

The nature of the visual environment induces implicit biases during language-mediated visual search

Falk Huettig · James M. McQueen

Published online: 2 April 2011
© Psychonomic Society, Inc. 2011

Abstract Four eyetracking experiments examined whether semantic and visual-shape representations are routinely retrieved from printed word displays and used during language-mediated visual search. Participants listened to sentences containing target words that were similar semantically or in shape to concepts invoked by concurrently displayed printed words. In **Experiment 1**, the displays contained semantic and shape competitors of the targets along with two unrelated words. There were significant shifts in eye gaze as targets were heard toward semantic but not toward shape competitors. In Experiments 2–4, semantic competitors were replaced with unrelated words, semantically richer sentences were presented to encourage visual imagery, or participants rated the shape similarity of the stimuli before doing the eyetracking task. In all cases, there were no immediate shifts in eye gaze to shape competitors, even though, in response to the **Experiment 1** spoken materials, participants looked to these competitors when they were presented as pictures (Huettig & McQueen, 2007). There was a late shape-competitor bias (more than

2,500 ms after target onset) in all experiments. These data show that shape information is not used in online search of printed word displays (whereas it is used with picture displays). The nature of the visual environment appears to induce implicit biases toward particular modes of processing during language-mediated visual search.

Keywords Attention · Eye movements · Semantic representations · Visual representations

When we hear or read a word and pay attention to it, we typically understand the meaning of that word if we have previously learned that meaning. The question we ask in this article is whether lexical access during language-mediated visual search leads to robust retrieval and the use of conceptual lexical representations, or whether this depends (at least partly) on the nature of the information in the visual environment.

In many situations that language users find themselves in, there is a tight coupling between visual processing and the high-level mental representations involved in memory and language. This is particularly evident from eye gaze behavior during language-mediated visual search (see Huettig, Olivers, & Hartsuiker, *in press*, and Huettig, Rommers, & Meyer, *in press*, for reviews; see also Cooper, 1974; Desimone & Duncan, 1995; Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy 1995; Wolfe, 1994). Cooper showed that participants, when presented with a visual display containing a number of common objects, tended to spontaneously fixate the visual referents of words that were heard concurrently. For instance, participants were more likely to fixate the picture of a lion when hearing “lion” (or even part of “lion”) than to fixate pictures of referents of unrelated control words.

This mapping behavior could reflect processing at one or more levels of representation. One possibility is that it reflects a phonological match (Alloppenna, Magnuson & Tanenhaus

F. Huettig (✉) · J. M. McQueen
Max Planck Institute for Psycholinguistics,
P.O. Box 310, Nijmegen 6500 AH, The Netherlands
e-mail: Falk.Huettig@mpi.nl

J. M. McQueen
e-mail: James.McQueen@mpi.nl

F. Huettig · J. M. McQueen
Donders Institute for Brain, Cognition, and Behavior,
Radboud University Nijmegen,
Kapittelweg 29,
6500 HB Nijmegen, The Netherlands

J. M. McQueen
Behavioural Science Institute, Radboud University Nijmegen,
Montessorilaan 3, 65,
Nijmegen 6525 HR, The Netherlands

1998). To be able to recognize the token of *lion* as an instance of the lexical type “lion,” the participant must retrieve stored knowledge about the phonological form of the word from long-term memory. The participant could also retrieve stored phonological knowledge on the basis of the information in the visual display (i.e., retrieval of the name of the pictured entity; see also Meyer, Belke, Telling & Humphreys 2007; Meyer & Damian, 2007). Fixations of the picture of a lion could thus reflect detection of a match at the phonological level between knowledge retrieved from the speech input and that retrieved from the visual input.

But this behavior could also reflect a match at the conceptual/semantic level. A match could be detected between knowledge about the concept associated with the spoken word “lion” and knowledge about the concept associated with the picture of the lion (Huettig & Altmann, 2005; Huettig, Quinlan, McDonald & Altmann 2006; Yee & Sedivy, 2006). Likewise, it could reflect a match at the level of visual shape, between knowledge about shape retrieved from the linguistic signal and knowledge retrieved during the process of recognizing the picture as being that of a lion (Dahan & Tanenhaus, 2005; Huettig & Altmann, 2004, 2007). Indeed, previous research has shown that with picture displays, fixations are determined by matches between knowledge retrieved on the basis of information in the linguistic and in the visual input at all three of these levels of representation. Shifts in overt attention are co-determined by the timing of cascaded processing in the word and object recognition systems and by the temporal unfolding of the spoken language (Huettig & McQueen, 2007). The present research investigates whether these different types of language-derived and display-derived representations are always accessed and used when someone is confronted simultaneously with speech and visual input, or whether the nature of the information in the visual environment influences this mapping behavior.

The data of Huettig and McQueen (2007) are of particular relevance for this issue. In four experiments, Dutch participants listened to spoken Dutch sentences (which included a critical word) while looking at visual displays containing four spatially distinct visual items. Each critical word was placed in a neutral sentence (e.g., for *beker* “beaker,” *Uiteindelijk keek ze naar de beker die voor haar stond* “Eventually she looked at the beaker that was in front of her”). The critical words were not predictable in these contexts. Even though the spoken sentences were identical across all four experiments, and even though the visual displays had exactly the same abstract content in each experiment (visual objects or their printed names), the eye movement behavior was radically different, in terms of both where participants looked and when they looked, across the four experiments. All that changed across the experiments was the relative timing of presentation of the linguistic and visual information and whether the displays consisted of pictures or printed words.

When participants had time to look at a display of four visual objects from the onset of the sentences (Huettig & McQueen, 2007, Exp. 1), attentional shifts to phonological competitors of the critical spoken words (e.g., to a beaver, *bever* in Dutch, on hearing *beker*) preceded attentional shifts to shape competitors (e.g., a bobbin) and semantic competitors (e.g., a fork). With only 200 ms of preview of the same picture displays prior to onset of the critical word (Exp. 2), participants did not look preferentially at the phonological competitors, but instead made more fixations to the shape competitors and then the semantic competitors. In other words, when there was ample time to view the display (Exp. 1), it appears that picture processing did advance as far as retrieval of the pictures’ names: There were fixations to all three types of competitors. But when there was only 200 ms of preview before the onset of the critical spoken word (Exp. 2), picture processing still involved retrieval of visual and semantic features to a degree sufficient to influence eye movements, but there was insufficient retrieval of the pictures’ names to influence behavior. Huettig and McQueen (2007) suggested that there were no preferential fixations to the phonological competitors under these conditions because, by the time a picture’s name could have been retrieved, the evidence in the speech signal had already indicated that that phonological competitor was not a part of the sentence. A third pattern was found when the pictures were replaced with printed words (the names of the same entities as before; cf. McQueen & Viebahn, 2007). Now attentional shifts were made only to the phonological competitors, both when there was only 200 ms of preview (Exp. 3) and when the displays appeared at sentence onset (Exp. 4).

One important conclusion from these data is that the storage and/or retrieval of phonological knowledge is independent of the storage/retrieval of conceptual knowledge (see also Norris, Cutler, McQueen & Butterfield 2006). If lexical knowledge were accessed in an all-or-none manner, such that retrieval of a word’s phonological form necessarily entailed retrieval of all other knowledge of a word, including visual and semantic features, then there could be no difference in the time courses of looks to the phonological and other types of competitors. The implication of this finding is that questions about the retrieval and use of conceptual/semantic knowledge are independent of similar questions concerning phonological knowledge. Moreover, these results also suggest that the nature of the visual stimuli (e.g., pictures or printed words) induces implicit biases toward particular modes of processing during language-mediated visual search.

The results of Huettig and McQueen (2007, Exp. 1; i.e., matches at visual-shape and semantic levels of processing during language-mediated visual search with picture displays) are in line with a wealth of evidence that we cannot switch off processes of visual object recognition and

conceptual processing of individual visual objects (e.g., Intraub, 1984; Loftus & Ginn, 1984; Potter, 1976; Smith & McGee, 1980) and that we cannot prevent the interpretation of meaningful visual scenes (e.g., Biederman, Mezzanotte & Rabinowitz 1982) or of displays of individual objects, even in the absence of a coherent visual scene (e.g., Biederman, Blickle, Teitelbaum & Klatsky 1988).

It is also uncontroversial that there is particularly easy access to phonological forms given orthographic input (Van Orden, Johnston & Hale 1988; see also Frost, 1998). Contrary to visual object processing, there is some evidence that conceptual retrieval during written word recognition is under strategic control. Stolz and Besner (1999), for example, argued that semantic processing depends on attentional control over how processing is distributed across different levels of representation. Stolz and Besner (1996) presented evidence that letter search on a prime word eliminates semantic priming in the lexical decision task. Their findings seem to suggest that letter search is under attentional control at a time when retrieval of the target word's meaning through priming would normally occur. They concluded that semantic activation is not an automatic product of a written word's presentation but that attentional control determines whether activation in the word recognition system will spread to the semantic level.

Here we sought to establish whether semantic and (stored) visual-shape representations are routinely retrieved from printed word displays and used *online* during language-mediated visual search. Huettig and McQueen's (2007) Experiments 1 and 2 using picture displays suggested that language-mediated visual attention involves multiple matches at phonological, visual feature, and semantic levels of processing. Importantly for the present study, however, search with printed word displays depended only on phonological matches. The strongest interpretation of these results is that retrieval of semantic and visual-shape knowledge is not an automatic consequence of exposure to printed words. An alternative interpretation, however, is that these knowledge types were retrieved during language-mediated visual search with printed word displays but not used, because of easier access to phonological representations given the orthographic input from the display (Van Orden et al., 1988) and/or because of access to orthographic representations given the concurrent phonological input from the spoken sentences (Salverda & Tanenhaus, 2010). The results of Huettig and McQueen thus leave open whether semantic and visual-shape knowledge can ever be used to direct visual attention around a printed word array. The critical question is thus whether the nature of the printed words and the presence of certain representational matches with the concurrent spoken words induces implicit biases during online language-mediated visual search with printed word displays.

We conducted four experiments. In Experiment 1, the displays consisted of semantic and shape competitors of the

critical spoken words, as well as two unrelated distractors. We used exactly the same spoken materials as in Huettig and McQueen (2007) and the same printed word displays as in their Experiments 3 and 4, except that the phonological competitors were replaced with a further set of unrelated distractors. If language-mediated shifts in attention contingent upon semantic and/or visual-shape information are, in principle, possible with printed word displays, then we should observe them in experimental trials in which matches of phonological information between spoken and printed words are not present. It is conceivable, however, that semantic and/or visual-shape competitors never attract increased overt attention with printed word displays—even if no other representational matches compete for attention. Such a finding, especially given that the same spoken materials and picture versions of the words in the display did lead to shifts in attention to conceptual competitors in the Huettig and McQueen study, would strongly suggest that the use of individual components of lexical knowledge is determined by the nature of the information in the visual display.

Experiment 1

Method

Participants A total of 26 members of the participant panel of the Max Planck Institute for Psycholinguistics, all native speakers of Dutch, were paid for their participation. All had normal or corrected-to-normal vision.

Stimuli Five sets of words were selected for 40 experimental trials (see Appx. A). Each set consisted of a critical spoken base word and four printed words, of which two were related words and two were unrelated distractor words. Each critical spoken word was placed in a neutral sentence context (e.g., for *beker*, *Uiteindelijk keek ze naar de beker die voor haar stond* “Eventually she looked at the beaker that was in front of her”). The critical words were not predictable on the basis of these contexts. The average acoustic onset of the critical word was 1,808 ms after sentence onset. The average target word duration was 437 ms.

The four printed words in each set were used in the visual displays. Each of these displays (see Fig. 1) contained a shape competitor word (unrelated in phonology and semantics), a semantic competitor word (unrelated in phonology and shape), and two printed words unrelated on all of these dimensions. The items were identical to those used in Huettig and McQueen (2007), except that the phonological competitors were replaced with unrelated distractors.

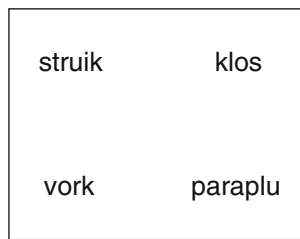


Fig. 1 Example of a visual stimulus used in Experiment 1. For the spoken sentence *Uiteindelijk keek ze naar de beaker die voor haar stond* “Eventually she looked at the **beaker** that was in front of her,” the display consisted of the printed words *klos* (“bobbin,” the visual-shape competitor) and *vork* (“fork,” a semantic competitor), as well as two unrelated distractor words (*struik* “bush,” *paraplu* “umbrella”)

A further 40 sets of four words were selected for filler trials. These sets included one target word that was placed in a neutral sentence context (like those used in experimental trials) and three unrelated distractor words. The visual display contained the printed form of the spoken target word, plus printed forms of the three distractors. Therefore, across all trials in the experiment, 50% of the 80 trials included a printed word that fully matched the spoken critical word (see Huettig & McQueen, 2007, for further details).

Rating study The present study used printed word displays only. The stimuli in Huettig and McQueen (2007), however, were normed using the visual objects rather than their printed names. It is conceivable that, when presented with the printed word stimuli, participants would not retrieve the intended visual shapes (the use of pictures in Huettig & McQueen, 2007, obviated this concern, since the forms were presented to the participants). We therefore carried out an additional rating study using the printed words rather than the visual objects. A total of 29 participants provided shape-similarity ratings. They were presented, over the Internet, with all 40 critical words, in printed form, each paired with their four printed words per trial (displayed as in Exp. 1; see Fig. 1). These participants were asked to judge how similar the physical shape of the concept of the critical word was to the physical shape of the concepts invoked by the four printed words, while ignoring any similarity in meaning, using an 11-point scale (0 representing *absolutely no similarity in physical shape* and 10 representing *identical physical shape*). The results are shown in Table 1. The shape competitors were judged to be significantly physically more similar to the critical words than to any of the other printed word stimuli (all $ps < .001$). We can thus be confident that the printed visual competitors did indeed evoke the intended default visual shapes.

Procedure The 40 experimental and 40 filler sentences were read aloud by a female native speaker of Dutch in a sound-damped booth. Digital recordings of these utterances, at a sampling rate of 44.1 kHz with 16-bit resolution, were stored directly on computer. The sentences were read with a neutral intonation contour such that, in particular, the critical words were not highlighted.

The participants were seated at a comfortable distance from the computer screen. One centimeter on the visual display corresponded to approximately 1° of visual arc. The eyetracking system was mounted and calibrated. Eye movements were monitored with an SR EyeLink eyetracking system, sampling at 250 Hz. The spoken sentences were presented to the participants through headphones. The parameters of each trial were as follows: First, a central fixation point appeared on the screen for 500 ms, followed by a blank screen for 600 ms. Then four printed words appeared on the screen as the auditory presentation of a sentence was initiated. The positions of the pictures were randomized across four fixed positions on a (virtual) 5 x 5 grid on every trial (Grid Positions 7, 9, 16, and 18, counting from left to right and from top to bottom).

Participants were asked to perform a “look-and-listen” task (Altmann & Kamide, 1999; Huettig & Altmann, 2005). They were told that they should listen to the sentences carefully, that they could look at whatever they wanted to, but that they should not take their eyes off the screen throughout the experiment (see Huettig & McQueen, 2007, for discussion). The participants’ fixations for the entire trial were thus completely unconstrained, and the participants were under no time pressure to perform any action.

Each participant was presented with all 80 trials. The experimental and filler trials were presented in random order. A central fixation point appeared on the screen after every five trials, allowing for drift correction in the calibration.

Data coding procedure The data from each participant’s right eye were analyzed and coded in terms of fixations,

Table 1 Results of the similarity-rating study: Means (and standard deviations in brackets) for the rated shape similarity between the critical words and each type of printed word

	Shape Competitors	Semantic Competitors	Unrelated Distractor 1	Unrelated Distractor 2
Shape similarity				
Mean rating (SD)	4.84 (1.69)	1.82 (0.81)	1.51 (0.75)	1.27 (0.88)

saccades, and blinks, using the algorithm provided in the EyeLink software. The timing of the fixations was established relative to the onset of the critical word in the spoken utterance. The visual displays were treated as being composed of four quadrants, and gaze position was categorized by quadrant. Fixations were coded as directed to the semantic competitor word, the shape competitor word, or to the unrelated distractor words.

Results

Figure 2 shows a time-course graph that illustrates the fixation proportions at 20-ms intervals to the various types of printed words over the course of the average experimental trial. On the y -axis, $p(\text{fixation})$ is the probability of a fixation to each of the types of onscreen words, and was computed by counting, over all participants and items, the proportion of fixations to each of the word types in a given 20-ms time window, relative to the total number of fixations in that time window. Zero represents the acoustic onset of the spoken critical word.

For the statistical analyses, we computed mean fixation proportions for each type of printed word. We calculated the ratio between the proportion of fixations to a particular competitor (semantic or shape) and the sum of the particular competitor- and distractor-fixation proportions (see Huettig & McQueen, 2007). A ratio greater than .5 shows that, of all the fixations directed toward a particular type of competitor and the unrelated

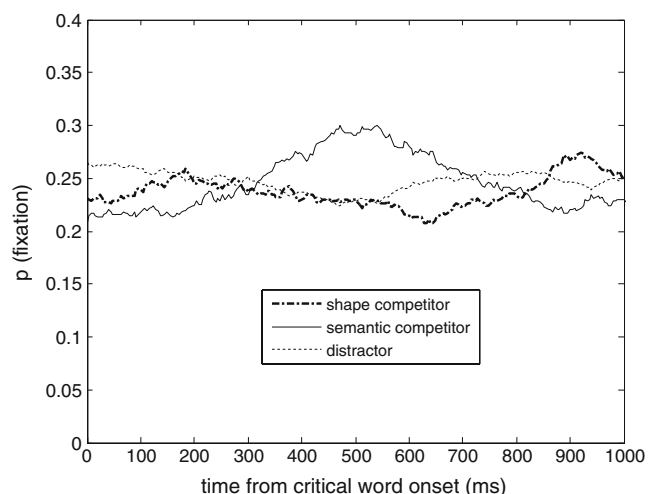


Fig. 2 Time-course graph showing fixation probabilities to semantic competitors, shape competitors, and unrelated distractors during the critical time period in Experiment 1. The x -axis shows time in milliseconds from the acoustic onset of the critical spoken word, for 1,000 ms

distractors, the competitors attracted more than half of those fixations. We computed mean ratios per participant and item over a time interval starting from the acoustic onset of the critical spoken word to 200 ms after this onset (as an estimate of the earliest point in time at which a fixation could reflect a response based on information in the critical word; cf. Matin, Shao & Boff 1993; Saslow, 1967), in order to obtain a baseline of competitor/distractor fixation ratios per type of printed word. We can assume that fixations during this 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 printed word *before* information from the critical word became available (indeed, Fig. 2 and Table 2 show that in Exp. 1 there was a slight initial bias in looks toward the unrelated distractors before the critical word had been heard). Calculating a mean ratio for the baseline time regions (and then comparing these ratios with the mean competitor/distractor ratios during later time regions) allows us to test for any shifts in overt attention to particular types of printed words during the time of interest. Thus, we also computed mean ratios per participant and item for two subsequent 200-ms time regions: from 201 to 400 ms and from 401 to 600 ms after the acoustic onset of the critical spoken word, a time period that (taking into account the time typically considered necessary to program an eye movement) reflects the approximate acoustic lifetimes of our critical spoken words (recall that the mean critical word duration was 437 ms). Table 2 shows that the mean ratios increased greatly over time for the semantic competitors but stayed approximately the same for the shape competitors. We tested whether the competitor/distractor ratios during the baseline time window were significantly different from the competitor/distractor ratios during the subsequent time windows.

Paired t tests showed that the mean semantic-competitor/distractor ratios during the baseline region differed significantly [$t_1(25) = 2.83, p = .009; t_2(39) = 3.89, p < .001$] from the mean ratios during the 401- to 600-ms time region. The mean shape-competitor/distractor ratios during the baseline region did not differ significantly [$t_1(25) = 0.97, p > .1; t_2(39) = 0.57, p > .1$] from the mean ratios during the 401- to 600-ms time region. We then divided the shape competitors on the basis of the shape ratings collected in the rating study into a set of 20 high-similarity items (mean shape similarity rating: 6.20, $SD = 0.66$) and a set of 20 low-similarity items (mean shape similarity rating: 3.65, $SD = 0.93$), and reran the shape-competitor/distractor analyses separately for each item set. Both of these analyses gave the same results as

Table 2 Competitor/distractor ratios in [Experiment 1](#)

Time Region (milliseconds from acoustic onset of critical word)	Competitor/Distractor Ratio	
	Semantic	Shape
0–200	.45	.47
201–400	.50	.49
401–600	.55	.49

the overall analysis (all $ps > .05$). Similar analyses during the 200- to 400-ms time window yielded no significant differences for the experimental items. During the filler trials (in which a printed word fully matched the spoken critical word), participants strongly shifted their eye gaze toward the matching printed word (e.g., 35% of fixation proportions during the 401- to 600-ms time region; chance level = 25%). This shows that participants engaged in the expected fixation behavior during the experiment.

Thus, in [Experiment 1](#) there were no preferential fixations toward the shape competitors during the time of interest (i.e., during the acoustic lifetimes of the target words and shortly after), even though the shape competitors had been rated as significantly more similar in visual shape to the critical words than to any of the unrelated distractors. We then analyzed the data in later time regions (from 1,000 to 4,000 ms after critical word onset). We did this for two related reasons. First, it appeared surprising that there were no shape effects in eyetracking with printed words, even though, in response to the same spoken materials used in [Experiment 1](#), participants looked at these competitors when they were presented as pictures (Huettig & McQueen, 2007). The question then was whether there were any late (“offline”) biases favoring the shape competitors. Had shape information been made salient? Second, these analyses provided the opportunity to check whether the shape similarity between the target referents and the shape-competitor referents was after all not strong enough to induce an early fixation preference for the shape competitors (i.e., even though the rating results had suggested that these relationships were strong enough). If a late bias for shape competitors were found, then the lack of an early (“online”) effect could not be attributed to weak shape overlap between the stimuli.

[Figure 3](#) plots the fixation proportions at 20-ms intervals to the competitors and the unrelated distractors, over the course of the average trial, for 4,000 ms from critical word onset. The figure shows that there was a late shift (more than 3,000 ms after critical word onset) toward the shape competitors.

Paired t tests on sequential 200-ms time regions showed that the mean shape-competitor/distractor ratios started to differ significantly from the mean ratios in the baseline region during the 3,401- to 3,600-ms time region [$t_1(25) = -2.21, p = .037$; $t_2(39) = -2.62, p = .013$]. This difference in eye gazes between shape competitors and distractors remained significant during the subsequent time regions. It thus appears that participants had a late bias to look at the shape competitors, and hence that shape similarity was salient in the experiment, and that there was sufficient shape similarity between these competitors and the referents of the critical spoken words.

Discussion

Huettig and McQueen (2007) found that with printed word visual displays, attentional shifts were made only to the phonological competitors. No increased overt attention was directed to semantic and shape competitors when phonological matches between spoken and written word were present, both when there was only 200 ms of preview ([Exp. 3](#)) and when the displays appeared at sentence onset ([Exp. 4](#)). The present [Experiment 1](#), with almost identical materials and displays that appeared at sentence onset, demonstrates that participants can use semantic knowledge rapidly to direct attention with printed word displays when phonological matches are absent on critical trials. Importantly, however, participants seem unable to use visual-shape knowledge in a similar rapid (online) manner to direct attention to arrays of printed words. Our rating study and the late (more than 3,000 ms after target word

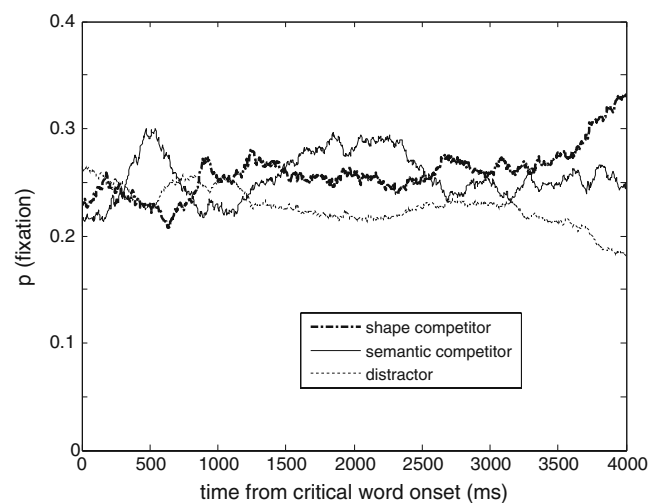


Fig. 3 Time-course graph showing early and late shifts in eye gaze to the different types of printed words in [Experiment 1](#). The x -axis shows time in milliseconds from the acoustic onset of the critical spoken word, for 4,000 ms

onset) bias in looks to the shape competitors rule out the theory that our printed word stimuli evoked different shapes than the intended ones (see the General Discussion for further discussion).

The finding that there were no immediate preferential fixations to the printed forms of the shape competitors leaves open the possibility that the bias toward the semantic competitors is due to an implicit bias toward the mapping of semantic information when phonological matches are not present on critical trials. Shape matches, however, could still in principle be used if semantic matches were no longer present, especially since in the Huettig and McQueen (2007) experiments with picture displays of the same materials, and the identical spoken sentences, participants did look at the shape competitors. In Experiment 2, therefore, we replaced the semantic competitors with a further set of unrelated distractors. Experiment 2 is thus a strong test of whether shape information can be used immediately in visual search of printed word displays: The only match was now one involving the typical shape of the spoken word's referent.

Experiment 2

Method

Participants A further 19 members of the Max Planck Institute for Psycholinguistics participant panel, all native speakers of Dutch, were paid for their participation. All had normal or corrected-to-normal vision. None had participated in Experiment 1.

Stimuli and procedure The same stimuli were used as in Experiment 1, except that each semantic competitor was replaced by an additional unrelated distractor. In order to minimize any similarity in shape and meaning among pictures of the four words in each display, some of the previously used unrelated distractors were paired in Experiment 2 with different critical spoken words than they had been paired with in Experiment 1. The new unrelated distractors are given in Appendix B. The procedure was identical to that in the earlier experiment.

Results

Figure 4 shows a time-course graph that illustrates, in the same way as Fig. 2, the fixation proportions at 20-ms intervals, for 1,000 ms from the critical word onset, to the shape competitors and the unrelated distractors over the course of the average trial. The figure shows that all types of printed words (shape competitors or unrelated distractors) were fixated with approximately equal probabilities.

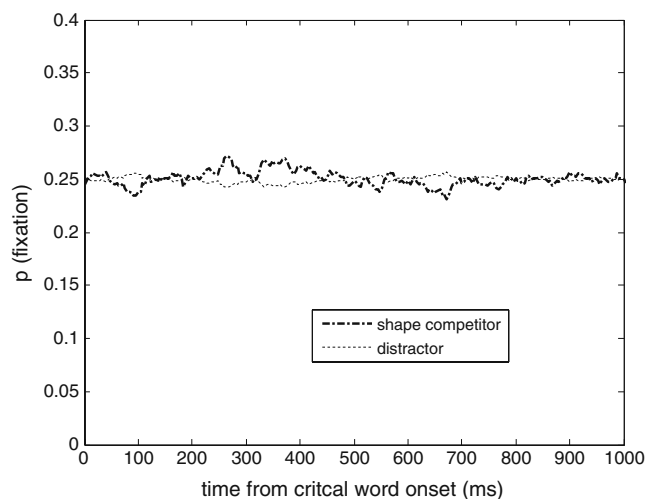


Fig. 4 Time-course graph showing fixation probabilities for the first 1,000 ms after word onset in Experiment 2

Table 3 lists the shape-competitor/distractor ratios for this and later experiments. Paired *t* tests showed that the mean shape-competitor/distractor ratios during the baseline region did not differ significantly [$t_1(18) = 0.79, p > .1$; $t_2(39) = 0.05, p > .1$] from the mean ratios during the 401- to 600-ms time region. We again divided the items on the basis of the shape ratings collected in the rating study into a set of 20 high-similarity items and a set of 20 low-similarity items. These analyses gave the same results as the overall analysis (all *ps* > .1). Similar analyses on the data from the 201- to 400-ms time window yielded no significant differences. During the filler trials (in which a printed word fully matched the spoken critical word), participants strongly shifted their eye gaze toward the matching printed word (e.g., 40% of fixation proportions during the 401- to 600-ms time region; chance level = 25%).

We again analyzed the data in later time regions from 1,000 to 4,000 ms after critical word onset. Figure 5 shows that, as in Experiment 1, there was a late shift in eye gaze toward the shape competitors. Paired *t* tests over subsequent 200-ms time windows showed that the mean shape-competitor/distractor ratios started to differ significantly from the mean ratios in the baseline region during the 2,801- to 3,000-ms time region [$t_1(18) = -2.80, p = .012$; $t_2(36) = -2.32, p = .026$].

Discussion

In Experiment 2, there was no hint of preferential fixations toward the shape competitors during the acoustic lifetime of the critical word, although the only possible match was one involving the typical physical shape of the spoken and printed words' referents. The rating study had demonstrated

Table 3 Shape-competitor/distractor ratios in Experiments 2–4

Time Region (milliseconds from acoustic onset of critical word)	Shape-Competitor/Distractor Ratio		
	Exp. 2	Exp. 3	Exp. 4
0–200	.50	.47	.47
201–400	.51	.45	.44
401–600	.49	.48	.43

that the printed word forms of the competitors did evoke the visual shapes that were intended. The late shape effect shows that there was, in principle, sufficient shape similarity to influence eye gaze, but there was no tendency among participants to look *immediately*, even at the shape competitors that were rated as the most visually similar to the target referents. The absence of an early shape effect is striking, given that participants did look at these shape competitors while listening to exactly the same sentences when the competitors were presented as pictures (Huettig & McQueen, 2007). This difference between the two studies suggests that the retrieval and use of physical shape knowledge depends on the nature of the information in the visual display (i.e., pictures or printed words).

One might wonder, however, whether our failure to find visual-shape effects in Experiments 1 and 2 was due to the sentence frames we used. Perhaps sentences such as “Eventually she looked at the beaker that was in front of her” focused participants insufficiently on the conceptual features of our target words, so that visual-feature knowledge was less likely to be accessed and/or used. In Experiment 3, therefore, we used the same printed word displays as in Experiment 2, but we created a new set of sentence frames for the critical spoken words. These carrier sentences were designed to encourage participants to use visual imagery when comprehending the utterance (e.g., for the critical word “moon,” the sentence “It was a cloudless night and they enjoyed the full moon”; the shape competitor in the display was the printed word “coin”; full moons and coins have a similar global physical shape). The aim was to encourage participants to focus more on the physical shape of the target words, and hence to encourage shape matching between spoken and written targets.

Experiment 3

Method

Participants A further 19 members of the Max Planck Institute for Psycholinguistics participant panel, all native

speakers of Dutch, were paid to take part. All had normal or corrected-to-normal vision. None had participated in the earlier experiments.

Stimuli The same visual stimuli were used as in Experiments 2. A new set of spoken sentences was created. The critical words (e.g., *maan* “moon”) were identical to those in the previous experiments, but each was placed in an imagery-loaded carrier sentence (e.g., *het was een wolkeloze avond en ze genoten van de volle maan* “it was a cloudless night and they enjoyed the full moon”; see Appx. B).

Procedure The procedure was the same as in Experiment 2.

Results

Figure 6 plots the data in the same way as in the earlier graphs. This figure shows that, at the acoustic onset and throughout the acoustic lifetime of the critical word, all types of pictures were fixated with approximately equal probabilities.

Paired *t* tests showed that the mean shape-competitor/distractor ratios (see Table 3) during the baseline region did not differ significantly [$t_1(18) = 0.58, p > .1$; $t_2(39) = 0.92, p > .1$] from the mean ratios during the 401- to 600-ms time region. We again divided the items into high- and low-similarity sets (20 items per set). Separate analyses for each set each revealed no differences between shape competitors and distractors. Similar analyses during the 201- to 400-ms time window yielded no significant differences. During the filler trials, however, participants

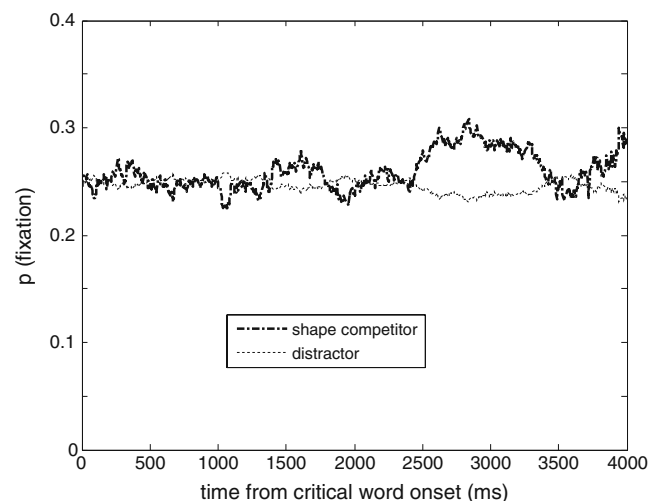


Fig. 5 Time-course graph showing fixation probabilities in milliseconds from the acoustic onset of the critical spoken word, for 4,000 ms, in Experiment 2

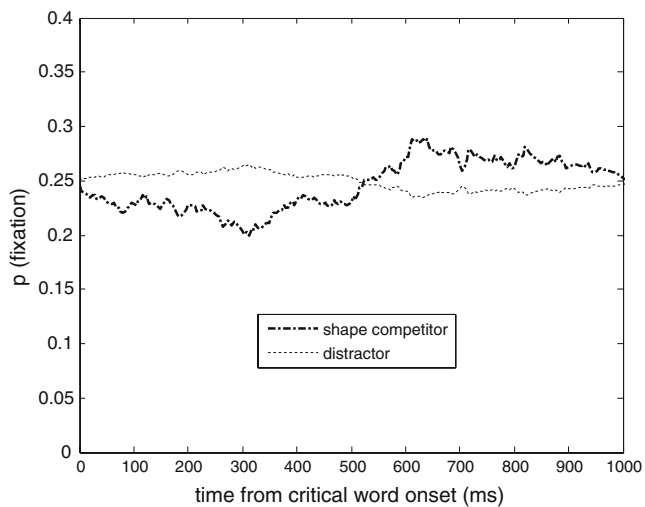


Fig. 6 Time-course graph showing fixation probabilities for the first 1,000 ms after word onset in [Experiment 3](#)

did shift their eye gaze toward the matching printed word (e.g., 51% of fixation proportions during the 401- to 600-ms time region; chance level = 25%). [Figure 7](#) shows that, as in the previous experiments, there was a late shift in eye gaze toward the shape competitors. Paired t tests showed that the mean shape-competitor/distractor ratios started to differ significantly from the mean ratios in the baseline region during the 3,401- to 3,600-ms time region [$t_1(18) = -2.89$, $p = .01$; $t_2(26) = -3.8$, $p = .001$].

Discussion

[Experiment 3](#) revealed that participants did not use visual-feature knowledge immediately to direct attention to arrays of printed words, even with the new set of spoken sentences that were designed to encourage deeper semantic processing and visual imagery. [Experiments 2 and 3](#) thus strongly suggest that participants are unable to establish early matches at the physical shape level of processing with printed word displays, even when no other matches are present. The finding that there are early shape effects with picture displays but not with printed word displays has parallels to findings reported in the perceptual priming literature. [Schreuder, Flores d'Arcais and Glazenberg \(1984\)](#) obtained significantly facilitated target naming times for perceptually related word pairs (e.g., button-coin) and proposed a model of lexical processing in which perceptual representations are activated very rapidly during word recognition. However, [Pecher, Zeelenberg and Raaijmakers \(1998\)](#) failed to replicate these results. Importantly, they obtained significant perceptual priming only when they presented participants with an explicit perceptual decision task (e.g., to judge whether the objects

had a flat surface). This task was intended to activate perceptual information prior to the naming task. Their results thus suggest that the experimental context has a strong influence on the retrieval of stored physical shape features.

In [Experiment 4](#), therefore, we went a step further to encourage shape mapping between the critical spoken and printed words. [Experiment 4](#) was identical to [Experiment 2](#), except that, just before the eyetracking experiment, we asked participants to judge (on a scale from 0 to 10) how similar the physical shape of the concept of the critical word (e.g., “beaker”) was to the physical shape of the concepts invoked by each of the printed words (e.g., “bobbin”). Participants were not presented with all four printed words (of the subsequent visual display) at once, but received items pairwise (e.g., beaker–bobbin, beaker–swan, beaker–umbrella, beaker–bush) in a randomized list of all items in the experiment. This was done so that participants would not stretch their imaginations to find some degree of similarity among at least one of the pairs in the four-word display if none came immediately to mind.

In sum, in [Experiment 4](#) we asked participants to rate the shape similarity between the physical shape of the concept of the spoken critical words and the physical shape of the concepts invoked by each of the printed words. Immediately after this rating task, the same participants then took part in an eyetracking experiment in which they were presented with the stimuli that they had just judged for shape similarity. Such a procedure ensured that shape similarity between critical items was salient to the participants. The question was whether, under these conditions of

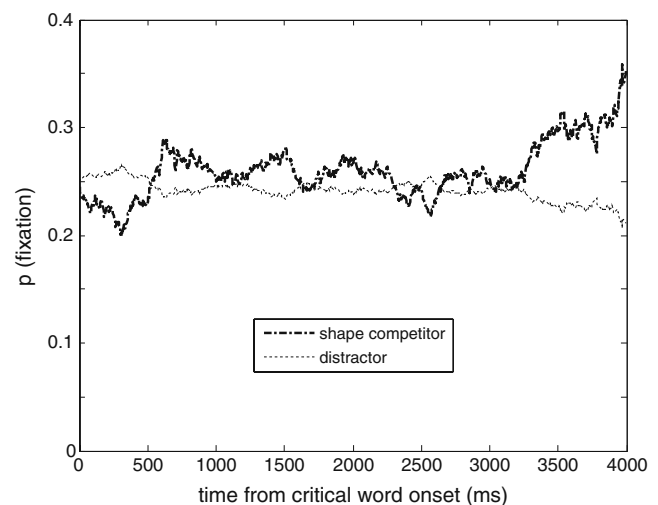


Fig. 7 Time-course graph showing fixation probabilities in milliseconds from the acoustic onset of the critical spoken word, for 4,000 ms, in [Experiment 3](#)

heightened awareness of shape similarity between the shape competitors and the critical words, participants would preferentially fixate those competitors as they heard those words.

Experiment 4

Method

Participants A further 20 members of the Max Planck Institute for Psycholinguistics participant panel, all native speakers of Dutch, were paid to take part. All had normal or corrected-to-normal vision, and none had participated in the earlier experiments.

Stimuli The same spoken sentences as in Experiment 1 and 2 and the same visual stimuli as in Experiments 2 and 3 were used.

Procedure Participants were asked to complete shape ratings prior to the main eyetracking experiment. They were presented, on a computer screen, with each of the 40 critical words, in printed form, each of them paired, one by one (e.g., beaker–bobbin, beaker–umbrella) with the four printed words that would appear later in the eyetracking display associated with that critical word. The order of the pairs was randomized. Participants were asked to judge the similarity in physical shape between the concepts invoked by the two printed words. They were also asked to ignore any similarity in meaning, using an 11-point scale (0 representing *absolutely no similarity in physical shape* and 10 representing *identical physical shape*). The results are shown in Table 4. The shape competitors were judged to be significantly physically more similar to the critical words than were any of the other printed word stimuli (all $ps < .001$).

Right after each participant had completed the ratings, the eyetracking experiment was conducted. The eyetracking procedure was the same as in the previous experiments.

Table 4 Results of the similarity-rating task in Experiment 4: Means (with standard deviations in parentheses) for the rated shape similarity between the critical words and each type of printed word

	Shape Competitors	Unrelated Distractor 1	Unrelated Distractor 2	Unrelated Distractor 3
Shape similarity				
Mean rating (SD)	4.52 (2.27)	0.64 (0.49)	0.72 (0.61)	0.82 (0.91)

Results and discussion

Figure 8 plots the data in the same way as in the earlier graphs. This figure shows that during the first 1,000 ms after critical word onset, all types of pictures were fixated with approximately equal probabilities.

Paired t tests showed that the mean shape-competitor/distractor ratios (see Table 3) during the baseline region did not differ significantly [$t_1(19) = 1.02, p > .1$; $t_2(39) = 0.72, p > .1$] from the mean ratios during the 401- to 600-ms time region. We then divided the shape competitors, on the basis of the shape-rating task of the critical items conducted prior to the eyetracking experiment, into a set of 20 high-similarity items (mean shape similarity rating: 6.43, $SD = 1.08$) and a set of 20 low-similarity items (mean shape similarity rating: 2.62, $SD = 1.33$), and reran the competitor/distractor analyses separately for each item set. Both of these analyses gave the same results as the overall analysis (all $ps > .1$). Similar analyses during the 201- to 400-ms time window yielded no significant differences.

Thus, in Experiment 4 there was again no hint of preferential fixations toward the shape competitors during the time of interest (i.e., during the acoustic lifetimes of the target words and shortly after), even though participants had completed shape-similarity ratings of the critical items right before the main eyetracking experiment, and had indeed rated the shape competitors as being more similar in visual shape to the critical words than were any of the unrelated distractors. We again analyzed the data in later time regions (from 1,000 to 4,000 ms after critical word onset), because it appeared surprising that there were no shape effects in eyetracking, even though shape similarity had been primed by the shape-rating task. These analyses provided another opportunity to check whether the shape similarity between the target referents and the shape-competitor referents was not strong enough to induce an early fixation preference for the shape competitors (even though the rating results had suggested that they were strong enough). Figure 9 shows that there was a late shift (starting around 2,500 ms after critical word onset or later) in overt attention toward the shape competitors.

Paired t tests showed that the mean shape-competitor/distractor ratios in Experiment 4 started to differ significantly from the mean ratios in the baseline region during the 2,601- to 2,800-ms time region [$t_1(19) = 2.78, p = .012$; $t_2(39) = 5.18, p < .001$]. This difference in eye gaze between shape competitors and distractors remained significant during the subsequent time regions. At the point of greatest magnitude of the difference in fixations between shape-competitors and distractors (at 3,800 ms after critical word onset), there was a significant correlation between shape ratings and the fixation proportions to the shape competitors (Pearson correlation, $r = .40, p = .01$).

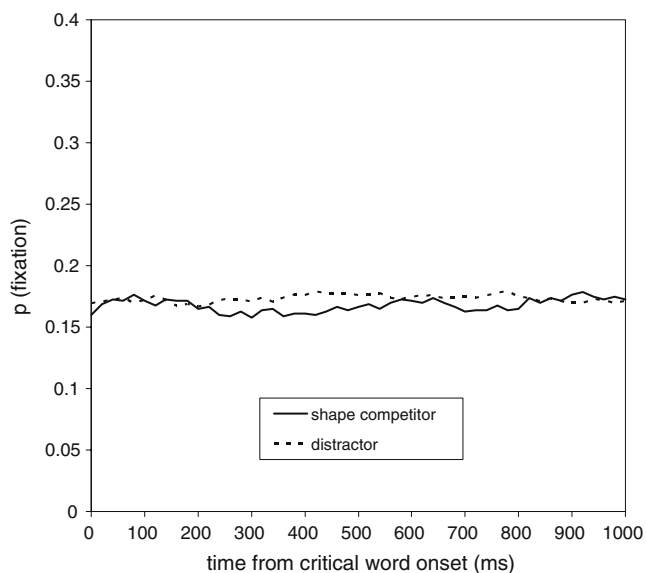


Fig. 8 Time-course graph showing fixation probabilities for the first 1,000 ms after word onset in [Experiment 4](#)

It thus appears that participants had a late bias to look at the shape competitors—hence, that shape similarity was salient—and that there was sufficient shape similarity between these competitors and the referents of the critical spoken words. Furthermore, the participants in [Experiment 4](#) had indeed rated the shape competitors as being similar in shape to the referents of the critical spoken words, and there were no other (semantic, phonological, or orthographic) matches present between the printed and spoken words that the participants could use to drive visual search. In spite of all this, they showed no tendency to prefer to look at the shape competitors as they heard the critical spoken words.

General discussion

In four experiments, Dutch participants listened to spoken Dutch sentences while looking at visual displays of four spatially distinct written words. In [Experiment 1](#), semantic competitors, physical shape competitors, and unrelated distractors were present in the display. We observed significant shifts in overt attention to the semantic competitors but no immediate preferential fixations on the shape competitors. In [Experiment 2](#), we replaced the semantic competitors with a further set of unrelated distractors. As in [Experiment 1](#), we did not observe any early shifts in overt attention to the visual competitors, although now the only possible match was one involving the typical physical shape of the spoken and printed words' referents. In [Experiment 3](#), we used more

semantically loaded sentence frames to encourage shape retrieval. Again there was no significant difference in overt attention between shape competitors and unrelated distractors during the acoustic lifetime of the target word. In [Experiment 4](#), participants took part in a shape-rating task of the critical items prior to the eyetracking experiment, yet again we observed no hint of immediate preferential fixations to the shape competitors. In all four experiments, however, participants showed late biases in overt attention toward the shape competitors, suggesting that the shape similarity between these competitors and the referents of the critical spoken words was strong enough to have driven early fixation behavior.

Our study used the look-and-listen variant of the paradigm. The results are unlikely to be due to this particular feature. Dahan and Tanenhaus (2005), for instance, using a task in which participants had to move named objects above or below a geometric shape adjacent to the object using a computer mouse, obtained physical shape effects very similar to those found with the look-and-listen task by Huettig and Altmann (2004, 2007). 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), again using the look-and-listen task.

The present set of experiments is further evidence that lexical knowledge is not accessed in an all-or-none manner, such that (partial) retrieval of a word necessarily entails retrieval of all semantic or physical shape knowledge (cf. Norris et al., 2006; see also Moss,

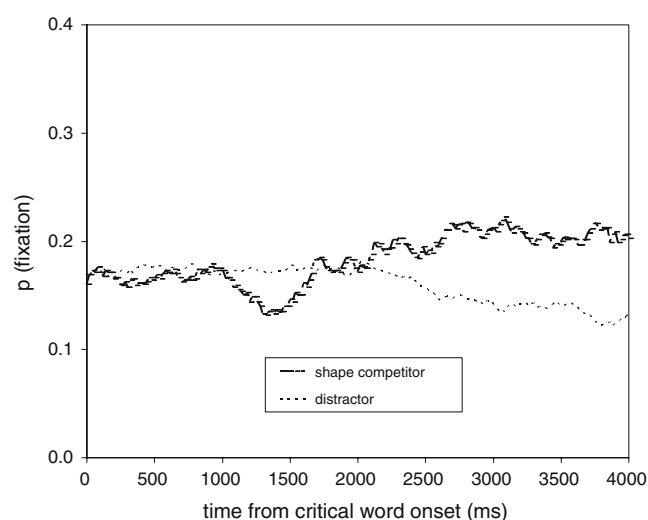


Fig. 9 Time-course graph showing fixation probabilities in milliseconds from the acoustic onset of the critical spoken word, for 4,000 ms, in [Experiment 4](#)

McCormick & Tyler 1997). If storage and/or retrieval of semantic representations were not at least partially independent from storage and/or retrieval of shape representations, then in [Experiment 1](#) we should have observed looks to shape competitors in addition to the shifts to semantic competitors.

More importantly, the present research, in conjunction with the data of Huettig and McQueen (2007), strongly suggests that the nature of the information in the visual environment induces implicit biases during language-mediated visual search. If the visual display consists of pictures of objects (and sufficient preview is given), participants retrieve shape, semantic (cf. Biederman et al., 1988; Biederman et al., 1982; Intraub, 1984; Loftus & Ginn, 1984; Potter, 1976; Smith & McGee, 1980), and phonological (cf. Meyer et al., 2007; Meyer & Damian, 2007) information. These (visually derived) representations are then available for the mapping process with spoken words (i.e., language-derived representations) during language-mediated visual search.

The situation is different if printed word displays are used. The results of [Experiments 3](#) and [4](#) of the Huettig and McQueen (2007) study suggest that printed word displays induce a bias toward a phonologically based mode of processing. Huettig and McQueen argued that the tendency for participants to look only at phonological competitors in a printed word display reflects matches between phonological representations. Salverda and Tanenhaus (2010) showed recently that this behavior can also be driven by orthographic rather than phonological overlap, and hence that it might sometimes reflect matches between orthographic rather than phonological representations. In either case, however, these matches reflect overlap in linguistic form, and not semantic or visual-shape overlap. Huettig and McQueen found no evidence that semantic- and shape-based information was used immediately to direct eye gaze around a printed word display. Our present [Experiment 1](#) suggests that if no phonologically based match is possible (because no phonological competitors are present in the display), participants do use semantic information quickly in the mapping process. Most interestingly, [Experiments 1–4](#) provide strong evidence that shape information is not used in a similar manner. Even though phonological and semantic matches were not present on the critical trials ([Exps. 2–4](#)), and, in addition, a set of semantically loaded sentences was used ([Exp. 3](#)), or, in addition, an explicit shape-rating task was administered to draw attention to the shape similarity between critical items ([Exp. 4](#)), participants did not use shape information during the acoustic lifetime of the target word to direct eye gaze.

We can rule out that the absence of early shape effects in the present set of experiments is due to insufficient visual similarity between the referents of the spoken and printed words. As already noted, in Huettig and McQueen (2007) we used the same set of shape competitors (picture displays instead of printed word displays), and participants showed strong early effects of physical shape competition upon hearing the spoken words. Furthermore, the rating data in [Experiments 1](#) and [4](#) rule out that the printed words we used here systematically evoke different visual shapes than the pictures of Huettig and McQueen. Moreover, in [Experiments 1, 2, and 4](#), we used exactly the same spoken sentences as in that study, yet observed no immediate shift in attention. Finally, [Experiment 4](#) revealed a significant correlation between participants' shape ratings and the late bias in eye gaze to the shape competitors.¹ Thus, the contrast between pictures and printed words in Huettig and McQueen and the present set of experiments suggests that the likelihood of early online (conceptual) visual shape mapping depends on the nature of the information in the visual environment.

Our finding of late biases in overt attention to the shape competitors reveals that participants can access shape information from printed words. The fact that these shifts occurred more than 2,000 ms after critical word onset, however, indicates that this information was not retrieved rapidly. Of course eye gaze is a measure of overt attention and not directly of “activation of underlying representations,” and thus we cannot rule out with absolute certainty that visual-shape representations were accessed immediately in all four experiments. According to this account, there would be fast and efficient retrieval of all lexical knowledge (including conceptual shape representations) from printed words, but this information would not be used immediately during language-mediated visual search with printed word displays. We consider this possibility very unlikely, however. This argument could explain why there were no attentional shifts to semantic competitors in the printed word experiments of Huettig and McQueen (2007)—namely, that the easy access of phonologically based information from printed words resulted in phonologically mediated (rather than semantically mediated) mapping behavior. Once phonologically (or orthographically) based matches were not present on critical trials ([Exp. 1](#) of the

¹ Future research could usefully explore whether the much-delayed effects of shape knowledge arise because it takes time to overcome an implicit bias against the retrieval and use of visual-shape information with printed word displays.

present study), participants used retrieved semantic knowledge to shift overt attention to semantic competitors in the visual display. This line of reasoning, however, fails to explain why there were no fast and efficient shifts to shape competitors in the present Experiments 2, 3, and 4, although no other representational matches were present. It fails to account for the finding that our participants seemed unable to use shape knowledge online, even though we designed Experiments 2–4 to encourage shape mapping. In the experimental trials, no fully matching targets were present (e.g., hearing “beaker” with the written target *beaker* in the display), a procedure that has been found to maximize the opportunity to observe competitor effects, because under these conditions participants search the display on all trials for matches between linguistic and visual information (see Huettig & Altmann, 2005; Huettig & McQueen, 2007). Moreover, in Experiments 2–4, no other representational matches were present, so the only match present between the spoken critical word and the printed word was one involving shape. Thus, there appears to be an important difference in how the nature of the visual environment induces the retrieval and use of visual-shape information. Our findings suggest that there is no fast and efficient retrieval of (conceptual) shape information from printed words, even in a task situation that makes shape similarity salient to participants (Exp. 4).

Why do our results differ from those observed by Schreuder et al. (1984)? One obvious reason is that there are crucial differences in the methodologies used. Schreuder et al. found priming between visually related word pairs (e.g., button–coin) in both a word naming task (“name the target word as quickly as possible, reading it aloud”) and a lexical decision task (“press the left or right button, respectively, when the target string spelled a Dutch word, otherwise press the other button”). They interpreted their results as reflecting fast access of perceptual (e.g., shape) information from written words. We are not the first to point out that there are problems with such an interpretation (see Moss, Ostrin, Tyler & Marslen-Wilson 1995; Pecher et al., 1998; Shelton & Martin, 1992; Williams, 1996). Pecher et al., for instance, argued that the procedure adopted by Schreuder et al. (1984) may have encouraged the use of strategies, because the written prime remained on the screen when the written targets were presented—a procedure that draws attention to the relation between prime and target. Other problematic issues are that the targets were presented four times (each time with a different prime), that word pairs were repeated later in the experiment if an incorrect response was made, and that some words were used as both target and prime

(see Pecher et al., 1998, for further discussion). Importantly (as we already mentioned), Pecher et al. failed to replicate the Schreuder et al. results in a series of experiments. Pecher et al. observed a priming effect for perceptually related word pairs only if participants had taken part in a perceptual decision task (i.e., judging whether the object referred to an oblong object and whether it had a flat surface) prior to the naming task. Interestingly, Pecher et al. found no priming effect when the lexical decision task was used, even with a preceding perceptual decision task. This finding fits well with the absence of immediate shifts in eye gaze to the shape competitors in the present experiments, even when an explicit shape-rating task was administered to draw attention to the shape similarity between critical items (Exp. 4).

We conjecture that the lack of an early preference for the shape competitors in all four experiments reported here appears because printed words induce an implicit bias against the rapid (online) use of (conceptual) visual-shape information. This bias is *implicit* in the sense that we suppose that participants are not explicitly choosing to ignore stored shape information. Indeed, the evidence that participants look much later at the shape competitors (not less than about 2 s after the acoustic onset of the critical spoken words) shows that they were not deliberately avoiding looking at those competitors. The bias thus appears to be driven implicitly, by the nature of the input. While information about the shape of objects is present in picture displays (and is used immediately in visual search of such displays), it must be retrieved from long-term memory when printed word displays are presented. It appears that there is no fast and efficient retrieval of this information when someone sees an array of printed words. It is not the case that retrieval of shape information is blocked, because visual search is being driven by other types of information (e.g., by phonological, orthographic, or semantic matches), since in Experiments 2–4 no other matches were present. Instead, the display itself (i.e., printed words rather than pictures) seems to signal that information about the typical shape of visual objects should not be retrieved rapidly or used to guide visual search.

To conclude, our findings are important with regard to the nature of information retrieval and information use during the processing of printed word and picture stimuli. What our results demonstrate is that different types of stimuli (e.g., pictures or printed words) and different visual environments (e.g., whether other representational matches are present in the concurrent visual surroundings) induce implicit biases toward particular modes of processing during language-mediated visual search.

Appendix A

Table 5 Experimental materials in Experiment 1

Critical Spoken Word	Shape Competitor	Semantic Competitor	Unrelated Distractor 1	Unrelated Distractor 2
boon (bean)	sabel (sword)	sla (lettuce)	muur (wall)	cello (cello)
hoefijzer (horseshoe)	magneet (magnet)	zadel (saddle)	boot (boat)	filter (filter)
peddel (paddle)	fluit (flute)	zeilboot (sailboat)	reiger (heron)	bril (glasses)
ballon (balloon)	zon (sun)	pop (doll)	muis (mouse)	deur (door)
raket (rocket)	fles (bottle)	vlieger (kite)	orgel (organ)	emmer (bucket)
arm (arm)	rietje (straw)	nier (kidney)	paddestoel (mushroom)	muts (hat)
hart (heart)	voetbal (football)	gebit (teeth)	klok (clock)	bloemkool (cauliflower)
ananas (pineapple)	boei (buoy)	pinda (peanut)	schaap (sheep)	spijker (nail)
paleis (palace)	kennel (kennel)	koning (king)	garnaal (shrimp)	slee (sledge)
bal (ball)	kers (cherry)	shuttle (shuttlecock)	bezem (broom)	hond (dog)
lelie (lily)	kroon (crown)	cactus (cactus)	podium (stage)	mossel (mussel)
kerk (church)	iglo (igloo)	graf (grave)	radio (radio)	pan (pot)
berg (mountain)	servet (napkin)	wolk (cloud)	paling (eel)	kies (tooth)
boor (drill)	pijl (arrow)	ladder (ladder)	tijger (tiger)	neus (nose)
bord (plate)	wiel (wheel)	karaf (pitcher)	schroef (screw)	aap (ape)
koffer (suitcase)	schilderij (picture)	tent (tent)	banaan (banana)	mug (mosquito)
tang (pliers)	broek (trousers)	fietspomp (bicycle pump)	boeket (bouquet)	oor (ear)
liniaal (ruler)	kam (comb)	kubus (cube)	enkel (ankle)	paprika (pepper)
das (tie)	veer (feather)	trui (jumper)	glijbaan (slide)	trommel (drum)
moer (nut)	donut (donut)	hamer (hammer)	prei (leek)	laars (boot)
dolk (dagger)	kurkretrekker (corkscrew)	kanon (cannon)	hamster (hamster)	television (TV)
fakkelt (torch)	ijsje (ice cream)	bom (bomb)	rasp (grater)	knoop (button)
schildpad (turtle)	ton (barrel)	haai (shark)	kampvuur (campfire)	penseel (paintbrush)
beker (beaker)	klos (bobbin)	vork (fork)	struik (shrub)	paraplu (umbrella)
hek (fence)	rail (railway line)	sleutel (key)	fornuis (stove)	tas (bag)
ketel (kettle)	slot (lock)	vijzel (jack)	mond (mouth)	vos (fox)
riem (belt)	slang (snake)	sandaal (sandal)	gordijn (curtain)	asbak (ashtray)
kogel (bullet)	ui (onion)	speer (spear)	rots (rock)	vest (waistcoat)
kano (canoe)	worst (sausage)	fontein (fountain)	rok (skirt)	tuba (tuba)
matras (mattress)	brief (letter)	kruk (stool)	gieter (watering can)	trompet (trumpet)
toren (tower)	beitel (chisel)	brug (bridge)	klomp (clog)	haas (hare)
boek (book)	kaart (playing card)	potlood (pencil)	snavel (beak)	zaag (saw)
silo (silo)	kasteel (castle)	tractor (tractor)	rugzak (rucksack)	bed (bed)
tol (top)	aardbei (strawberry)	baby (baby)	stempel (stamp)	bus (bus)
maan (moon)	gulden (guilden)	tornado (tornado)	nijlpaard (hippopotamus)	voet (foot)
zwaard (sword)	pincet (tweezers)	pistool (gun)	gans (goose)	ster (star)
vijl (file)	zuil (column)	schaar (scissors)	pleister (plaster)	safe (safe)
soldaat (soldier)	robot (robot)	bijl (axe)	rivier (river)	piano (piano)
dokter (doctor)	kabouter (gnome)	sput (syringe)	ezel (donkey)	mand (basket)
pen (pen)	sigaret (cigarette)	bureau (desk)	mier (ant)	anker (anchor)

Appendix B

Table 6 Additional materials used in Experiments 2–4

Critical Spoken Word	Shape Competitor	Additional Unrelated Distractor (Exps. 2–4)	Imagery-Loaded Sentence(Exp. 3)
boon (bean)	sabel (sword)	vogel (bird)	in het met aarde gevulde bakje stopt hij een <i>boon</i> he put a bean in the tub that was filled with dirt
hoefijzer (horseshoe)	magneet (magnet)	drumstel (drum set)	om geluk af te dwingen hing boven de deur een <i>hoefijzer</i> there was a horseshoe above the door for good luck
peddel (paddle)	fluit (flute)	sinaasappel (orange)	de roeier wil vaart maken en pakt zijn <i>peddel</i> the rower wants to pick up speed and grabs his paddle
ballon (balloon)	zon (sun)	fluit (flute)	hij zag dat het ding dat in de takken vastzat een <i>ballon</i> was the noticed that the object that was stuck in the branches was a balloon
raket (rocket)	fles (bottle)	huis (house)	ze werden de ruimte ingeschoten met een <i>raket</i> they were launched into space in a rocket
arm (arm)	rietje (straw)	doos (can)	bovenop de dekens lag haar <i>arm</i> her arm was lying on top of the blankets
hart (heart)	voetbal (football)	vaas (vase)	als teken van zijn liefde voor zijn vriendin tekende hij een <i>hart</i> he drew a heart as a sign of his love for his girlfriend
ananas (pineapple)	boei (buoy)	ketting (necklace)	de hawaiaanse dame droeg een <i>ananas</i> the Hawaiian lady was carrying a pineapple
paleis (palace)	kennel (kennel)	bezem (broom)	de toeristen keken vol bewondering naar het <i>paleis</i> the tourists looked at the palace in admiration
bal (ball)	kers (cherry)	trein (train)	hij schopt tegen de <i>bal</i> he kicks the ball
lelie (lily)	kroon (crown)	ring (ring)	op de vijver drijft een <i>lelie</i> there is a lily floating on the pond
kerk (church)	iglo (igloo)	vork (fork)	op het plein stond een statige <i>kerk</i> there is an imposing church on the square
berg (mountain)	servet (napkin)	radio (radio)	vanuit de helikopter had zij een mooi uitzicht op de <i>berg</i> from the helicopter she had a great view of the mountain
boor (drill)	pijl (arrow)	hoed (hat)	hij maakt een gat in de muur met een <i>boor</i> he makes a hole in the wall with a drill
bord (plate)	wiel (wheel)	agent (policeman)	op de tafel stond een kandelaar en een <i>bord</i> there were a candlestick and a plate on the table
koffer (suitcase)	schilderij (picture)	schroef (screw)	de man op het station zat op zijn <i>koffer</i> the man at the station was sitting on his suitcase
tang (pliers)	broek (trousers)	konijn (rabbit)	hij trok de spijker uit de muur met een <i>tang</i> he pulled the nail from the wall with pliers
liniaal (ruler)	kam (comb)	kam (computer)	hij tekent de raaklijn met zijn potlood en <i>liniaal</i> he draws the tangent with his pencil and ruler
das (tie)	veer (feather)	glijbaan (slide)	er stonden allemaal cartoons op zijn <i>das</i> there were lots of cartoons on his tie
moer (nut)	donut (donut)	gitaar (guitar)	op de grond in de schuur zag hij een <i>moer</i> he noticed a nut on the floor of the shed
dolk (dagger)	kurkentrekker (corkscrew)	jas (jacket)	uit de borst van het slachtoffer steekt een <i>dolk</i> there is a dagger sticking out of the victim's chest
fakkel (torch)	ijsje (ice cream)	dorp (village)	de gids liep voorop en verlichtte de grot met een <i>fakkel</i> the guide led the way and lit up the cave with his torch
schildpad (turtle)	ton (barrel)	fiets (bicycle)	het dier dat over het strand loopt is een <i>schildpad</i> the animal walking on the beach is a turtle

Table 6 (continued)

Critical Spoken Word	Shape Competitor	Additional Unrelated Distractor (Exps. 2–4)	Imagery-Loaded Sentence(Exp. 3)
beker (beaker)	klos (bobbin)	zwaan (swan)	hij drinkt zijn kop koffie altijd uit een blauwe <i>beker</i> he always drinks his coffee from a blue beaker
hek (fence)	rail (railway line)	sla (salade)	het privestrand werd van het gewone strand gescheiden door een groot <i>hek</i> the private beach was separated from the regular beach by a big fence
ketel (kettle)	slot (lock)	vlieg (fly)	om straks thee te zetten staat op het gasfornuis alvast een <i>ketel</i> there's already a kettle on the stove to make tea a bit later
riem (belt)	slang (snake)	brood (bread)	om te voorkomen dat zijn broek op zijn knieën hangt gebruikt hij een <i>riem</i> he uses a belt to stop his pants from dropping to his knees
kogel (bullet)	ui (onion)	aanstecker (sticker)	met zijn tanden ving de goochelaar de <i>kogel</i> the magician caught the bullet with his teeth
kano (canoe)	worst (sausage)	lift (escalator)	de indiaan reist stroomafwaarts in zijn <i>kano</i> the Indian travels downstream in his canoe
matras (mattress)	brief (letter)	koe (cow)	na het afhalen van het bedlaken zag hij het gat in het <i>matras</i> after removing the sheets he saw the hole in the mattress
toren (rook)	beitel (chisel)	tomaat (tomato)	het enige stuk dat zij behalve de koning nog op het bord had staan was de <i>toren</i> apart from the king she only had the rook left on the board
boek (book)	kaart (playing card)	varken (pig)	uit haar tas viel een <i>boek</i> a book fell from her bag
silo (silo)	kasteel (castle)	pistool (pistol)	naast de boerderij stond een torenhoge <i>silo</i> next to the farm stood a very high silo
tol (top)	aardbei (strawberry)	hond (dog)	hij bleef maar kijken naar de ronddraaiende <i>tol</i> he just kept watching the spinning top
maan (moon)	gulden (guilder)	brug (bridge)	het was een wolkeloze avond en ze genoten van de volle <i>maan</i> it was a cloudless night and they enjoyed the full moon
zwaard (sword)	pincet (tweezers)	karaf (carafe)	om zichzelf te verdedigen trok de ridder zijn <i>zwaard</i> the knight pulled his sword to defend himself
vijl (file)	zuil (column)	kubus (cube)	de gedetineerde vond in zijn taart een <i>vijl</i> the prisoner found a file in his cake
soldaat (soldier)	robot (robot)	bever (beaver)	de ingang van de kazernes werd bewaakt door een enkele <i>soldaat</i> the entrance to the barracks was guarded by a single soldier
dokter (doctor)	kabouter (gnome)	bloemen (flowers)	de persoon met de stethoscoop en het grijze haar is de <i>dokter</i> the person with the stethoscope and the grey hair is the doctor
pen (pen)	sigaret (cigarette)	kaas (cheese)	als cadeau voor haar jubileum gaf haar baas een zilveren <i>pen</i> her boss gave her a silver pen for her anniversary

References

- Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419–439.
- Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247–264.
- Biederman, I., Blickle, T. W., Teitelbaum, R. C., & Klatsky, G. J. (1988). Object search in nonscene displays. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 14, 456–467.
- Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, 14, 143–177.
- 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.
- 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.

- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.
- Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, *123*, 71–99.
- 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*, 23–32.
- 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., & McQueen, J. M. (2007). The tug of war between phonological, semantic, and shape information in language-mediated visual search. *Journal of Memory and Language*, *54*, 460–482.
- Huettig, F., Olivers, C. N. L., & Hartsuiker, R. J. (in press). Looking, language, and memory: Bridging research from the visual world and visual search paradigms. *Acta Psychologica*. doi:10.1016/j.actpsy.2010.07.013
- Huettig, F., Quinlan, P. T., McDonald, S. A., & Altmann, G. T. M. (2006). Models of high-dimensional semantic space predict language-mediated eye movements in the visual world. *Acta Psychologica*, *121*, 65–80.
- Huettig, F., Rommers, J., & Meyer, A. S. (in press). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*. doi:10.1016/j.actpsy.2010.11.003
- Intraub, H. (1984). Conceptual masking: The effects of subsequent visual events on memory for pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 115–125.
- Loftus, G. R., & Ginn, M. (1984). Perceptual and conceptual masking of pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 435–441.
- Matin, E., Shao, K., & Boff, K. (1993). Saccadic overhead: Information processing time with and without saccades. *Perception & Psychophysics*, *53*, 372–380.
- McQueen, J. M., & Viebahn, M. (2007). Tracking recognition of spoken words by tracking looks to printed words. *The Quarterly Journal of Experimental Psychology*, *60*, 661–671.
- Meyer, A. S., Belke, E., Telling, A. L., & Humphreys, G. W. (2007). Early activation of object names in visual search. *Psychonomic Bulletin & Review*, *14*, 710–716.
- Meyer, A. S., & Damian, M. F. (2007). Activation of distractor names in the picture-picture interference paradigm. *Memory & Cognition*, *35*, 494–503.
- Moss, H. E., McCormick, S., & Tyler, L. K. (1997). The time course of activation of semantic information during spoken word recognition. *Language and Cognitive Processes*, *12*, 695–731.
- Moss, H. E., Ostrin, R. K., Tyler, L. K., & Marslen-Wilson, W. D. (1995). Accessing different types of lexical semantic information: Evidence from priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1–21.
- Norris, D., Cutler, A., McQueen, J. M., & Butterfield, S. (2006). Phonological and conceptual activation in speech comprehension. *Cognitive Psychology*, *53*, 146–193.
- 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.
- Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *2*, 509–522.
- Salverda, A. P., & Tanenhaus, M. K. (2010). Tracking the time course of orthographic information in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1108–1117.
- Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of the Optical Society of America*, *57*, 1030–1033.
- Schreuder, R., Flores d'Arcais, G. B., & Glazenberg, G. (1984). Effects of perceptual and conceptual similarity in semantic priming. *Psychological Research*, *45*, 339–354.
- Shelton, J. R., & Martin, R. C. (1992). How semantic is automatic semantic priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1191–1210.
- Smith, M. C., & McGee, L. E. (1980). Tracing the time-course of picture-word processing. *Journal of Experimental Psychology: General*, *109*, 373–392.
- Stolz, J. A., & Besner, D. (1996). Role of set in visual word recognition: Activation and activation blocking as nonautomatic processes. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1166–1177.
- Stolz, J. A., & Besner, D. (1999). On the myth of automatic semantic activation in reading. *Current Directions in Psychological Science*, *8*, 61–65.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.
- Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 371–386.
- Williams, J. N. (1996). Is automatic priming semantic? *European Journal of Cognitive Psychology*, *8*, 113–161.
- Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, *1*, 202–238.
- 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.

We thank Laurence Bruggeman, Lies Cuijpers, Holger Mitterer, Vera Hoskam, and Willemijn van den Berg for their assistance, and three anonymous reviewers for constructive comments on a previous version of this article. Parts of this research were presented at the AMLaP 2008 conference in Cambridge, England.