of semantic and visual orienting biases. Language-induced attentional biases are thus not simply caused by an automatic cascade of activation, but are subject to task-specific priority settings.

Notes

1. Not all visual search models completely exclude meaning. Some allow for the influence of spatial statistics associated with the context provided by entire scenes. For example, when looking for a person in an urban scene, attention may be biased to the middle areas of the picture, or areas where one is likely to find pavements. However, such overall statistical spatial biases are different from the biases we refer to here, which are driven by individual object meaning.
2. Note that estimated peak times from the curve fitting deviate somewhat from those suggested by the graphs. A closer examination of the individual fits revealed that some participants showed multiple peaks, some of which had a later maximum than the overall group.

Acknowledgements

Disclosure Statement

No potential conflict of interest was reported by the authors.

Funding

References

Bae, G. Y., & Luck, S. J. (2016). *Two ways to remember: Properties of visual representations in active and passive working memory*. Paper presented at the VSS annual meeting, St. Pete Beach, FL.

Belke, E., Humphreys, G. W., Watson, D. G., Meyer, A. S., & Telling, A. L. (2008). Top-down effects of semantic knowledge in visual search are modulated by cognitive but not perceptual load. *Perception & Psychophysics*, 70(8), 1444–1458.

Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R., & Haynes, J. D. (2017). The distributed nature of working memory. *Trends in Cognitive Sciences*, 21(2), 111–124.

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(1), 84–107. doi:10.1016/0010-0285(74)90005-x

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, 1(1), 42–45.

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(3), 453–459. doi:10.3758/BF03193787

de Groot, F., Huettig, F., & Olivers, C. N. L. (2016a). When meaning matters: The temporal dynamics of semantic influences on visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 42(2), 180–196. doi:10.1037/xhp0000102

de Groot, F., Huettig, F., & Olivers, C. N. L. (2016b). Revisiting the looking at nothing phenomenon: Visual and semantic biases in memory search. *Visual Cognition*, 24(3), 226–245. doi:10.1080/13506285.2016.1221013

de Groot, F., Koelewijn, T., Huettig, F., & Olivers, C. N. L. (2016). A stimulus set of words and pictures matched for visual and semantic similarity. *Journal of Cognitive Psychology*, 28(1), 1–15. doi:10.1080/20445911.2015.1101119

Gunseli, E., Olivers, C. N. L., & Meeter, M. (2016). Task-irrelevant memories rapidly gain attentional control with learning. *Journal of Experimental Psychology: Human Perception and Performance*, 42(3), 354–362.

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(1), B23–B32. doi:10.1016/j.cognition.2004.10.003

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(8), 985–1018. doi:10.1080/13506280601130875

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*, 57(4), 460–482. doi:10.1016/j.jml.2007.02.001

Huettig, F., Olivers, C. N. L., & Hartsuiker, R. J. (2011). Looking, language, and memory: Bridging research from the visual world and visual search paradigms. *Acta Psychologica*, 137(2), 138–150. doi:10.1016/j.actpsy.2010.07.013

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, 137(2), 151–171. doi:10.1016/j.actpsy.2010.11.003

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95(4), 492–527.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). Opensesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314–324. doi:10.3758/s13428-011-0168-7

Mayberry, M. R., Crocker, M. W., & Knoeferle, P. (2009). Learning to attend: A connectionist model of situated language comprehension. *Cognitive Science*, 33, 449–496.

McQueen, J. M., & Huettig, F. (2014). Interference of spoken word recognition through phonological priming from visual objects and printed words. *Attention, Perception, & Psychophysics*, 76(1), 190–200. doi:10.3758/s13414-013-0560-8

Mirman, D., & Magnuson, J. S. (2009). Dynamics of activation of semantically similar concepts during spoken word recognition. *Memory & Cognition*, 37, 1026–1039. doi:10.3758/MC.37.7.1026

Mishra, R. K., Olivers, C. N. L., & Huettig, F. (2013). Spoken language and the decision to move the eyes: To what extent are language-mediated eye movements automatic? In C. Pammi & N. Srinivasan (Eds.), *Progress in brain research* (Vol. 202, pp. 135–149). Amsterdam: Elsevier.

Moores, E., Laiti, L., & Chelazzi, L. (2003). Associative knowledge controls deployment of visual selective attention. *Nature Neuroscience*, 6(2), 182–189. doi:10.1038/nn996

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, 132, 297–326.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, 4(2), 61–64.

Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Research*, 45(2), 205–231.

Olivers, C. N. L. (2009). What drives memory-driven attentional capture? The effects of memory type, display type, and search type. *Journal of Experimental Psychology: Human Perception and Performance*, 35(5), 1275–1291. doi:10.1037/a0013896

Olivers, C. N. L., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences*, 15(7), 327–334. doi:10.1016/j.tics.2011.05.004

Rommers, J., Meyer, A. S., Praamstra, P., & Huettig, F. (2013). The contents of predictions in sentence comprehension: Activation of the shape of objects before they are referred to. *Neuropsychologia*, 51(3), 437–447. doi:10.1016/j.neuropsychologia.2012.12.002

Salverda, A. P., & Altmann, G. T. M. (2011). Attentional capture of objects referred to by spoken language. *Journal of Experimental Psychology: Human Perception and Performance*, 37(4), 1122–1133. doi:10.1037/a0023101

Schmidt, J., & Zelinsky, G. J. (2009). Search guidance is proportional to the categorical specificity of a target cue. *The*

Quarterly Journal of Experimental Psychology, 62(10), 1904–1914. doi:10.1080/17470210902853530

Schmidt, J., & Zelinsky, G. J. (2011). Visual search guidance is best after a short delay. *Vision Research*, 51(6), 535–545. doi:10.1016/j.visres.2011.01.013

Smith, A. C., Monaghan, P., & Huettig, F. (2013). An amodal shared resource model of language-mediated visual attention. *Frontiers in Psychology*, 4, 528. doi:10.3389/fpsyg.2013.00528

Soto, D., & Humphreys, G. W. (2007). Automatic guidance of visual attention from verbal working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 33(3), 730–737. doi:10.1037/0096-1523.33.3.730

Spivey, M. (2008). *The continuity of mind*. Vol. 40. New York, NY: Oxford University Press.

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 (5217), 1632–1634. doi:10.1126/science.7777863

Telling, A. L., Kumar, S., Meyer, A. S., & Humphreys, G. W. (2010). Electrophysiological evidence of semantic interference in visual search. *Journal of Cognitive Neuroscience*, 22(10), 2212–2225. doi:10.1162/jocn.2009.21348

Telling, A. L., Meyer, A. S., & Humphreys, G. W. (2010). Distracted by relatives: Effects of frontal lobe damage on semantic distraction. *Brain and Cognition*, 73(3), 203–214.

Walenchok, S. C., Hout, M. C., & Goldinger, S. D. (2016). Implicit object naming in visual search: Evidence from phonological competition. *Attention, Perception, & Psychophysics*, 78(8), 2633–2654.

Wolfe, J. M. (2007). Guided search 4.0: Current progress with a model of visual search. In W. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99–119). New York, NY: Oxford University Press.

Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research*, 44(12), 1411–1426. doi:10.1016/j.visres.2003.11.024

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review*, 115(4), 787–835.

Appendix A. Description of the target absent trials

The last four columns are the intended names of the pictures in Dutch (and within parentheses the English translation). The trials with an asterisk are an extension of the 100 trials of the stimulus set described in de Groot, Koelewijn, Huettig, and Olivers (2016).

Trial	Spoken word	Visual	Semantic	Unrelated1	Unrelated2
1	aardappel (potato)	bowlingbal (bowling ball)	maïskolf (corn cob)	batterij (battery)	potlood (pencil)
2*	antenne (antenna)	sigaret (cigarette)	televisie (television)	trampoline (trampoline)	kiwi (kiwi)
3*	arm (arm)	boemerang (boomerang)	hersenen (brain)	waterscooter (jet ski)	plakbandhouder (tape incl. holder)
4	asbak (ashtray)	jojo (yoyo)	pijp (pipe)	dennenappel (pinecone)	rozen (roses)
5	bad (bath tub)	slee (sleigh)	kraan (faucet)	honkbalhandschoen (baseball glove)	kwast (brush)
6	badpak (bathing suit)	kruik (hot water bottle)	slippers (flip flops)	nietjes (staples)	koffiezetapparaat (coffee maker)
7	bakblik (oven tin)	cassettebandje (cassette tape)	taart (pie)	schaats (ice skate)	stropdas (tie)
8	bal (ball)	tomaat (tomato)	voetbalschoenen (soccer cleats)	waterpijp (hookah)	schep (shovel)
9	ballon (balloon)	kers (cherry)	cadeau (present)	kaasschaaf (cheese slicer)	koffiebonen (coffee beans)
10	banaan (banana)	kano (canoe)	aap (monkey)	tamboerijn (tambourine)	hoed (hat)
11	basketbal (basketball)	kokosnoot (coconut)	badmintonracket (badminton racket)	steekwagen (handtruck)	stanleymes (box cutter)
12	beker (mug)	garen (thread)	vork (fork)	pen (pen)	duikbril (goggles)
13	blokken (blocks)	toffee (toffee)	hobbelpaard (rocking horse)	saxofoon (saxophone)	beer (bear)
14	bolhoed (bowler hat)	citruspers (juicer)	wandelstok (walking stick)	vlees (meat)	olifant (elephant)
15	boom (tree)	wc-borstel (toilet brush)	bijl (axe)	magnetron (microwave)	magneet (magnet)
16	boor (drill)	pistool (hand gun)	rolmaat (measuring tape)	ballon (balloon)	bureaustoel (office chair)
17	boot (boat)	klomp (clog)	anker (anchor)	chocolade (chocolate)	honkbal (baseball)
18	bot (bone)	halter (dumb-bell)	puppy (puppy)	bezem (broom)	narcis (narcissus)
19*	brievenbus (mailbox)	broodrooster (toaster)	postzegels (stamps)	ijslepel (ice cream scooper)	fluit (recorder)
20	bril (glasses)	bh (bra)	telescoop (telescope)	scheermes (razor)	sleutel (key)
21	buggy (buggy)	tractor (tractor)	flesje (baby bottle)	sneeuwschuiver (snow shovel)	zonnebloem (sunflower)
22	cd (cd)	reddingsboei (life saver)	diskette (floppy disk)	holster (holster)	duimstok (yard stick)
23	drol (turd)	ijsje (ice cream cone)	luier (diaper)	kan (jar)	pompoen (pumpkin)
24	druiven (grapes)	biljartballen (billiard balls)	wijnglas (wine glass)	kettingzaag (chainsaw)	bel (bell)
25	drumstel (drum kit)	weegschaal (scale)	elektrischegitaar (electric guitar)	katapult (sling shot)	speelkaarten (playing cards)
26	ei (egg)	wol (yarn)	haan (rooster)	tandenborstel (toothbrush)	xylofoon (xylophone)
27	fles (bottle)	kegel (pin)	kurk (cork)	broek (pants)	kerstbal (bauble)
28	fluit (recorder)	deegroller (rolling pin)	harp (harp)	badeend (rubber duck)	ton (barrel)
29	garde (whisk)	borstel (hair brush)	schaal (bowl)	speldenkussen (pincushion)	pillen (pills)
30	gloeilamp (light bulb)	avocado (avocado)	lichtschakelaar (light switch)	adelaar (eagle)	mand (basket)
31	handboeien (handcuffs)	trappers (pedals)	politiepet (police hat)	scheerkwast (shaving brush)	hijskraan (crane)
32	handboog (longbow)	ijzerzaag (hacksaw)	kanon (cannon)	ananas (pineapple)	nagellak (nail polish)
33*	handdoek (towel)	zonnescherm (sunshade)	bad (bath tub)	monitor (monitor)	vogelhuisje (birdhouse)
34	hark (rake)	spatel (spatula)	heggenschaar (hedge trimmer)	dynamiet (dynamite)	zwemband (inner tube)
35	helm (helmet)	mango (mango)	motor (engine)	blik (dustpan)	ijshoorntje (cone)
36*	hersenen (brains)	bloemkool (cauliflower)	neus (nose)	koekje (cookie)	nijptang (pincers)
37	hijskraan (crane)	giraf (giraffe)	cementwagen (cement truck)	kopje (cup)	bramen (blackberries)
38	Hoefijzer (horseshoe)	koptelefoon (headphones)	zadel (saddle)	teddybeer (teddy bear)	brie (brie)
39	ipod (ipod)	kompas (compass)	radio (radio)	watermeloen (watermelon)	flesopener (bottle opener)
40*	jas (coat)	tuitbeker (sippy cup)	want (mitten)	platenspeler (turntable)	snoepjes (candy)
41	jerrycan (jerry can)	paprika (bell pepper)	benzinepomp (petrol pump)	ventilator (fan)	telefoon (phone)
42*	joystick (joystick)	tol (top (toy))	toetsenbord (keyboard)	klamboe (mosquito net)	kuiken (chick)
43	kleerhanger (clothes hanger)	triangel (triangle)	kapstok (coat hanger)	luidspreker (megaphone loudspeaker)	driewieler (tricycle)
44	klokhuis (apple core)	vaas (vase)	aardbei (strawberry)	portemonnee (wallet)	hamer (hammer)
45	koekje (cookie)	pleister (band aid)	chips (potato chips)	boog (bow)	thermometer (thermometer)
46	koelkast (refrigerator)	mobiel toilet (portapotty)	ijskristal (snow flake)	skeeler (roller blade)	naald (needle)
47	koffer (suitcase)	lantaarn (lantern)	trein (train)	stoel (chair)	olijf (olive)
48	krijtjes (chalks)	spelden (pins)	palet (palette)	kikker (frog)	trommel (drum)
49	krokodil (crocodile)	augurk (pickle)	uil (owl)	bokshandschoenen (boxing gloves)	tandartsstoel (dental chair)
50	kussen (pillow)	ravioli (ravioli)	schommelstoel (rocking chair)	leeuw (lion)	asbak (ashtray)
51	lampion (Chinese lantern)	bandoneon (accordion)	zaklamp (flashlight)	peultje (sugar snap pea)	hagedis (lizard)
52	lasso (lasso)	waterslang (water hose)	cowboyhoed (cowboy hat)	stemvork (tuning fork)	tas (bag)
53	liniaal (ruler)	kam (comb)	perforator (hole puncher)	pannenkoeken (pancakes)	drinkzak (camel bag)

Continued.

Trial	Spoken word	Visual	Semantic	Unrelated1	Unrelated2
54	lippenstift (lipstick)	aansteker (lighter)	parfum (perfume)	cruiseschip (cruise ship)	zak (paper bag)
55	loep (lens)	tafeltennisbatje (ping pong paddle)	microscoop (microscope)	prullenbak (trash can)	reddingsvest (life vest)
56	medaille (medal)	bord (plate)	trofee (trophy)	garnaal (shrimp)	schroevendraaier (screwdriver)
57	meloen (melon)	rugbybal (rugby ball)	bananen (bananas)	golfclub (golf club)	raket (rocket)
58	mes (knife)	peddel (paddle)	theepot (teapot)	poederdoos (face powder box)	babybedje (play crib)
59	microfoon (microphone)	pizzasnijder (pizza cutter)	boxjes (speakers)	ketel (kettle)	vuilniszakken (garbage bags)
60*	mijter (mitre)	pylon (pylon)	staf (staff)	bergschoen (mountain boot)	fax (fax machine)
61	milkshake (milk shake)	walkietalkie (walkie talkie)	friet (french fries)	wetsuit (wet suit)	snelheidsmeter (speedometer)
62	monitor (monitor)	dienblad (tray)	muis (mouse)	notenkraker (nutcracker)	rietjes (straws)
63	naald (needle)	dwarsfluit (flute)	vingerhoedje (thimble)	fiets (bicycle)	boek (book)
64	oog (eye)	globe (globe)	haar (wig)	broccoli (broccoli)	politieauto (police car)
65	oor (ear)	croissant (croissant)	voet (foot)	schildersezel (easel)	vrachtwagen (truck)
66	oven (oven)	kastje (cabinet)	koekenpan (frying pan)	honkbalknuppel (baseball bat)	tijger (tiger)
67	pannenkoek (pancake)	klok (clock)	brood (bread)	ketting (chain)	vijl (nail file)
68	paraplu (umbrella)	krukje (stool)	regenlaarzen (rain boots)	veiligheidsspelden (safety pins)	kruiwagen (wheelbarrow)
69	piano (piano)	barcode (barcode)	trompet (trumpet)	riem (belt)	bureaulamp (desk light)
70	pinguïn (penguin)	champagne (champagne)	ijsbeer (polar bear)	tissues (tissues)	bureau (desk)
71	pinpas (debit card)	envelop (envelope)	euro (euro)	blad (leaf)	zwaan (swan)
72	plakband (scotch tape)	toiletpapier (toilet paper)	paperclip (paper clip)	pijl (arrow)	zonnebril (sunglasses)
73	plant (plant)	feesttoeter (party horn)	gieter (watering can)	nagelknipper (nail clipper)	controller (controller)
74*	portemonnee (wallet)	kussen (pillow)	geld (money)	zebra (zebra)	gong (gong)
75	potlood (pencil)	schroef (screw)	puntenslijper (pencil sharpener)	skelet (skeleton)	kat (cat)
76*	raam (window)	schilderij (painting)	schoorsteen (chimney)	vishaak (lure)	zalmmoot (salmon fillet)
77	radiator (radiator)	dranghek (fence)	kachel (heater)	boon (bean)	nietmachine (stapler)
78	raket (rocket)	vuurtoren (lighthouse)	tank (tank)	toilettas (toiletry bag)	dalmatiër (Dalmatian)
79*	rasp (grater)	wolkenkrabber (skyscraper)	kaas (cheese)	backpack (backpack)	brandweerhelm (fireman's helmet)
80	rat (rat)	stekkerdoos (extension cord)	muizenval (mousetrap)	horloge (watch)	brug (bridge)
81	riem (belt)	slang (snake)	sokken (socks)	dartbord (dartboard)	cappuccino (cappuccino)
82	ring (ring)	donut (donut)	oorbellen (earrings)	telraam (abacus)	prei (leek)
83	rog (stingray)	vliegtuig (plane)	zeepaardje (sea horse)	bierflesje (beer bottle)	discobal (disco ball)
84*	schaakbord (chessboard)	theedoek (dishcloth)	dobbelstenen (dice)	mixer (mixer)	bloempot (flower pot)
85*	scheermes (razor)	fietspomp (bicycle pump)	zeeppompje (soap dispenser)	piramide (pyramid)	tram (tram)
86	schildpad (tortoise)	noot (nut)	viskom (fishbowl)	vaatwasser (dishwasher)	winkelwagen (shopping cart)
87	schoen (shoe)	strijkijzer (iron)	pet (baseball cap)	propeller (propeller)	pakket (packet)
88*	schoorsteen (chimney)	trechter (funnel)	dak (roof)	dubbeldekker (double decker bus)	peer (pear)
89	shuttle (badminton birdie)	gloeilamp (light bulb)	tennisbal (tennis ball)	pasta (pasta)	dunschiller (potato peeler)
90	sinaasappel (orange)	golfbal (golf ball)	courgette (zucchini)	kalf (calf)	snijplank (cutting board)
91	ski's (skis)	pincet (tweezers)	muts (beanies)	ezel (donkey)	peper (pepper)
92	sleutel (key)	kurkentrekker (corkscrew)	kluis (safe)	basketbal (basketball)	spinnewiel (spinning wheel)
93	slof (slipper)	cavia (guinea pig)	badjas (bathrobe)	filmrol (film)	strijkplank (ironing board)
94	snijplank (cutting board)	laptop (laptop)	hakmes (cleaver)	kerstkrans (Christmas wreath)	jas (jacket)
95	snoep (candy)	knikkers (marbles)	hamburger (hamburger)	wasmachine (washing machine)	fototoestel (camera)
96	spaghetti (spaghetti)	touw (rope)	vergiet (colander)	verkeerslicht (traffic light)	klarinet (clarinet)
97	speen (pacifier)	pion (pawn)	babypakje (onesies)	picknicktafel (picnic table)	dolfijn (dolphin)
98*	sperzieboon (butter bean)	sabel (saber)	ui (onion)	spiegel (mirror)	douchekop (shower head)
99	spook (ghost)	shuttle (badminton birdie)	grafsteen (tombstone)	hondenriem (dog leash)	koffiemolen (coffee grinder)
100	spuit (injection)	Dartpijl (dart)	stethoscoop (stethoscope)	dominostenen (dominoes)	fornuis (stove)
101*	stift (pin)	pipet (pipette)	notitieboekje (notebook)	vliegenmepper (fly swatter)	kist (chest)
102	stijgbeugel (stirrup)	stamper (masher)	paard (horse)	hotdog (hot dog)	palmboom (palm tree)
103*	stopcontact (socket)	knoop (button)	stekker (plug)	sjaal (scarf)	luchtballon (hot air balloon)
104	strijkplank (ironing board)	keyboard (keyboard)	wasmand (laundry basket)	bloem (flower)	hand (hand)
105*	stropdas (tie)	vlieger (kite)	trui (sweater)	rolstoel (wheelchair)	videoband (videotape)
106	surfplank (surfboard)	veer (feather)	badpak (bathing suit)	bizon (bison)	graafmachine (excavator)
107	sushi (sushi)	duct tape (duct tape)	eetstokjes (chopsticks)	kruisboog (crossbow)	step (scooter)
108	tamboerijn (tambourine)	pizza (pizza)	viool (violin)	wattenstaafje (cotton swab)	kruk (stool)

(*Continued*)

Continued.

Trial	Spoken word	Visual	Semantic	Unrelated1	Unrelated2
109	televisie (television)	schoolbord (blackboard)	afstandsbediening (remote control)	trombone (trombone)	cowboylaarzen (cowboy boots)
110*	tent (tent)	geodriehoek (protractor)	gasflesje (camping burner)	neushoorn (rhino)	brandweerauto (fire truck)
111	theepot (teapot)	kandelaar (candle holder)	lepel (spoon)	sportschoenen (sneakers)	bretels (suspenders)
112	toffee (toffee)	vlinderdas (bow tie)	gebit (teeth)	agenda (agenda)	hout (wood)
113	trappers (pedals)	verfroller (paint roller)	wiel (wheel)	haai (shark)	glijbaan (slide)
114	visnet (fishnet)	zeef (sieve)	kreeft (lobster)	lantaarnpaal (lamp post)	scheerapparaat (electric razor)
115	vlieger (kite)	voorrangsbord (traffic sign)	springtouw (jump rope)	geweer (rifle)	printer (printer)
116	vliegtuig (airplane)	kruis (cross)	label (label)	worst (sausage)	muffin bakvorm (muffin tin)
117	vlinder (butterfly)	gereedschapskist (tool box)	rups (caterpillar)	rijst (rice)	slot (lock)
118*	wortel (carrot)	schelp (shell)	appel (apple)	usb-stick (USB stick)	frisbee (frisbee)
119	zaklamp (flashlight)	ontstopper (plunger)	kaars (candle)	ijsblokjesvorm (ice cube tray)	flippers (flippers)
120	zweep (whip)	hengel (fishing rod)	cap (derby hat)	verrekijker (binocular)	framboos (raspberry)

Appendix B. The set of plastic figurines that were used as targets on target present trials in the accessory condition.