

Interference of spoken word recognition through phonological priming from visual objects and printed words

James M. McQueen · Falk Huettig

Published online: 17 October 2013
© Psychonomic Society, Inc. 2013

Abstract Three cross-modal priming experiments examined the influence of preexposure to pictures and printed words on the speed of spoken word recognition. Targets for auditory lexical decision were spoken Dutch words and nonwords, presented in isolation (Experiments 1 and 2) or after a short phrase (Experiment 3). Auditory stimuli were preceded by primes, which were pictures (Experiments 1 and 3) or those pictures' printed names (Experiment 2). Prime–target pairs were phonologically onset related (e.g., *pijl–pijn*, *arrow–pain*), were from the same semantic category (e.g., *pijl–zwaard*, *arrow–sword*), or were unrelated on both dimensions. Phonological interference and semantic facilitation were observed in all experiments. Priming magnitude was similar for pictures and printed words and did not vary with picture viewing time or number of pictures in the display (either one or four). These effects arose even though participants were not explicitly instructed to name the pictures and where strategic naming would interfere with lexical decision making. This suggests that, by default, processing of related pictures and printed words influences how quickly we recognize spoken words.

Keywords Picture recognition · Priming · Phonological interference · Semantic facilitation

J. M. McQueen
Behavioural Science Institute, Radboud University Nijmegen,
Nijmegen, The Netherlands

J. M. McQueen · F. Huettig
Donders Institute for Brain, Cognition and Behaviour, Radboud
University Nijmegen, Nijmegen, The Netherlands

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

A characteristic aspect of human cognition is our ability to integrate rapidly visual or auditory input with stored linguistic and nonlinguistic mental representations (for reviews, see Anderson, Chiu, Huette, & Spivey, 2011; Huettig, Mishra, & Olivers, 2012; Huettig, Olivers, & Hartsuiker, 2011a; Lupyán, 2012a, 2012b). Our shifts in eye gaze, for instance, are frequently, be it consciously or subconsciously, directed by spoken language input. A mother may tell her child to mind the step or to look at the beautiful flower, a visitor may ask for a glass of water, or via mobile phone, we may be given directions to find an unknown location. Here, we focus on the other side of the coin: Does exposure to visual stimuli influence how we recognize spoken words? More specifically, we ask whether viewing objects (e.g., the line drawing of a beaver) or printed words (the printed word BEAVER) results in retrieval of the associated phonological form (“beaver”) and, consequently, affects the recognition of related spoken words, such as “beaker.”

Some previous eye-tracking research suggests that viewers often access the names of objects, even when they are not required to name them. In one study (Huettig & McQueen, 2007), Dutch participants were asked to listen to sentences while looking at displays of four pictures. Given a sentence including the Dutch word *beker*, ‘beaker,’ for example, the visual display contained phonological (a beaver, *bever*), shape (a bobbin, *klos*), and semantic (a fork, *vork*) competitors and an unrelated distractor (an umbrella, *paraplu*) which served as a baseline. The participants’ task was to listen carefully to the sentences. They were told that they were free to look at whatever they wanted but that they should not take their eyes off the screen (see Huettig, Rommers, & Meyer, 2011b, for further discussion of this task). When the display appeared at sentence onset, there were preferential fixations during the acoustic lifetime of ‘beaker’ to all three types of competitors over the unrelated distractors, and fixations to phonological competitors preceded fixations to shape and semantic competitors. This was presumably because the preview gave participants sufficient

time to retrieve the names of the four objects in the displays, and thus, a match in phonological representations activated by both spoken words and visual objects resulted in an increased likelihood of a saccade toward the location of the (partially) matching visual object. Experiment 2 (Huettig & McQueen, 2007) provided further support for this account of the results. When display onset was only 200 ms before hearing, for example, *beker*, fixations were directed to shape and then semantic competitors, but not phonological competitors (i.e., participants no longer preferentially fixated the phonological competitor over unrelated distractors). In other words, with very short preview, picture processing still involved retrieval of visual and semantic features, but there was insufficient time for information to reach the phonological level (i.e., to retrieve the pictures' names), and thus word–object mapping could occur only at semantic and visual, but not phonological, levels.

This account of the Huettig and McQueen (2007) results is in line with models of information processing during language production according to which information processing at one level (e.g., the conceptual level) cascades to later levels (e.g., the phonological level; see Caramazza, 1997; Costa, Caramazza, & Sebastian-Galles, 2000; Costa, Santesteban, & Cano, 2005; Dell, 1986; Griffin & Bock, 1998; Harley, 1993; Navarrete & Costa, 2005; Rapp & Goldrick, 2000). According to an influential model of language production (Levelt, Roelofs, & Meyer, 1999), on the other hand, there is no cascaded processing between conceptual and phonological levels. Levelt et al. argued that speech production is a serial, two-stage process. A concept is first activated at the conceptual level, but its phonology is retrieved only at the word-form level. Critically, the phonology of no other word is retrieved. A picture's name is thus retrieved only if it has been selected for production (e.g., in picture naming, only if the participant intends to name the picture). This account conflicts with our interpretation of the Huettig and McQueen (2007) eye-tracking data.

It could be argued, however, that eye-tracking experiments do not involve passive picture viewing. Even though the participants in the Huettig and McQueen (2007) study had no physical task to perform (e.g., they were not required to click on the pictures with a computer mouse, as in many visual-world studies), they still were engaged in a task that encouraged word–object mapping, and thus it cannot be ruled out that they chose implicitly to name the objects (see Huettig et al., 2011b, for an extensive discussion).

Moreover, there is some language production work that could be taken to be inconsistent with the view that object names are retrieved during object processing. Jescheniak, Oppermann, Hantsch, Wagner, Maedebach and Schriefers (2009) failed to observe phonological priming effects from context pictures that were phonologically related to a to-be-named picture (but see Morsella & Miozzo, 2002, who observed such an effect). Bloem, van den Boogaard, and La Heij (2004; see also Bloem & La Heij, 2003) used a word-

translation task in which Dutch–English bilinguals were required to translate an English target word (e.g., *witch*, Dutch *heks*) as quickly as possible while ignoring a phonologically related context picture (the picture of a hedge, Dutch *hek*) or a phonologically unrelated context picture (the picture of a pepper, Dutch *paprika*). Reaction times (RTs) did not differ between phonologically related and unrelated conditions. The authors took these results to argue that the name of the context picture was not retrieved and, thus, could not facilitate participant performance. Note, however, that the question of whether context pictures automatically activate phonological representations when participants are engaged in an explicit *word production* task, such as picture naming or word translation, is an interesting one but a different one from whether there is phonological priming from visual objects and printed words during spoken *word recognition*. It is conceivable that attentional control could limit phonological retrieval when the current task is to speak and that, in such a situation, retrieval of phonology is limited to the to-be-named word (Levelt et al., 1999). This does not mean that phonological representations are not retrieved during object processing in the absence of an explicit naming task.

Inconsistent results with regard to whether information processing during object identification reaches levels at which object names are retrieved comes also from nonlinguistic experiments. On the one hand, Noizet and Pynte (1976) found that French participants gazed approximately 200 ms longer at objects with multisyllable names (e.g., *hélicoptère*) than at objects with one-syllable names (e.g., *main*) when asked to shift eye gaze to three objects, one after another, and to identify them silently. Thus, as in Huettig and McQueen (2007), participants were not asked to name these objects, yet robust evidence for object name retrieval was obtained. Meyer, Roelofs, and Levelt (2003, Experiment 2), on the other hand, found no significant latency difference between objects with monosyllabic and disyllabic names in a task in which participants were required to decide as quickly as possible whether a presented item was an object or a pseudoobject—that is, a nonexisting object that resembled a real object (although numerically, participants were 11 ms faster for objects with monosyllabic names). However, it is conceivable that decisions in speeded object verification tasks can be made before information cascades to the phonological level, and thus, the absence of an effect in such tasks does not rule out a cascade in principle (especially in tasks in which participants are put under less time pressure). Zelinsky and Murphy (2000) contrasted objects with one-syllable names (e.g., *ball*) and three-syllable names (e.g., *elephant*). In their Experiment 1, participants were presented with displays of 2 one-syllable and 2 three-syllable objects, which they were allowed to look at as long as they wished. After a 2.5-s blank screen interval, they were presented with one centrally positioned object and had to indicate by pressing a button whether it had been

present in the previous display. Zelinsky and Murphy observed more fixations and longer gaze durations on objects with three-syllable names. In a second experiment, participants were first asked to associate faces with names until all participants could correctly name all of the faces. Then a recognition task and a visual search task followed. In the recognition task (similar to Experiment 1), participants had to study a display with four faces (two faces with one-syllable names and two with three-syllable names) and subsequently had to indicate whether the face had been present in the previous study display. In the search task, participants first saw a target face and then had to indicate whether it was present in the subsequently presented display. An effect of syllable length of the name of the face was observed in the recognition task, but not in the search task. Zelinsky and Murphy concluded that depending on the task people “supplement their visual representations with verbal encoding (p. 130).”

In sum, it is still unclear whether (or at least, to what extent and under which circumstances) preexposure to visual stimuli leads to retrieval of phonological representations associated with those visual stimuli. Here, we present three cross-modal priming experiments that examined the influence of preexposure to pictures and printed words on the speed of spoken word recognition. Our primary interest was in phonological overlap between prime and target. As a control condition, we included prime–target pairs that were semantically related.

The notion that preexposure to related stimuli speeds up retrieval of target stimuli is typically referred to as *priming*. Many forms of priming have been demonstrated. In the field of psycholinguistics, priming is often measured using the lexical decision task. In the lexical decision task, words and nonwords (i.e., wordlike stimuli) are presented, and participants are required to determine, as quickly as possible, whether or not the presented item is an existing word. For example, it has consistently been found that participants’ responses to recognize *butter* as a word after hearing “bread” are faster than to recognize semantically unrelated control words as words. The main assumption is that priming (Warren, 1972) amounts to a degree of preactivation of a target word because of the similarity between the prime and target. The primed word *butter*, for example, is assumed to be already active to some degree because of its semantic similarity with *bread* (via featural overlap or an active connection in a semantic network), and thereby, it takes less time for it to reach the threshold for triggering a response (e.g., to decide whether *butter* is an existing word or not).

Although the magnitude of priming tends to be larger within (e.g., auditory word/nonword to auditory word/nonword) modalities (see Carlesimo et al., 2003, for a review), priming has also been observed across modalities (e.g., auditory word/nonword to printed word/nonword). Cross-modal priming (the form of priming we are concerned with here) is assumed to reflect recoding of a stimulus presented in one modality

(e.g., the auditory modality) to another (e.g., visual) modality (e.g., Downes et al., 1996; McClelland & Pring, 1991). There have been many cross-modal priming studies in psycholinguistics (typically, with spoken word primes and printed word or nonword targets; e.g., Gow & Gordon, 1995; Shillcock, 1990; Swinney, 1979; Tabossi, Burani, & Scott, 1995); however, there have been few relevant studies that have explored explicitly whether *visual stimuli* facilitate or interfere with the recognition of *spoken words*.

If there is default retrieval of conceptual knowledge from picture displays, as appears to be the case (Dell’Acqua & Grainger, 1999; see also Biederman, Bickler, Teitelbaum, & Klatsky, 1988; Biederman, Mezzanotte, & Rabinowitz, 1982; Intraub, 1984; Loftus & Ginn, 1984; Potter, 1976; Smith & McGee, 1980), semantic (category) overlap between picture primes and auditory targets should *facilitate* target responses. For example, deciding that *zwaard*, *sword*, is a word (in Dutch) would be easier if a picture of an arrow has already caused weapon knowledge to be retrieved.

The key prediction is that onset phonological overlap between picture names and targets should interfere with target responses if pictures routinely call up their names. For example, deciding that *pijn*, *pain*, is a Dutch word should be harder if seeing a picture of an arrow typically results in retrieval of its name *pijl*, a phonological competitor of *pijn*.

The present study

We conducted three cross-modal priming experiments to investigate this issue. In Experiment 1, Dutch participants saw picture primes for 1 s and then made lexical decisions to spoken Dutch word and nonword targets. Phonological and semantic overlap between primes and targets was manipulated. Critically, participants were not instructed to process the pictures in any way, and proportions of different trial types were selected to discourage participants from naming the pictures.

Cross-modal priming effects involving printed and spoken words have previously been observed. There is facilitation of printed word lexical decisions by semantically related spoken primes (a seminal finding; e.g., Swinney, Onifer, Prather, & Hirshkowitz, 1979) and interference of printed word lexical decisions by spoken-word primes with the same phonological onsets as the targets (e.g., Marslen-Wilson, 1990). Interference of responses to auditory targets by phonologically related printed word primes has also been found (in shadowing for items sharing onset phonemes, Slowiaczek & Hamburger, 1992; and in lexical decision for primes that are pseudohomophones of the targets, Grainger, Diependaele, Spinelli, Ferrand, & Farioli, 2003). But semantic facilitation with printed primes and spoken targets has not, to our knowledge, been documented. Experiment 2 therefore sought to establish whether

this kind of semantic facilitation can indeed be observed and whether, in auditory lexical decision (as in shadowing), there is phonological interference when primes and targets overlap phonologically only at onset.

In eye-tracking experiments, however, target words typically appear in sentence contexts, and pictures appear at sentence onset. Experiment 3 was thus designed to investigate whether phonological and semantic activation is detectable with more time between prime pictures and target words. The experiment was identical to Experiment 1, but the auditory targets appeared at the end of a short carrier sentence. In Experiment 3a, a single picture prime was presented (as in Experiment 1). In Experiment 3b, four pictures were presented (the prime picture and three completely unrelated pictures) in order to make it as similar as possible to typical eye-tracking studies.

Experiment 1

Method

Participants

Participants were 32 paid members of the MPI for Psycholinguistics subject panel. All were native speakers of Dutch, none had known hearing problems, and all had normal or corrected-to-normal visual acuity.

Materials and design

Forty-eight experimental picture primes (see Fig. 1 and the Appendix Table 2) were paired with a semantically related target (a category coordinate; e.g., *pijl*–*zwaard*, *arrow*–*sword*) and a phonologically related target (beginning with the same sounds, at least up to the vowel of the first or only syllable; e.g., *pijl*–*pijn*, *arrow*–*pain*; mean onset overlap = 2.3 phonemes). These 96 related targets also served as unrelated targets: Each pair of them was combined with another, unrelated picture (e.g., *zwaard* and *pijn* combined with *huis*, *house*). There were also 48 fillers (prime pictures combined with nonword targets that were phonologically unrelated to the prime names). Participants were assigned to four counterbalanced groups. Each group were presented with all 48 experimental primes, paired with 12 semantic, 12 phonological, and 24 unrelated word targets (such that, across groups, all four prime–target combinations for each prime were presented), mixed randomly with all 48 fillers.

Procedure

Primes (for experimental and filler trials) were selected from the MPI for Psycholinguistics line-drawing database. The

auditory materials were spoken in isolation in a sound-damped room by a female native speaker of Dutch who was unaware of the experiment's purpose. Digital recordings of each stimulus were selected and measured using Praat (Boersma & Weenink, 2005). Mean durations were 604 ms (semantically related targets) and 590 ms (phonologically related targets).

Participants were tested individually in a sound-damped booth. They were told that they would see pictures of objects, each followed by a spoken word. They were asked to decide whether the spoken items were real Dutch words and to respond as quickly and as accurately as possible by pressing “yes” or “no” response keys. They were not instructed to name the pictures. Stimulus presentation and timing were controlled by a computer. The picture primes were presented for 1,000 ms on the computer screen and were followed immediately by the spoken targets. RTs were recorded from target onsets, but target durations were subtracted from the raw RTs prior to analysis to control for effects of word duration on RT. Responses above the 98th or below the 2nd percentile were identified as outliers and removed from the analysis.

Results

Mean correct lexical decision RTs and mean errors are shown in Table 1, and RT priming effects (differences between the related and unrelated conditions) are plotted in Fig. 2. As can be seen, there was semantic facilitation (faster and more accurate responses to words preceded by semantically related picture primes than to those preceded by unrelated primes) and phonological interference (responses to targets were slower and less accurate after phonologically related than after unrelated picture primes). There were two factors in the analyses: prime–target relationship type (semantic or phonological; within subjects but between items) and relatedness (related vs. unrelated; within subjects and items). In the RT analysis, there was a main effect of relationship type, $F_1(1, 28) = 52.32, p < .001, \eta^2_p = 1,678$; $F_2(1, 88) = 7.88, p < .01, \eta^2_p = 17,415$, but not of relatedness, F_1 and $F_2 < 1$, and an interaction, $F_1(1, 28) = 11.98, p < .01, \eta^2_p = 1,330$; $F_2(1, 88) = 11.55, p < .01, \eta^2_p = 1,591$. Pairwise comparisons indicated that there was both semantic facilitation (e.g., faster “yes” decisions to *zwaard* [*sword*] after seeing an arrow than after seeing a house), $F_1(1, 28) = 5.50, p < .05, \eta^2_p = 1,295$; $F_2(1, 44) = 5.29, p < .05, \eta^2_p = 1,627$, and phonological interference (e.g., slower “yes” decisions to *pijn* after seeing an arrow [a *pijl*] than after seeing a house), $F_1(1, 28) = 5.98, p < .05, \eta^2_p = 1,481$; $F_2(1, 44) = 6.29, p < .05, \eta^2_p = 1,554$.

In the error analysis there was again an effect of relationship type, $F_1(1, 28) = 25.65, p < .001, \eta^2_p = 31$; $F_2(1, 94) = 5.53, p < .05, \eta^2_p = 212$, but not of relatedness, F_1 and $F_2 < 1$. These factors interacted, $F_1(1, 28) = 12.55, p < .005, \eta^2_p = 39$;

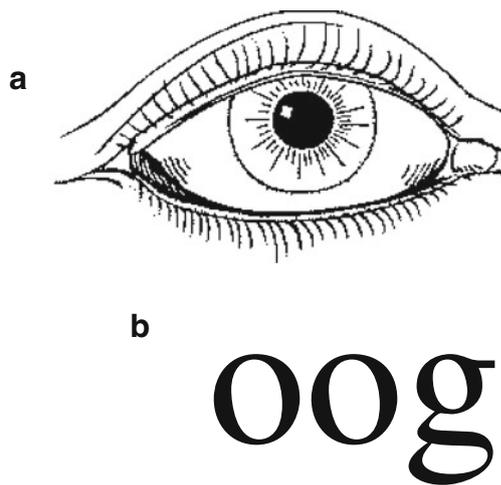


Fig. 1 Visual stimuli. **a** Example of a picture used in Experiments 1 and 3. **b** Example of a printed word used in Experiment 2

$F2(1, 94) = 18.26, p < .001, \eta^2_p = .40$. This interaction reflected facilitation in the semantic condition (more accurate responses in the related condition), $F1(1, 28) = 7.52, p < .05, \eta^2_p = .21$; $F2(1, 47) = 9.40, p < .005, \eta^2_p = .25$, and interference in the phonological condition (fewer accurate responses in the related condition), $F1(1, 28) = 6.27, p < .05, \eta^2_p = .56$; $F2(1, 47) = 9.54, p < .005, \eta^2_p = .55$.

Discussion

Semantic category overlap between pictures and spoken words facilitated target responses. More important, phonological onset overlap between picture names and spoken words interfered with target responses. The latter effect arose in a situation where

Table 1 Mean lexical decision reaction times (RTs, in milliseconds; with standard errors of the means in parentheses) between related and unrelated targets (from target offset) and mean errors (in percentages; with standard errors of the means in parentheses)

Experiment	Measure	Condition			
		Semantic		Phonological	
		Related	Unrelated	Related	Unrelated
1	Mean RT	221 (18)	242 (18)	296 (19)	272 (19)
	Mean Error	1.0 (0.5)	4.2 (1.1)	9.9 (1.8)	5.2 (1.0)
2	Mean RT	228 (17)	247 (14)	301 (15)	287 (15)
	Mean Error	1.0 (0.5)	2.1 (0.6)	6.5 (1.1)	5.5 (1.3)
3a	Mean RT	206 (14)	221 (15)	286 (18)	256 (17)
	Mean Error	1.6 (0.7)	1.6 (0.7)	6.5 (1.6)	4.9 (0.8)
3b	Mean RT	314 (22)	329 (24)	383 (27)	360 (29)
	Mean Error	1.3 (0.7)	2.1 (0.7)	6.5 (1.3)	4.9 (1.2)

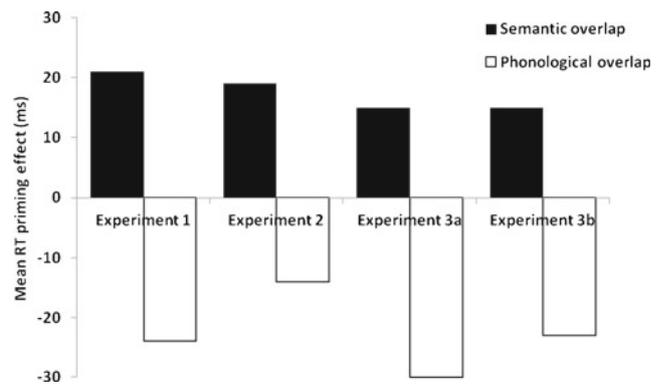


Fig. 2 Mean priming effects (related – unrelated differences in reaction times, in milliseconds) in the semantic and phonological overlap conditions across experiments

subvocal naming of primes would not make lexical decisions easier (50 % of all targets were phonologically unrelated non-words; 75 % of the word targets were phonologically unrelated to the primes). Naming was thus unlikely to be a deliberate participant strategy.

In Experiment 2, we replaced the picture primes with their printed Dutch names. Otherwise, the experiment was identical to Experiment 1. The primary goal of Experiment 2 was to test whether priming effects with printed-word primes are equivalent (in nature and magnitude) to those observed with matched picture primes. It is often assumed that presentation of a printed word results in fast and efficient retrieval of its phonology; indeed, this may be necessary for visual-word recognition (e.g., Rastle & Brysbaert, 2006). Similarly, many researchers have claimed that presentation of a printed word results in fast and efficient retrieval of semantic knowledge (e.g., Neely, 1991). If the semantic facilitation and phonological interference effects found in Experiment 1 were replicated in Experiment 2, this would suggest that picture recognition reflects similar retrieval operations to those postulated for visual-word recognition.

Experiment 2

Method

Thirty-two new participants from the population used in Experiment 1 took part. The 48 experimental and 48 filler picture primes from Experiment 1 were replaced by their printed Dutch names, presented in 72-point Times New Roman font (see Fig. 1). All other materials and the procedure were as before.

Results

Mean correct RTs and error rates are given in Table 1, and RT priming effects are shown in Fig. 2. Semantically related

primes again speeded up lexical decisions, and phonologically related primes again slowed them down. In the RT analysis, there was an effect of relationship type, $F_1(1, 28) = 67.10$, $p < .001$, $\eta^2_p = 1,516$; $F_2(1, 88) = 12.20$, $p < .001$, $\eta^2_p = 14, 501$, but not of relatedness, F_1 and $F_2 < 1$, and these factors interacted, $F_1(1, 28) = 6.32$, $p < .05$, $\eta^2_p = 1,314$; $F_2(1, 88) = 7.03$, $p < .01$, $\eta^2_p = 2,491$. Pairwise comparisons revealed weak semantic facilitation, $F_1(1, 28) = 3.58$, $p = .07$, $\eta^2_p = 1,525$; $F_2(1, 44) = 4.21$, $p < .05$, $\eta^2_p = 2,511$, and weak phonological interference, $F_1(1, 28) = 2.38$, $p = .13$, $\eta^2_p = 1,272$; $F_2(1, 44) = 2.87$, $p = .10$, $\eta^2_p = 2,470$. But combined analyses of Experiments 1 and 2 revealed no interaction of experiment and relatedness in either the semantic or the phonological condition (all $F_s < 1$). In the errors, although there was again a main effect of relationship type, $F_1(1, 28) = 27.16$, $p < .001$, $\eta^2_p = 23$; $F_2(1, 94) = 4.41$, $p < .05$, $\eta^2_p = 214$, there was no effect of relatedness and no interaction (all $F_s < 2$). Pairwise analyses of the errors confirmed that there was neither semantic facilitation nor phonological interference (all $F_s \leq 2$).

Discussion

The results with prime words paralleled those with prime pictures in Experiment 1. Furthermore, the magnitudes of the semantic facilitation and phonological interference effects in RTs that arose from pictures were very similar to those that arose from printed words. Evidence from masked phonological priming (Rastle & Brysbaert, 2006), for instance, suggests that retrieval of phonology from print is fast and efficient. The similarity in results across experiments suggests that retrieval of phonology from pictures can be just as fast and efficient. As the error analysis revealed, however, phonological interference was greater with pictures than with words. Retrieval of word phonology upon seeing a picture is thus at least as strong as when the printed form of that word is seen. Furthermore, from a strategic point of view, processing of primes to the phonological level was as disadvantageous when the primes were pictures as when they were words.

In Experiment 3, we investigated whether priming is still detectable with more time between picture primes and target words, as in typical visual-world eye-tracking experiments.

Experiment 3

Method

New participants from the same population as before were recruited (32 in each subexperiment). The materials in

Experiment 3a were identical to those in Experiment 1, except that the auditory targets appeared at the end of a short carrier phrase (“*Het volgende woord is . . .*,” *The next word is . . .*) rather than in isolation. The tokens of the targets used in Experiment 1 were spliced, using Praat, onto the end of the carrier (which was spoken by the same talker and was 1, 400 ms long). The Experiment 3b materials were identical to those in Experiment 3a, except that four pictures were presented on each trial, rather than one. These visual displays contained images of four spatially distinct objects, one in each quadrant of the computer screen. On each trial, the display included the same prime picture as in the corresponding trial in Experiment 3a and three unrelated pictures that did not share any semantic or phonological similarity with the targets. Picture locations were randomized. The same procedure as before was used, with two exceptions. First, the displays remained visible until responses were initiated. There was 2,400 ms from onset of the prime display until spoken target onset (1-s display preview plus carrier phrase duration). Second, participants were told to make lexical decisions to the carrier-final stimuli.

Results

Table 1 shows mean RTs and error rates, and Fig. 2 shows the RT priming effects. The pattern of semantic facilitation and phonological interference was replicated again, in both subexperiments. In the RT analysis, there were main effects of subexperiment (responses were faster with single-picture primes than with four-picture primes), $F_1(1, 56) = 12.85$, $p < .001$, $\eta^2_p = 54,254$; $F_2(1, 88) = 254.20$, $p < .001$, $\eta^2_p = 4,256$, and relationship type (responses to the semantic targets were again faster than those to the phonological targets), $F_1(1, 56) = 106.71$, $p < .001$, $\eta^2_p = 1,275$; $F_2(1, 88) = 8.69$, $p < .005$, $\eta^2_p = 31,152$, but not of relatedness, $F_1(1, 56) = 1.34$, $p > .02$, $\eta^2_p = 1,625$; $F_2(1, 88) = 1.34$, $p > .02$, $\eta^2_p = 3,199$. Critically, as in both previous experiments, there was an interaction of relatedness and relationship type, $F_1(1, 56) = 16.43$, $p < .001$, $\eta^2_p = 1,733$; $F_2(1, 44) = 10.50$, $p < .005$, $\eta^2_p = 3,199$. There was no modulation of this interaction across subexperiments, F_1 and $F_2 < 1$, suggesting that RTs did not differ depending on whether one or four pictures were presented. Separate analyses by relationship type once again revealed facilitation for semantically related targets, $F_1(1, 56) = 7.10$, $p < .05$, $\eta^2_p = 1,049$; $F_2(1, 44) = 2.44$, $p = .13$, $\eta^2_p = 2,848$, and interference for phonologically related targets, $F_1(1, 56) = 10.06$, $p < .005$, $\eta^2_p = 2,309$; $F_2(1, 44) = 8.72$, $p < .01$, $\eta^2_p = 3,550$. There was no interaction of subexperiment and relatedness in either of these analyses (all $F_s < 1$). A parallel analysis of the errors revealed no effects of relatedness or subexperiment (all

$F_s < 1$) and no interaction of relatedness with relationship type, $F_1(1, 56) = 2.21, p = .143, \eta_p^2 = .28$; $F_2(1, 94) = 2.51, p = .117, \eta_p^2 = .36$, but there was again an effect of relatedness type, $F_1(1, 56) = 35.93, p < .001, \eta_p^2 = .30$; $F_2(1, 94) = 6.83, p < .05, \eta_p^2 = .236$. Error analyses for each type of relationship separately showed that there was no semantic facilitation (both $F_s < 1$) and no phonological interference, $F_1(1, 56) = 2.19, p = .144, \eta_p^2 = .36$; $F_2(1, 47) = 2.10, p = .154, \eta_p^2 = .56$.

Discussion

In a situation where the targets were delayed by 2.4 s relative to picture onset, there was once again semantic facilitation and, more clearly, phonological interference. These effects did not differ depending on whether one picture (as in cross-modal priming experiments) or four pictures (as in eye-tracking studies) were presented.

General discussion

Spoken Dutch words and nonwords, presented either in isolation (Experiments 1 and 2) or at the end of a short carrier phrase (Experiment 3), were preceded by picture primes (Experiments 1 and 3) or the printed names of those primes (Experiment 2). Primes were phonologically related (beginning with the same sounds), were semantically related to the targets (they were semantic category coordinates), or were unrelated on both dimensions. The number of pictures in the displays was also varied (one vs. four pictures in Experiments 3a and 3b, respectively). Facilitative semantic priming of similar magnitude was found in all experiments. These results suggest that exposure to pictures and printed words typically results in retrieval of conceptual knowledge (whether that is a picture or a printed word) and facilitates the recognition of semantically related spoken words. This conclusion is consistent with previous claims about access of semantic information during picture recognition (Dell'Acqua & Grainger, 1999; see also Biederman et al., 1988; Biederman et al., 1982; Intraub, 1984; Loftus & Ginn, 1984; Potter, 1976; Smith & McGee, 1980) and visual-word recognition (e.g., Neely, 1991). The key theoretical question concerned phonological overlap. Phonological interference was found in all experiments and was again of similar magnitude across experiments. These lexical decision findings converge with our previous eye-tracking results (Huettig & McQueen, 2007). Both methods provide evidence that phonological information can be retrieved during picture viewing.

Before we discuss the theoretical implications of the results, we will address two potential caveats. One might argue

that we obtained phonological interference because participants named the prime pictures strategically during the auditory lexical decision task and that, without such strategic naming, object names are not routinely retrieved. We consider this possibility very unlikely. First, the present results indicate that seeing visual objects can result in phonological interference even when participants are not explicitly instructed to name those objects and even when a strategy of implicit naming would tend to make it harder for participants to perform the lexical decision task (although all nonword targets were phonologically unrelated to the prime names, 75 % of the word targets were also unrelated). Second, in Experiment 3b, we presented participants with four-object displays and obtained essentially the same priming results as in Experiment 3a (in which only one prime picture was presented). This is inconsistent with the notion that an explicit picture-naming strategy during the lexical decision task can account for our data, since participants in Experiment 3b would have had much less time to strategically name the pictures than those in Experiment 3a.

The second alternative account is that the effect in the phonological condition could, in fact, be due to semantic interference. On this account, hearing the (phonological) target word *pijn* (*pain*) activates the phonological cohort competitor *pijl* (*arrow*) via the speech signal (cf. Marslen-Wilson & Zwitserlood, 1989). This will activate the semantic representations of *pijl* (*arrow*), which are already preactivated because of the exposure to the prime picture of the arrow. The enhanced processing of the concept *arrow* from two sources (picture prime and spoken target word) may then interfere with the decision about whether *pijn* (an unrelated concept) is a real word. We cannot rule out this possibility with certainty. However, note that there is also activation of unrelated concepts and, thus, possible interference in the unrelated condition (e.g., a house as a picture prime and *pijn* as the target). The concept *house* will be activated by the prime, and the concept *arrow* by the target (just as in the related condition). The semantic interference account thus rests on the assumption, untested as far as we know, that one unrelated concept activated by multiple sources interferes more with auditory lexical decision than do two unrelated concepts. In contrast, the phonological interference account rests on the well-established finding, as implemented in all models of spoken-word recognition, that phonologically overlapping words compete with each other (McQueen, Norris, & Cutler, 1994; see McQueen, 2007, for a review). Responses to *pijn*, for example, are slowed down because of enhanced competition between the word forms *pijn* and *pijl*, where *pijl* has been preactivated by the prime picture. We therefore believe that phonological

interference is the more parsimonious explanation of the present data.

We conclude that retrieval of phonological information from pictures (or printed words) does not appear to depend on the participant's intention to name those pictures (or repeat the words). The present results thus join a growing body of data (Griffin & Bock, 1998; Huettig & McQueen, 2007; Kuipers & La Heij, 2009; Mani & Plunkett, 2010; Meyer, Belke, Telling, & Humphreys, 2007; Meyer & Damian, 2007; Morsella & Miozzo, 2002; Navarrete & Costa, 2005; Peterson & Savoy, 1998; but see Bloem et al., 2004; Damian & Bowers, 2003; Jescheniak et al., 2009) suggesting that information flows from conceptual to phonological levels of processing during object processing. We suggest, therefore, that processing typically does extend to the phonological level even with passive picture viewing. By extension, phonological effects with passive viewing in visual-world eye-tracking studies also likely reflect routine retrieval processes rather than deliberate (subvocal) naming (Huettig & McQueen, 2007; Huettig et al., 2011b).

How extensive is phonological activation during object processing likely to be, for instance, in cluttered scenes with many objects? It is important to emphasize that we do not claim that name retrieval during object identification cannot be under attentional control. There are alternative accounts of how attentional control might mediate retrieval of phonological information. One possibility is that, in the default state of the system, there is no retrieval of phonological representations for object names the individual does not intend to utter but that task demands can sometimes cause information flow to the phonological level to be switched on. Alternatively, retrieval of phonological information may typically occur, but task demands can cause this to be switched off. The present data suggest that the latter account is more plausible. Participants had no reason to retrieve phonological representations, since name retrieval would tend to make it harder to make lexical decisions (and it did indeed do so). It is thus more likely that picture names are routinely retrieved, unless doing so would interfere more substantially with the task at hand. Under such circumstances, information flow to the phonological level may be switched off. One obvious situation where attentional control could limit phonological retrieval is when the current task is to speak. As already noted, retrieval of the phonology of words that the speaker does not intend to utter is likely to make speaking harder. Hence, during speaking, retrieval of phonology may tend to be limited to the to-be-named word (Levelt et al., 1999; but see La Heij, Starreveld, & Kuipers, 2007, for the contrasting view that in object-naming tasks, it may be hard to prevent cascaded processing of context stimuli). This argument is further supported by experiments on dual-task performance, which show that retrieval of phonological

word forms is open to cognitive control (Ferreira & Pashler, 2002).

Moreover, phonological effects in visual-world situations are most likely mediated by working memory (cf. Huettig et al., 2011a). Although our results suggest that information flows routinely from conceptual to phonological levels of processing during object processing, this does not mean that phonological competition effects during on-line word-object mapping are limitless. Working memory appears to have a limit of four objects (Cowan, 2001; see also Atkinson, Campbell, & Francis, 1976; Burkell & Pylyshyn, 1997; Luck & Vogel, 1997; Mandler & Shebo, 1982; Pashler, 1988; Phillips, 1974; Pylyshyn & Storm, 1988; Sperling, 1960; Trick & Pylyshyn, 1994; Yantis & Johnson, 1990). If the notion is correct that working memory is the nexus that binds visual, linguistic (e.g., phonological), and spatial information (Huettig et al., 2011a; cf. Ferreira, Apel, & Henderson, 2008; Knoeferle & Crocker, 2007; Richardson & Spivey, 2000; Spivey & Geng, 2001; Spivey, Richardson, & Fitneva, 2004), then phonological word-object mapping will be subject to working memory capacity limitations, particularly in situations such as cluttered scenes with numerous objects.

To conclude, our data are important with regard to the question of whether (and under which circumstances) preexposure to visual stimuli (as in visual-world eye-tracking experiments and, indeed, in many real-world situations) primes the retrieval of spoken words. Our findings provide further evidence that viewing pictures facilitates the retrieval of semantically related spoken words. With regard to phonological information (i.e., pictures' names), our data indicate that attentional factors, driven by the demands of the participant's task, determine retrieval. In the situation examined here, participants had no reason to block retrieval of picture names, and thus, in spite of the fact that this made it a little harder for them to make lexical decisions, they went ahead and accessed those picture names. We suggest that retrieval of picture names is not mandatory but will typically happen unless control processes intervene. Similarly, retrieval of semantic and phonological information from printed words occurs (and primes the recognition of semantically and phonologically related spoken words) unless task demands block such retrieval (cf. Frost, 1998; Huettig & McQueen, 2011; Stolz & Besner, 1996).

Acknowledgments We thank Laurence Bruggeman for assistance with running the experiments and Jennifer Stolz, Wido La Heij, and Gary Lupyan for constructive feedback on a previous version of the manuscript. Correspondence may be addressed to: Falk Huettig, Max Planck Institute for Psycholinguistics, PO Box 310, 6500 AH Nijmegen, The Netherlands, or by email to either author: falk.huettig@mpi.nl; j.mcqueen@pwo.ru.nl

Appendix

Table 2 Critical stimuli

Picture primes	Spoken targets	
	Semantic	Phonological
boom (tree)	rivier (river)	boot (boat)
trommel (drum)	klarinet (clarinet)	trolley (trolley)
rok (skirt)	jeans (jeans)	rots (rock)
zaag (saw)	moersleutel (spanner)	zaad (seed)
vos (fox)	bever (beaver)	vonk (spark)
oog (eye)	lever (liver)	oom (uncle)
hond (dog)	geit (goat)	hof (court)
hart (heart)	been (leg)	ham (ham)
vlinder (butterfly)	kakkerlak (cockroach)	flits (flash)
aardbei (strawberry)	sinaasappel (orange)	aartsbisschop (archbishop)
pop (doll)	bal (ball)	pot (pot)
spin (spider)	rups (caterpillar)	spits (peak)
vlieger (kite)	adelaar (eagle)	vliering (attic)
leeuw (lion)	rat (rat)	leed (sorrow)
appel (apple)	druiven (grapes)	applaus (applause)
auto (car)	scooter (scooter)	autist (person with autism)
vliegtuig (airplane)	mug (mosquito)	vlierbes (elderberry)
wortel (carrot)	radijs (radish)	wonder (miracle)
sjaal (scarf)	bloes (blouse)	sjabloon (stencil)
laars (boot)	sok (sock)	laan (avenue)
schoen (shoe)	overhemd (shirt)	schoepenrad (turbine)
peer (pear)	framboos (raspberry)	pees (tendon)
potlood (pencil)	tang (pliers)	polder (polder)
citroen (lemon)	kers (cherry)	cipier (jailor)
boek (book)	krant (newspaper)	boer (farmer)
huis (house)	iglo (igloo)	huid (skin)
aardappel (potato)	kool (cabbage)	aarzeling (hesitation)
trompet (trumpet)	viool (violin)	trottoir (pavement)
tent (tent)	flat (apartment)	tekst (text)
arm (arm)	long (lung)	arts (doctor)
klomp (clog)	sandaal (sandal)	klok (clock)
kruiwagen (wheelbarrow)	grasmaaier (lawnmower)	kruimeldief (petty thief)
paprika (pepper)	selderie (celery)	pakpapier (wrapping paper)
boog (bow)	granaat (grenade)	boon (bean)
stoel (chair)	plank (shelf)	stoep (pavement)
schroevendraaier (screwdriver)	beitel (chisel)	schroeiplek (scorch mark)
giraf (giraffe)	hert (deer)	gigolo (gigolo)
kameel (camel)	gorilla (gorilla)	kanaal (canal)
tomaat (tomato)	spinazie (spinach)	toneel (stage)
vork (fork)	kan (jug)	vocht (liquid)
broek (trousers)	muts (hat)	broer (brother)
harp (harp)	trombone (trombone)	hal (hall)
banana (banana)	perzik (peach)	barak (barracks)
bloemkool (cauliflower)	asperge (asparagus)	bloedbad (bloodbath)
ananas (pineapple)	bosbes (blueberry)	animo (eagerness)
fiets (bicycle)	helikopter (helicopter)	fierheid (pride)
hamer (hammer)	meetlat (ruler)	hagel (hail)
pijl (arrow)	zwaard (sword)	pijn (pain)

References

- Anderson, S. E., Chiu, E., Huette, S., & Spivey, M. J. (2011). On the temporal dynamics of language-mediated vision and vision-mediated language. *Acta Psychologica*, *137*(2), 181–189.
- Atkinson, J., Campbell, F. W., & Francis, M. R. (1976). The magic number 4 +/- 0. *Perception*, *5*, 327–334.
- Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*, 143–177.
- 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.
- Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. *Journal of Memory and Language*, *48*, 468–488.
- Bloem, I., van der Boogaard, S., & La Heij, W. (2004). Semantic facilitation and semantic interference in language production: Further evidence for the conceptual selection model of lexical access. *Journal of Memory and Language*, *51*, 307–323.
- Boersma, B., & Weenink, D. (2005). Praat: Doing phonetics by computer (Version 4.3.14) [Computer program]. Retrieved May 26, 2005, from <http://www.praat.org/>
- Burkell, J. A., & Pylyshyn, Z. W. (1997). Searching through subsets: A test of the Visual Indexing Hypothesis. *Spatial Vision*, *11*, 225–258.
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*, 177–208.
- Carlesimo, A., Turriziani, P., Paulesu, E., Gorini, A., Caltagirone, C., Fazio, F., & Perani, D. (2003). Brain activity during intra- and cross-modal priming: New empirical data and review of literature. *Neuropsychologia*, *42*, 14–24.
- Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for the models of lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *26*, 1283–1296.
- Costa, A., Santesteban, M., & Cano, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language*, *94*, 94–103.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185.
- Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychonomic Bulletin & Review*, *10*, 111–117.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283–321.
- Dell'Acqua, R., & Grainger, J. (1999). Unconscious semantic priming from pictures. *Cognition*, *73*, B1–B15.
- Downes, J. J., Davis, E. J., De Mornay Davies, P., Perfect, T. J., Wilson, K., Mayes, A. R., & Sagar, H. J. (1996). Stem-completion priming in Alzheimer's disease: The importance of target word articulation. *Neuropsychologia*, *34*, 63–75.
- Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Sciences*, *12*, 405–410.
- Ferreira, V., & Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology: Human Learning and Memory*, *28*, 1187–1199.
- Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, *123*, 71–99.
- Gow, D. W. J., & Gordon, P. C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 344–359.

- Grainger, J., Diependaele, K., Spinelli, E., Ferrand, L., & Farioli, F. (2003). Masked repetition and phonological priming within and across modalities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 1256–1269.
- Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between processing levels in spoken word production. *Journal of Memory and Language*, *38*, 313–338.
- Harley, T. A. (1993). Phonological activation of semantic competitors during lexical access in speech production. *Language and Cognitive Processes*, *8*, 291–309.
- Huetting, 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*, 460–482.
- Huetting, F., & McQueen, J. M. (2011). The nature of the visual environment induces implicit biases during language-mediated visual search. *Memory and Cognition*, *39*, 1068–1084.
- Huetting, F., Mishra, R. K., & Olivers, C. N. (2012). Mechanisms and representations of language-mediated visual attention. *Frontiers in Psychology*, *2*, 394.
- Huetting, F., Olivers, C. N. L., & Hartsuiker, R. J. (2011a). Looking, language, and memory: Bridging research from the visual world and visual search paradigms. *Acta Psychologica*, *137*, 138–150.
- Huetting, F., Rommers, J., & Meyer, A. S. (2011b). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, *137*, 151–171.
- 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.
- Jescheniak, J. D., Oppermann, F., Hantsch, A., Wagner, V., Maedebach, A., & Schriefers, H. (2009). Do perceived context pictures automatically activate their phonological code? *Experimental Psychology*, *56*, 56–65.
- Knoeferle, P., & Crocker, M. W. (2007). The influence of recent scene events on spoken comprehension: Evidence from eye-movements. *Journal of Memory and Language*, *57*, 519–543.
- Kuipers, J.-R., & La Heij, W. (2009). The limitations of cascading in the speech production system. *Language and Cognitive Processes*, *24*, 120–135.
- La Heij, W., Starreveld, P. A., & Kuipers, J.-R. (2007). Structural complexity is not the (big) issue: A reply to Roelofs (2007). *Language and Cognitive Processes*, *22*, 1261–1280.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75.
- Loftus, G. R., & Ginn, M. (1984). Perceptual and conceptual masking of pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 435–441.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.
- Lupyan, G. (2012a). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Cognition*, *3*(54).
- Lupyan, G. (2012b). What do words do? Towards a theory of language-augmented thought. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 57, pp. 255–297). Academic Press.
- Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, *111*, 1–22.
- Marslen-Wilson, W. (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of Language Processing: The Second Sperlonga Meeting* (pp. 148–172). Cambridge, MA: MIT Press.
- Meyer, A. S., Belke, E., Telling, A., & Humphreys, G. W. (2007). Early activation of object names in visual search. *Psychonomic Bulletin and Review*, *14*, 710–716.
- Meyer, A. S., & Damian, M. F. (2007). Activation of distractor names in the picture-picture interference paradigm. *Memory and Cognition*, *35*, 494–503.
- Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, *48*, 131–147.
- McClelland, A. G. R., & Pring, L. (1991). An investigation of cross-modality effects in implicit and explicit memory. *Quarterly Journal of Experimental Psychology*, *43A*, 19–33.
- McQueen, J. M. (2007). Eight questions about spoken-word recognition. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 37–53). Oxford: Oxford University Press.
- McQueen, J. M., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 621–638.
- Mani, N., & Plunkett, K. (2010). In the infant's mind's ear: Evidence for implicit naming in infancy. *Psychological Science*, *21*, 908–913.
- Marslen-Wilson, W. D., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 576–585.
- Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 555–563.
- Navarrete, N., & Costa, A. (2005). Phonological activation of ignored pictures: Further evidence for a cascade model of lexical access. *Journal of Memory and Language*, *53*, 359–377.
- Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic Processes in Reading: Visual Word Recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.
- Noizet, G., & Pynte, J. (1976). Implicit labeling and readiness for pronunciation during the perceptual process. *Perception*, *5*, 217–223.
- Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*, 369–378.
- Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological coding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 539–557.
- Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, *16*, 283–290.
- Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *2*, 509–522.
- Pylyshyn, Z., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*, 179–197.
- Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*, 460–499.
- Rastle, K., & Brysbaert, M. (2006). Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychology*, *53*, 97–145.
- Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood Squares: Looking at things that aren't there anymore. *Cognition*, *76*, 269–295.
- Shillcock, R. (1990). Lexical hypotheses in continuous speech. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing* (pp. 24–49). Cambridge, MA: MIT Press.
- Slowiaczek, L. M., & Hamburger, M. B. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1239–1250.
- Smith, M. C., & McGee, L. E. (1980). Tracing the time-course of picture-word processing. *Journal of Experimental Psychology: General*, *109*, 373–392.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, *74*(11), whole issue.
- Spivey, M., & Geng, J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, *65*, 235–241.

- Spivey, M. J., Richardson, D. C., & Fitneva, S. A. (2004). Memory outside of the brain: Oculomotor indexes to visual and linguistic information. In J. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.
- Stolz, J. A., & Besner, D. (1996). The role of set in visual word recognition: Activation and activation blocking as nonautomatic processes. *Journal of Experimental Psychology: Human Perception & Performance*, *22*, 1166–1177.
- Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*, 645–659.
- Swinney, D., Onifer, W., Prather, P., & Hirshkowitz, M. (1979). Semantic facilitation across modalities in the processing of individual words and sentences. *Memory & Cognition*, *7*, 159–165.
- Tabossi, P., Burani, C., & Scott, D. (1995). Word identification in fluent speech. *Journal of Memory and Language*, *34*, 440–467.
- Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, *101*, 80–102.
- Warren, R. E. (1972). Stimulus encoding and memory. *Journal of Experimental Psychology*, *94*, 90–100.
- Yantis, S., & Johnson, D. N. (1990). Mechanisms of attentional priority. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 812–825.
- Zelinsky, G. J., & Murphy, G. L. (2000). Synchronizing visual and language processing: An effect of object name length on oculomotor behavior. *Psychological Science*, *11*, 125–131.