Capacity Demands of Phoneme Selection in Word Production: New Evidence From Dual-Task Experiments

Amy E. Cook and Antje S. Meyer
University of Birmingham

Three dual-task experiments investigated the capacity demands of phoneme selection in picture naming. On each trial, participants named a target picture (Task 1) and carried out a tone discrimination task (Task 2). To vary the time required for phoneme selection, the authors combined the targets with phonologically related or unrelated distractor pictures (Experiment 1) or words, which were clearly visible (Experiment 2) or masked (Experiment 3). When pictures or masked words were presented, the tone discrimination and picture naming latencies were shorter in the related condition than in the unrelated condition, which indicates that phoneme selection requires central processing capacity. However, when the distractor words were clearly visible, the facilitatory effect was confined to the picture naming latencies. This pattern arose because the visible related distractor words facilitated phoneme selection but slowed down speech monitoring processes that had to be completed before the response to the tone could be selected.

Keywords: word production, picture naming, phoneme selection, dual-task paradigm, speech monitoring

The present research concerns the capacity demands of speech production. Planning the content of utterances—deciding what to say and how to order the information to be conveyed—can be cognitively demanding (e.g., Bock, 1982; Levelt, 1989). This is probably the main reason why speakers pause or hesitate, repair or rephrase their utterances (e.g., Clark & Wasow, 1998; Goldman-Eisler, 1968; Maclay & Osgood, 1959; see also Kubose et al., 2006; Strayer & Drews, 2007). By contrast, speakers do not normally experience linguistic planning—selecting words and combining them into sentences—as particularly effortful, and the received view in the literature appears to be that these processes are considerably less demanding than is planning the content of utterances (Bock, 1982; Levelt, 1989).

Apart from analyses of hesitations and pauses (for a review see Levelt, 1989), few studies have addressed the capacity demands of speech production. The experiments reported below concern a key component of this process: access to the mental lexicon. More specifically, we examined whether selecting the phonemes constituting object names requires cognitive capacity. Before turning to the experiments, we briefly review the main characteristics of current models of lexical access and describe the paradigm we used and a closely related study by Ferreira and Pashler (2002) that motivated the present research.

An Overview of Lexical Access

Word production involves three distinct types of processes: prelinguistic planning processes, lexical access, and postlexical phonetic and articulatory planning processes (for reviews see Johnson, Paivio, & Clark, 1996; Rapp & Goldrick, 2000, 2006; Vigliocco & Hartsuiker, 2002). In object naming, the prelinguistic planning processes are the visual and conceptual processes leading to the identification of the object. Lexical access is the retrieval of the object name from the mental lexicon; and during the phonetic and articulatory planning processes, speakers select and combine the articulatory gestures leading to overt speech.

Most models of lexical access postulate separate representations of semantic, syntactic, and word form information. The models differ in the properties ascribed to these representations and in the ways the representations are connected. In the independent network model proposed by Caramazza (1997; Caramazza & Miozzo, 1997), activation spreads in parallel from a lexical–semantic network to (a) a syntactic network, (b) phonological lexemes, and (c) orthographic representations. Phonological lexemes are unitary representations of word forms. They are linked to the syntactic properties of the words, specified in the syntactic network, and to units representing individual phonemes. An important property of this model is that lexical–semantic units have direct links to both syntactic and word form representations. By contrast, in the model proposed by Levelt, Roelofs, and Meyer (1999; see also Roelofs, 1992, 1997), semantic representations (called lexical concepts in that model) have direct links to only one other type of units, called lemmas, which are linked to (a) nodes specifying the syntactic features of words, (b) nodes specifying the morphological representations (which are similar to Caramazza's phonological lexemes), and (c) nodes representing the orthographic forms of words. Morphological representations are linked to units representing individual phonemes. The family of models proposed by Dell and

Amy E. Cook and Antje S. Meyer, School of Psychology, University of Birmingham, Birmingham, United Kingdom.

Amy E. Cook is now at the Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom.

Correspondence concerning this article should be addressed to Antje S. Meyer, School of Psychology, University of Birmingham, Edgbaston, Birmingham B30 1LH, United Kingdom. E-mail: a.s.meyer@bham.ac.uk

colleagues (Dell, 1986; Dell, Burger, & Svec, 1997; Foygel & Dell, 2000) has a similar architecture.

Models of lexical access differ not only in the processing units they postulate and the links between them, but also in the assumptions about the information flow during lexical access (for further discussion see Morsella & Miozzo, 2002; Rapp & Goldrick, 2000). The model proposed by Levelt and colleagues (Levelt et al., 1999; Roelofs, 1992, 1997) is a serial stage model (for related models see also Bloem & La Heij, 2003; Starreveld & La Heij, 1996), which means that information about the morphological and phonological properties of a word becomes available only after the corresponding lemma has been selected. In the other models mentioned above, activation spreads continuously across the lexical network so that word form information can become activated before the semantic processing of a word has been completed. In some models, including those proposed by Dell and colleagues (Dell, 1986; Dell et al., 1997; Foygel & Dell, 2000), the links between units are bidirectional, allowing for forward and limited backward flow of activation.

For the present purposes, the differences between the models are less important than the communalities. All models assume that lexical access can be distinguished from conceptual and from phonetic and articulatory planning processes, and that it consists of several distinct processing steps during which different types of information about a word are retrieved. According to all models, speakers generate the phonological representations of words by selecting and combining individual phonemes. This assumption—that phonological forms are generated out of individual phonemes rather than being retrieved as units—is well supported by results of speech error analyses and experimental studies (e.g., A. S. Meyer & Belke, 2007; Wilshire & Saffran, 2005). In their study of the capacity demands of lexical access, Ferreira and Pashler (2002) used a novel version of the dual-task paradigm, which is described in the next section.

The Dual-Task Paradigm

On each trial of a dual-task experiment, participants carry out two tasks (Task 1 and Task 2) by responding as quickly as possible to two stimuli. The stimulus onset asynchrony (SOA) is varied (for reviews see Navon & Miller, 2002; Pashler, 1994). A ubiquitous finding, the PRP effect, is that the mean Task 2 reaction time increases with decreasing SOA, whereas the mean Task 1 reaction time remains fairly constant across SOAs (e.g., Kahneman, 1973; Telford, 1931; Welford, 1952, 1959). This pattern has been attributed to the existence of a central capacity bottleneck (e.g., Johnston, McCann, & Remington, 1995; Pashler, 1984, 1994). According to this account, the preparation of a response requires the engagement of central decision or response selection processes, which are capacity demanding and can be devoted to only one task at a time. These capacity demanding processes are preceded by stimulus recognition processes and followed by response execution processes, which are not subject to capacity restrictions (e.g., McCann & Johnston, 1992). If two tasks are to be carried out in succession, the capacity demanding processes for Task 2 can be initiated only after the corresponding processes for Task 1 have been completed. At short SOAs this can lead to a bottleneck effect: The prebottleneck processing of the second stimulus is likely to be completed before the central bottleneck processes for Task 1, and

therefore the central bottleneck processes for Task 2 must be postponed until the central processes for Task 1 have been completed. This situation is less likely to arise at longer SOAs, which is why the mean response latency for Task 2 tends to decrease with increasing SOA. The response latency for Task 1 tends to be unaffected by the SOA because this task typically claims the bottleneck first.

In order to determine whether a component of Task 1 requires central processing capacity, one can vary its duration and observe whether this affects the reaction time for only that task or whether, at short SOAs, it also affects the reaction time for Task 2. If a component is part of the central bottleneck processes or precedes these processes, prolonging its duration should prolong the delay of the central processes for Task 2 and delay the Task 2 reaction time. By contrast, if a component is part of the postbottleneck processes of Task 1, variations in its duration should not affect the Task 2 reaction times. This is because the postbottleneck processes for Task 1 can run in parallel with the central bottleneck processes for Task 2.

A theoretical alternative to the central bottleneck model is that central processing capacity need not be allocated to only one task at a time but can be distributed over two tasks (Kahneman, 1973; McLeod, 1977; Navon & Gopher, 1980; Navon & Miller, 2002; Tombu & Jolicoeur, 2003, 2005; see also D. E. Meyer & Kieras, 1997a, 1997b, for a related theory). When two tasks are carried out simultaneously, they compete for limited resources. Compared with a single-task situation, this is likely to lead to a performance decrement for at least one of the tasks (Tombu & Jolicoeur, 2003; but see Schumacher et al., 2001). On the shared capacity view, the typical PRP pattern—the decrease of the reaction time for Task 2 but not Task 1 with increasing SOA—arises because participants tend to allocate most of their processing resources to the task associated with the first stimulus.

A comparison and appraisal of the various versions of bottleneck and shared capacity models is beyond the scope of the present article (for in-depth discussion see Hazeltine, Ruthruff, & Remington, 2006; Pashler, 1994, 1998). Many of the key findings obtained in the dual-task paradigm can be explained within both frameworks (Navon & Miller, 2002; Tombu & Jolicoeur, 2003). For the present purposes it is important to note that the shared capacity framework also predicts that at short SOAs variations in the time required for Task 1 processing components that are high in capacity demands will carry forward to Task 2 latencies. By contrast, variations in the time needed for response execution processes for Task 1 that require little or no processing capacity should have a measurable effect on only the Task 1 but not the Task 2 response latencies.

Ferreira and Pashler's (2002) Study of Capacity Demands in Lexical Access

In their study of the capacity demands of lexical access, Ferreira and Pashler (2002) adopted a working model according to which a word is accessed in three steps: the selection of a syntactically specified lemma, the selection of a unitary phonological word form, and the selection of the word's phonemes. To assess the capacity demands of each of these steps, they carried out two dual-task experiments in which they combined picture naming (Task 1) with a tone discrimination task (Task 2). In their first experiment, they varied the time required for lemma selection by

presenting the target pictures following context sentences such as "Bob was tired so he went to . . ." or "She saw a picture of a . . .," which either strongly or weakly constrained the nature of the picture. In addition, they varied the time required for phonological word form selection by presenting pictures with high-frequency or low-frequency names. Tones were presented 50, 150, or 900 ms after picture onset.

As expected, the tone discrimination (Task 2) latencies increased with decreasing picture–tone SOA, which replicates the classic PRP effect. The picture naming latencies were shorter after highly constraining than less constraining sentences. In addition, there was an interaction of context and morpheme frequency, with the frequency effect being present in only the low-constraint condition. Importantly, the tone discrimination latencies showed the same pattern. Thus, differences in the time required for lemma selection and for word form selection led to corresponding differences in the tone discrimination latencies. Ferreira and Pashler (2002) concluded that lemma and word form selection were capacity demanding processes.

In their second experiment, they combined target pictures with written distractor words that were semantically related to the picture names (e.g., *couch–bed*), phonologically related (e.g., *bend–bed*), or unrelated. Semantically related distractor words have been shown to slow down picture naming relative to unrelated ones, whereas phonologically related distractor words facilitate picture naming (Cutting & Ferreira, 1999; Damian & Bowers, 2003; Damian & Martin, 1999; A. S. Meyer & Schriefers, 1991; Schriefers, Meyer, & Levelt, 1990). The semantic interference effect has often been allocated at the level of lemma selection (Bloem & La Heij, 2003; La Heij, Kuipers, & Starreveld, 2006; Roelofs, 1992; but see Finkbeiner & Caramazza, 2006a, 2006b; Mahon, Costa, Peterson, Vargas, & Caramazza, 2007, for an alternative account). The main locus of the phonological effect is likely to be the selection of the target phonemes (A. S. Meyer & Belke, 2007; A. S. Meyer & Schriefers, 1991; Roelofs, 2002; Wilshire & Saffran, 2005). A written or spoken distractor word activates the corresponding phonemes. In the phonologically related condition, some of these phonemes are also part of the phonological representation of the target. Because of the activation that these phonemes receive from the distractor, they can be selected faster than in the unrelated condition, where target and distractor do not have any phonemes in common.

Ferreira and Pashler (2002) replicated the semantic and the phonological effect on the picture naming latencies: Compared with unrelated distractors, semantically related distractors slowed down picture naming, and phonologically related ones facilitated it. However, only the semantic effect carried forward to the tone discrimination latencies. Ferreira and Pashler concluded that lemma selection required central processing capacity, whereas phoneme selection was a response execution process, which could be carried out in parallel with the selection of the response to the tone.

The results are important for several reasons: First, they demonstrate the usefulness of the dual-task paradigm for studying word production; second, they show that two main components of lexical access—lemma selection and word form selection—require central processing capacity; and third, they suggest that this is not true for phoneme selection. This difference in capacity demands of different components of lexical access is not predicted by any of the models of lexical access mentioned above.

The Present Study

Ferreira and Pashler's (2002) results suggest that the speakers could select the response to the tone while they were selecting the phonemes of the picture name. One might therefore predict that speakers will be able to carry out other capacity demanding activities, including, for instance, utterance planning processes, during phoneme selection as well. However, the results of recent eye tracking experiments do not support this prediction. When speakers name several objects, they typically look at each object in the order of mention (e.g., Griffin, 2001, 2004; Griffin & Bock, 1998; A. S. Meyer, Sleiderink, & Levelt, 1998). The amount of time they spend looking at an object depends on the time they need to identify it and to select the lemma, the phonological word form, and the phonemes of the object name (Korvorst, Roelofs, & Levelt, 2007; A. S. Meyer, Belke, Häcker, & Mortensen, 2007; A. S. Meyer, Roelofs, & Levelt, 2003). In studies by A. S. Meyer and van der Meulen (2000) and by Mortensen, Meyer and Humphreys (in press), the time required for phoneme selection was varied in the same way as in Ferreira and Pashler's (2002) Experiment 2, namely by combining target pictures with phonologically related or unrelated distractor words. In both studies, the speakers looked at them for a shorter time and named the objects faster when related rather than unrelated distractors were presented. One account of these findings is that speakers look at each object and process it with priority until all capacity demanding components of the naming process, which would include phoneme selection (Griffin, 2004; Wheeldon, Meyer, & van der Meulen, 2007; see also Roelofs, 2007), have been completed. Obviously, the eye tracking studies did not directly assess the capacity demands of object naming. Nevertheless, their results and those reported by Ferreira and Pashler (2002) invite different conclusions about the capacity demands of phoneme selection (see also Griffin, 2004). This motivated the present research.

We report three experiments in which we used the same basic paradigm as in Ferreira and Pashler's (2002) Experiment 2. On each trial the participants named a picture (Task 1) and categorized a tone as high, medium, or low in pitch (Task 2). The tones were presented 50, 150, or 900 ms after each picture. The pictures were accompanied by phonologically related, unrelated, or identical distractors. We expected the picture naming latencies to be shorter in the related and identical conditions than in the unrelated condition. The most important question was whether these facilitatory effects would carry forward to the tone discrimination latencies.

In Experiment 1, we used pictorial distractors. They appeared as red line drawings superimposed on green targets. The two pictures had phonologically related or unrelated names (e.g., *bed–bell* or *bed–nut*) or, in the identical condition, were mirror images of each other. The participants had to name the objects shown in green and ignore the objects shown in red.

Several studies (A. S. Meyer & Damian, 2007; Morsella & Miozzo, 2002; Navarrete & Costa, 2005) have shown that, compared with unrelated distractor pictures, phonologically related ones facilitate target naming. This effect can be explained in a similar fashion as the effect of phonologically related distractor words: The distractor picture activates the associated linguistic information, including the phonemes constituting its name. In the related condition, some of the distractor phonemes are also part of the target name. These phonemes benefit from the activation

received from the distractor and can therefore be selected faster than in the unrelated condition.

Given that the effects of distractor pictures on target naming have been shown to be similar to those of distractor words, we expected to replicate the results reported by Ferreira and Pashler (2002): The picture naming latencies should be shorter in the related than in the unrelated condition, whereas the tone discrimination latencies in the two conditions should not differ from each other.

The identical condition was included as a benchmark condition. Since the picture pairs shown in this condition were visually similar, represented the same concept, and had the same name, the identification of the target object and the retrieval of its name should be faster than in the unrelated condition, where target and distractor represented different concepts and had different names. If any of the processes involved in picture naming require cognitive capacity, the tone discrimination latencies should be shorter in the identical than in the unrelated condition as well.

We found that the picture naming latencies were indeed much shorter in the identical than in the unrelated condition and that this effect carried forward to the tone discrimination latencies. More importantly, the phonologically related distractors also facilitated picture naming, and deviating from Ferreira and Pashler's (2002) results, the effect also carried forward to the tone discrimination latencies.

In order to explore the basis for this discrepancy in the results, we conducted two further experiments using distractor words. The results of the first of these experiments suggested that (a) related distractor words facilitated phoneme selection but slowed down the speakers' self-monitoring processes, and (b) the tone discrimination latencies were sensitive to both of these effects. Experiment 3 tested this hypothesis. As in the preceding experiment, the target pictures were combined with words, but in order to minimize their effects on self-monitoring processes, we presented them for only 50 ms and had them preceded and followed by pattern masks. As will be shown, this yielded a clearer picture of the capacity demands of phoneme selection.

Experiment 1

Method

Participants. Experiment 1 was carried out with 48 undergraduate students of the University of Birmingham (United Kingdom). They were native speakers of British English and had normal or corrected-to-normal vision. They were paid or received course credits for participating in the study.

Materials. One hundred sixteen line drawings of objects with monosyllabic names were selected from a picture gallery available in our lab. Thirty-three of the pictures stemmed from the Snodgrass and Vanderwart (1980) corpus. Seventy-two pictures were used on experimental trials. This experimental picture set consisted of 18 pairs of pictures whose names shared the onset consonant or consonant cluster and the following vowel (e.g., *bed–bell*) and 18 pairs with monosyllabic rhyming names (e.g., *nest–vest*). Each member of an experimental pair was used both as a target and as a distractor. Therefore, there were 72 related target–distractor pairs (see Appendix for a listing of the materials). Also, 72 phonologically and semantically unrelated pairs were

generated by recombining the pictures. In the identity condition, target and distractor were mirror images of each other. Each experimental picture was also combined with one of 36 additional unrelated distractor pictures. These unrelated target–distractor pairs were each shown twice in a practice block. Finally, there were four unrelated picture pairs, which were shown on warm-up trials.

The targets were shown as green line drawings and the distractors as superimposed red line drawings, all on a gray background. The pictures appeared in the center of the screen and were scaled to fit into 9.5-cm by 9.5-cm frames (8.5° by 8.5° when seen from the participant's position). Three tones of 180 Hz (low), 500 Hz (medium), and 1,200 Hz (high) were used. The tone duration was 285 ms.

Design. There were six test blocks of 40 trials each (4 practice trials followed by 36 experimental trials). Each experimental picture was shown as a target in Blocks 1, 3, and 5 or in Blocks 2, 4, and 6. In each block, 12 targets were combined with related, unrelated, and identical distractors. Each target was shown with a different distractor in each of the three blocks in which it appeared. Thus, each participant saw each target with a related, an unrelated, and an identical distractor. The order of the blocks was counterbalanced across participants with a Latin square design. The order of the items within blocks was random, with the provision that distractors of the same type did not appear on more than two consecutive trials. A different sequence of trials was used for each participant.

Three picture–tone SOAs—50, 150, and 900 ms—were used. Each SOA was used 12 times on the experimental trials of each block and once or twice on warm-up trials. Within each block, the same SOA was used for all trials featuring the same distractor type. For instance, in one block the SOA of 50 ms was used when identical distractors were presented, the SOA of 150 ms when unrelated distractors were presented, and the SOA of 900 ms when related distractors were presented. Across the six blocks, the combinations of SOA and distractor type were counterbalanced, such that by the end of the experiment all possible combinations of distractor type and SOA had been used equally often. Each of the three tones was used 12 times on the experimental trials of each block (four times in each distractor condition) and once or twice on warm-up trials. Tone type varied randomly across trials, with the provision that the same tone was not played on more than two consecutive trials.

The experimental blocks were preceded by three practice blocks. In the first practice block, which included 24 trials, the participants practiced only the tone discrimination task. The three tones were played eight times each in a random order. In the second block, the participants saw each warm-up and experimental picture once and heard the tones at the same SOAs as in the experimental blocks. They practiced naming the pictures and responding to the tones. Finally, in the third practice block, each warm-up and experimental picture was shown two more times accompanied by unrelated practice distractors. The participants' task was to name the target pictures and to categorize the tones.

Procedure. Participants were tested individually in a sound-attenuated booth. Before the experiment, they received written instructions and a picture booklet showing the target pictures and the names they should use in the experiment. They were asked to familiarize themselves with the materials and were then tested

orally by the experimenter to ensure that they would use the expected picture names.

The participants were told that they would see pairs of line drawings and that they should name the object shown in green and ignore the object shown in red. In addition, they should promptly categorize the tones as low, medium, or high by pressing the appropriate button on the push-button panel. Both tasks were given equal emphasis. No instructions about the order of the responses were given, but the analyses showed that in all experiments, the participants named the picture first on more than 97% of the valid trials.

At the beginning of each trial, a fixation point was presented in the center of the screen for 1 s. After a blank interval of 500 ms, the target and distractor picture were presented simultaneously for 850 ms. A tone was played 50, 150, or 900 ms after picture onset. The trial ended as soon as the voice key was activated and a button-press response was registered but not earlier than 2 s after picture onset. The maximal trial duration was 3.5 s. After 500 ms the next trial began.

Apparatus. The experiment was controlled by the experimental package NESU 5.1 (Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands). The pictures were presented on a Samtron 95 Plus 19-in. (48-cm) screen. The tones were heard through Beyerdynamic DT 931 headphones. The participants' speech was recorded with a Sony ECM-MS907 microphone. Speech onset latencies were measured with the NESU voice key. A custom-made three-button response panel was used to record the tone discrimination latencies. This setup was used in all experiments of the present study.

Analysis. We recorded the rates of naming errors (incorrect picture names and utterances including hesitations and repairs) and tone discrimination errors as well as the latencies for correct naming and tone discrimination responses. Separate analyses of variance (ANOVAs) were carried out for naming and tone discrimination errors with distractor type (related vs. unrelated) and SOA (50, 150, or 900 ms) as independent variables.

The picture naming and tone discrimination latencies were analyzed in joint ANOVAs that included distractor type, SOA, and task (picture naming vs. tone discrimination) as independent variables. This was done because our main goal was to determine whether the distractors had similar or different effects on the performance in the two tasks. In other words, our main interest was to determine whether task and distractor type interacted with each other. To correct for differences in the mean picture naming versus tone discrimination latencies and variances, we z-transformed the latencies prior to analysis, which yielded standard scores with zero mean and unit standard deviation (Winer, Brown, & Michels, 1991). Because of this transformation, the main effect of task could not be assessed. In separate analyses, participants (F_1) and items (F_2) were used as random variables.

The related target–distractor pairs shared either word-initial or word-final phonemes. Latency analyses including the position of the shared phonemes as an additional variable yielded no main effect of this variable and no interactions with other variables. Therefore we combined the two item sets for the analyses reported here.

Table 1
Results of Experiment 1

Condition	Picture naming response		Tone discrimination response	
	Latency	Error rate	Latency	Error rate
SOA = 50 ms				
Identical	723 (113)	0.9 (0.03)	1,140 (203)	5.4 (0.05)
Related	739 (110)	1.7 (0.03)	1,147 (203)	3.2 (0.05)
Unrelated	761 (124)	1.7 (0.04)	1,171 (212)	4.3 (0.08)
SOA = 150 ms				
Identical	732 (129)	1.7 (0.03)	1,068 (183)	3.8 (0.05)
Related	758 (117)	1.5 (0.03)	1,073 (214)	4.3 (0.05)
Unrelated	778 (119)	1.6 (0.05)	1,087 (190)	3.4 (0.04)
SOA = 900 ms				
Identical	706 (177)	2.3 (0.04)	628 (134)	4.6 (0.05)
Related	745 (211)	2.1 (0.03)	656 (128)	2.0 (0.03)
Unrelated	756 (194)	1.8 (0.06)	654 (143)	3.1 (0.04)

Note. Mean picture naming and tone discrimination latencies (in ms, based on participant means, with standard deviations in parentheses) and error rates (%, means and standard deviations) for each SOA and distractor type. SOA = stimulus onset asynchrony.

Results and Discussion

The results are summarized in Table 1. In the ANOVAs of the rates of picture naming errors, neither the main effects nor their interaction was significant. For the rates of tone discrimination errors, we obtained a significant main effect of distractor type, $F_1(2, 94) = 5.31$, $p < .01$; $F_2(2, 142) = 4.56$, $p < .05$. Participants made more errors in the identical condition (4.6%) than in the related and unrelated conditions (3.2% and 3.6%, respectively). We do not know how this effect arose, but it should be noted that it was not replicated in the following experiments.

Trials on which a picture naming or tone discrimination error had occurred were eliminated from the latency analyses. Further trials were excluded because the picture naming latency was longer than 2,000 ms or shorter than 300 ms (2.5% of the trials) or because the tone discrimination latency was longer than 3,000 ms (1.8% of the trials). Finally, 2.0% of the trials were excluded for technical reasons, mainly because the voice key was triggered by noise in the environment.

For the remaining latencies we obtained a main effect of SOA, $F_1(2, 94) = 120.87$; $F_2(2, 142) = 169.43$, both $ps < .001$, and a significant interaction between response type and SOA, $F_1(2, 94) = 202.61$; $F_2(2, 142) = 1,364.57$, both $ps < .001$. This interaction arose because the tone discrimination latencies decreased across SOAs, from a mean latency of 1,153 ms at the SOA of 50 ms, to 1,076 ms at the SOA of 150 ms, and 646 ms at the SOA of 900 ms, $F_1(2, 94) = 708.76$; $F_2(2, 142) = 617.05$, both $ps < .001$, for the simple main effect of SOA, whereas the picture naming latencies were not significantly affected by the SOA (means: 741, 756, and 736 ms for the SOAs of 50, 150, and 900 ms, respectively); $F_1(2, 94) < 1$; $F_2(2, 142) = 2.53$, $p < .10$. As explained in the introduction, a marked decrease in Task 2 (tone discrimination) latencies with increasing SOA indicates that some of the processes involved in Task 1 (picture naming) required central processing capacity.

As predicted, the participants named the pictures faster when they were accompanied by identical or phonologically related
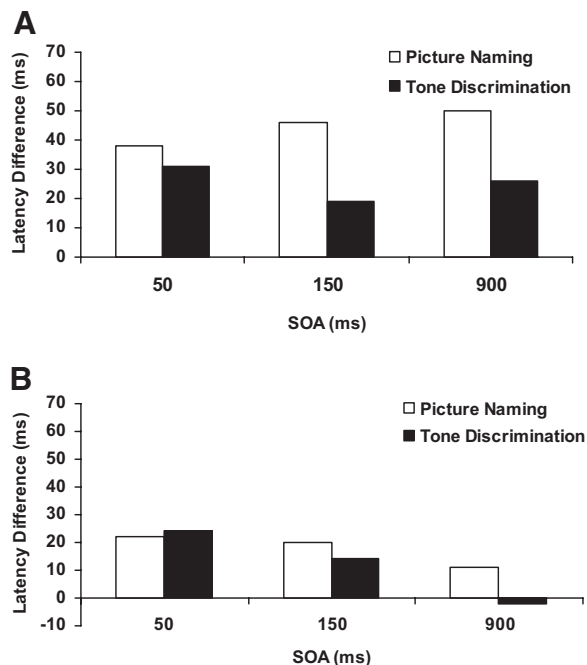
Figure 1. Results of Experiment 1. Effects of identical distractors (A) and phonologically related distractors (B) compared with unrelated distractors on picture naming and tone discrimination latencies.

distractors than when accompanied by unrelated distractors (see Figure 1). Across SOAs, the facilitatory effects of identical and related distractors were 45 ms and 18 ms, respectively. The results obtained for the tone discrimination latencies were similar, though the facilitatory effects were somewhat smaller: Across SOAs, identical distractors sped the reactions by, on average, 26 ms and phonologically related ones by 12 ms. This yielded a significant main effect of distractor type, $F_1(2, 94) = 38.65$; $F_2(2, 142) = 11.59$, both $ps < .001$, and a significant interaction between distractor type and response type, $F_1(2, 94) = 15.88$; $F_2(2, 142) = 10.53$, both $ps < .01$.

The interaction arose primarily because the identical distractors had a stronger effect on the picture naming latencies than on tone discrimination latencies (see Figure 1). An analysis including only the identical and unrelated distractor conditions yielded a significant main effect of distractor type, $F_1(1, 47) = 98.02$; $F_2(1, 71) = 23.27$, both $ps < .001$, and a significant interaction of distractor type and response type, $F_1(1, 47) = 36.36$; $F_2(1, 71) = 16.65$, both $ps < .001$. However, analyses of simple effects showed that the facilitatory effect of the identical distractors was significant for the picture naming latencies, $F_1(1, 47) = 121.83$; $F_2(1, 71) = 35.39$, both $ps < .001$, and for the tone discrimination latencies, $F_1(1, 47) = 17.85$, $p < .001$; $F_2(1, 71) = 5.18$, $p < .05$. This pattern suggests that the identical distractors facilitated several components of the target naming process. Some of these components required processing capacity, and the facilitatory effect of the identical distractors on these components propagated to the tone discrimination latencies. In addition, the identical distractors apparently facilitated processes that were not subject to capacity restrictions, and this effect did not propagate to the tone discrimination latencies. Therefore, a stronger facilitatory effect

was seen for the picture naming latencies than for the tone discrimination latencies.

An analysis including only the phonologically related and unrelated distractor conditions yielded a significant main effect of distractor type, $F_1(1, 47) = 12.87$, $p < .001$; $F_2(1, 71) = 4.45$, $p < .05$, and an interaction of distractor type and response type, which was significant by participants only, $F_1(1, 47) = 5.40$, $p < .05$; $F_2(1, 71) = 2.18$, $p < .15$. As Figure 1 shows, at the two shortest SOAs, facilitatory effects of a similar size were found for the two response types. At the SOA of 900 ms, only the effect on the picture naming latencies was maintained. As explained above, such a pattern can arise because capacity bottleneck effects are more likely to occur at short rather than long SOAs. Analyses including only the related and unrelated distractors and only the two shortest SOAs yielded a main effect of distractor type, $F_1(1, 47) = 9.49$, $p < .01$; $F_2(1, 71) = 6.37$, $p < .05$, but no interaction of distractor type and response type, $F_1(1, 47) = 2.67$; $F_2(1, 71) < 1$.

In sum, Experiment 1 showed that phonologically related distractor pictures facilitated target naming and that at the two shortest SOAs this effect propagated to the tone discrimination latencies. This contrasts with Ferreira and Pashler's (2002) finding that phonologically related distractor words facilitated only target naming but not the responses to the tones. It is difficult to interpret this discrepancy in the results of the two studies because they differed not only in the type of distractors (words vs. pictures) but also in several other ways, including the choice of target pictures and distractors, the timing of the trials, and the presence or absence of semantically related and identical distractors. Nevertheless, the pattern suggests that distractor pictures and words had similar effects on the picture naming latencies but differed in their effects on the tone discrimination latencies. Experiment 2 assessed this hypothesis by replacing the distractor pictures with distractor words (i.e., the names of the pictures shown in Experiment 1). We predicted that with these materials the pattern of results obtained by Ferreira and Pashler would be reproduced.

Experiment 2

Method

Participants. The experiment was carried out with 48 participants.

Materials, design, and procedure. The same materials and design were used as in Experiment 1 except that distractor words were presented instead of pictures. These were the names of the pictures presented in Experiment 1. They appeared in red print (18-point Arial type font) in the center of the screen, superimposed upon the target pictures.

Procedure. The same procedure was used as in Experiment 1, except that the participants were told that they would see picture–word pairs rather than picture pairs. They were asked to name the pictures, ignoring the words, and to categorize the tones as quickly and accurately as possible.

Results

Table 2 summarizes the findings. The analysis of the picture naming errors yielded a significant effect of distractor type, $F_1(2, 94) = 10.03$; $F_2(2, 142) = 9.12$, both $ps < .001$, which arose

Table 2
Results of Experiment 2

Condition	Picture naming response		Tone discrimination response	
	Latency	Error rate	Latency	Error rate
SOA = 50 ms				
Identical	693 (138)	0.2 (0.01)	1,245 (248)	5.5 (0.05)
Related	792 (139)	0.8 (0.02)	1,409 (247)	7.1 (0.06)
Unrelated	817 (141)	0.9 (0.02)	1,343 (245)	7.1 (0.07)
SOA = 150 ms				
Identical	704 (143)	0.1 (0.01)	1,144 (240)	5.4 (0.06)
Related	795 (162)	1.6 (0.03)	1,302 (254)	5.2 (0.07)
Unrelated	836 (149)	0.9 (0.02)	1,292 (249)	4.6 (0.06)
SOA = 900 ms				
Identical	744 (292)	0.8 (0.04)	756 (235)	4.0 (0.05)
Related	778 (250)	1.7 (0.05)	765 (222)	6.3 (0.07)
Unrelated	841 (264)	1.6 (0.07)	763 (255)	3.4 (0.04)

Note. Mean picture naming and tone discrimination latencies (in ms, based on participant means, with standard deviations in parentheses) and error rates (%, means and standard deviations) for each SOA and distractor type. SOA = stimulus onset asynchrony.

because there were fewer errors in the identical condition (0.4%) than in the related and unrelated conditions (1.7% and 1.1%, respectively). The analysis of the tone discrimination errors also yielded a significant effect of distractor type, $F_1(2, 94) = 6.90$; $F_2(2, 142) = 6.75$, both $ps < .01$. The participants made slightly more errors in the phonologically related condition (6.2%) than in the remaining two conditions (5% for both conditions). There was also a significant effect of SOA, $F_1(2, 94) = 3.64$; $F_2(2, 142) = 3.74$, both $ps < .05$. The rate of the tone discrimination errors decreased across SOAs (means: 6.5%, 5.1%, and 4.6% for SOAs of 50, 150, and 900 ms, respectively).

Error trials were excluded from the latency analyses. We excluded a further 1.5% of the trials because the picture naming latency was longer than 2,000 ms or shorter than 300 ms, and 1.5% of the trials because the tone discrimination latency was longer than 3,000 ms. Also, 2.5% of the trials were excluded for technical reasons.

ANOVAs of the remaining latencies yielded a main effect of SOA, $F_1(2, 94) = 86.39$; $F_2(2, 142) = 683.05$, both $ps < .001$, and a significant interaction between response type and SOA, $F_1(2, 94) = 163.53$; $F_2(2, 142) = 2,298.22$, both $ps < .001$. As in Experiment 1, the tone discrimination latencies decreased across SOAs, from a mean of 1,332 ms at the SOA of 50 ms to 1,246 ms at the SOA of 150 ms and 761 ms at the SOA of 900 ms, $F_1(2, 94) = 331.91$, $F_2(2, 142) = 2,488.07$, both $ps < .001$, for the simple effect, whereas the picture naming latencies were stable (means: 767, 778, and 787 ms for SOAs of 50, 150, and 900 ms, respectively; both $Fs < 1$).

There was a significant main effect of distractor type, $F_1(2, 94) = 96.85$; $F_2(2, 142) = 91.64$, both $ps < .001$, and a significant interaction between distractor type and response type, $F_1(2, 94) = 35.59$; $F_2(2, 142) = 45.50$, both $ps < .001$. As Figure 2 shows, the identical distractors strongly facilitated picture naming, and at the two shortest SOAs this effect propagated to the tone discrimination latencies. At the longest SOA, the tone discrimination latencies in the identical and unrelated condition differed by only 7 ms. In

ANOVAs including only the identical and unrelated condition, this pattern yielded a significant main effect of distractor type, $F_1(1, 47) = 148.20$; $F_2(1, 71) = 172.38$, both $ps < .001$, and significant interactions between response type and distractor type, $F_1(1, 47) = 44.40$; $F_2(1, 71) = 95.08$, both $ps < .001$, and between response type, distractor type, and SOA, $F_1(2, 94) = 4.81$; $p < .01$; $F_2(2, 142) = 3.12$, $p < .05$. Analyses of simple effects showed that, as in Experiment 1, the identical distractors facilitated both types of responses: picture naming latencies, $F_1(1, 47) = 139.55$; $F_2(1, 71) = 249.32$, both $ps < .001$, and tone discrimination latencies, $F_1(1, 47) = 56.77$; $F_2(1, 71) = 53.59$, both $ps < .001$.

The phonologically related distractors also facilitated picture naming, by, on average, 43 ms across the three SOAs. However,
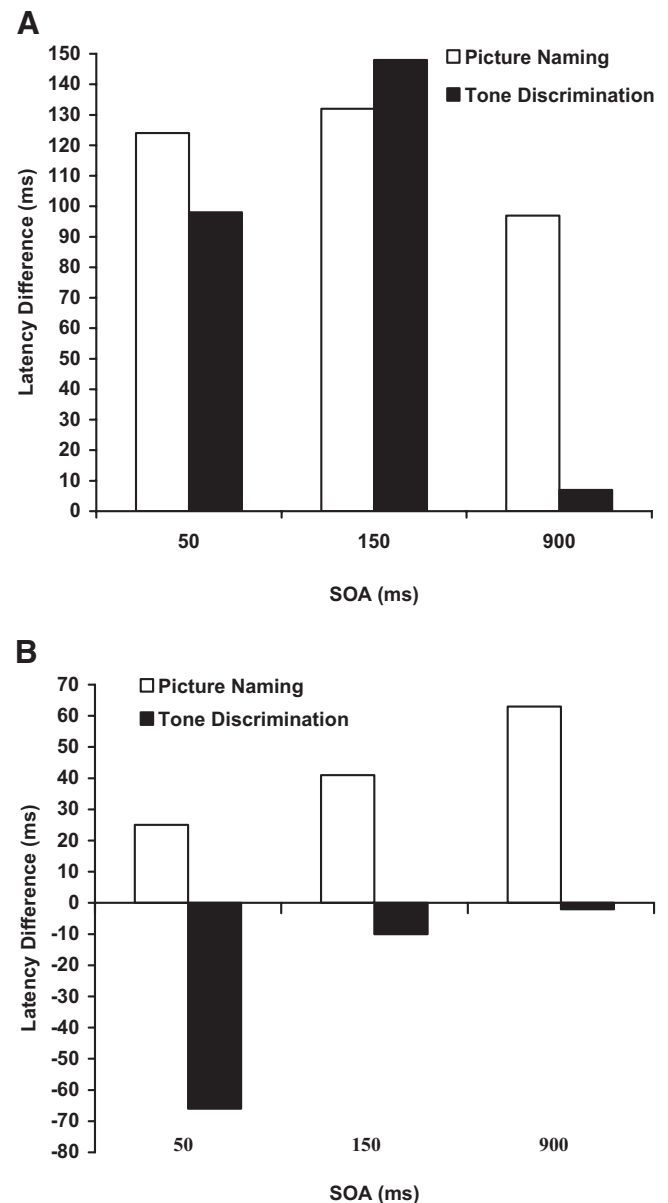

Figure 2. Results of Experiment 2. Effects of identical distractors (A) and phonologically related distractors (B) compared with unrelated distractors on picture naming and tone discrimination latencies.

this effect did not propagate to the tone discrimination latencies: These were longer by, on average, 26 ms when phonologically related as opposed to unrelated distractors were presented. In the ANOVA including only the phonologically related and unrelated conditions, we obtained a significant main effect of distractor type, $F_1(1, 47) = 4.56$; $F_2(1, 71) = 4.40$, both $ps < .05$, and a significant interaction of response type and distractor type, $F_1(1, 47) = 70.14$; $F_2(1, 71) = 63.35$, both $ps < .001$. Analyses of simple effects showed that the facilitatory effect of distractor type on the picture naming latencies was significant, $F_1(1, 47) = 28.81$; $F_2(1, 71) = 28.20$, both $ps < .001$. The inhibitory effect on the tone discrimination latencies was significant in the analysis by participants and approached significance in the analysis by items, $F_1(1, 47) = 5.59$, $p < .05$; $F_2(1, 71) = 3.16$, $p < .10$.

Discussion

Experiments 1 and 2 were identical in all respects except for the modality of the distractors. The results of both experiments showed that phonologically related distractor words and pictures facilitated target naming relative to unrelated distractors, with pictures yielding a more substantial effect (43 ms) than did words (18 ms). However, the facilitatory effect carried forward to only the tone discrimination latencies when distractor pictures were used. The confinement of the effect of related words to the picture naming latencies replicates a key finding of the study by Ferreira and Pashler (2002).

Why might the effects of phonologically related distractor pictures and words on the tone discrimination latencies be so different? One possibility is that the two types of distractors affected different processes. Specifically, related distractor pictures could facilitate capacity demanding processes, which could not be carried out in parallel with the selection of the response to the tone, whereas related distractor words could facilitate response execution processes, which could run in parallel with selection of the response to the tone. Following Ferreira and Pashler (2002), the effect of related distractor words could be allocated at the phoneme level and the effect of related distractor pictures at the level of the phonological word form.

However, within the existing models of lexical access, it is difficult to explain how related distractor pictures and words came to affect different processes. Given that the related distractors shared some phonemes but not the entire phonological word form with the targets, the most likely origin of the facilitatory effect is the phoneme level. This holds for distractor words and pictures (A. S. Meyer & Damian, 2007; A. S. Meyer & Schriefers, 1991; Morsella & Miozzo, 2002; Navarrete & Costa, 2005; Roelofs, Meyer, & Levelt, 1996; Schriefers et al., 1990; Starreveld, 2000). However, effects arising at other processing levels may have contributed to the observed latency differences. If there are bidirectional links between phonemes and superordinate units, as, for instance, proposed by Dell (1986; Dell & Gordon, 2003; see also Vitevitch, 2002), then activation of some of the target phonemes by a related distractor will lead to additional activation of the phonological word form of the target. Therefore, word form selection may be facilitated as well as phoneme selection. This holds equally for both distractor types. In addition, the phonological word form and lemma of the target might become briefly activated during the recognition of a phonologically related distractor word

(Damian & Martin, 1999; Roelofs et al., 1996; Starreveld & La Heij, 1995). However, this should not happen when a distractor picture is processed. Thus, while the written word *bell* might directly activate the orthographic and phonological word form representations of *bed,* there is no reason to assume that a picture of a bell would activate the lemma or phonological word form of *bed.* This is because, according to current models of lexical access, there are no direct facilitatory links between the concepts, lemmas, or word forms of phonologically similar words; such words can activate each other only via links to shared phonemes. In short, it is not obvious why related distractor pictures would have a stronger effect at the phonological word form level than would related distractor words.

An alternative account is that related distractor words and pictures affected lexical access in the same way and that this effect carried forward to the tone discrimination latencies, but that the related distractor words had an additional effect, which was reflected only in the tone discrimination but not in the picture naming latencies. Specifically, we propose that related distractor words slowed down speech monitoring processes, which occurred after speech onset but before the response to the tone.

This hypothesis was derived from the results of an eye movement study by Mortensen et al. (in press), in which participants named object triplets in noun phrases (e.g., *bed, mouse, tree*) and heard distractor words that were phonologically related or unrelated to the name of the first object (e.g., *bell* or *nut*). As expected, the participants looked at the first object for a shorter period of time and named it faster when a related rather than unrelated distractor was presented. However, the second object (*mouse* in the example) was inspected for a longer period of time and named more slowly in the related rather than unrelated condition.

Mortensen et al. (in press) proposed that the related distractors facilitated the selection of the phonemes of the first object name but slowed down the speakers' speech monitoring processes. Speakers continuously monitor their speech plan (the phonological or phonetic representation) and their overt speech for accuracy and adequacy (e.g., Blackmer & Mitton, 1991; Hartsuiker, Corley, & Martensen, 2005; Postma, 2000; Slevc & Ferreira, 2006). The speakers in the experiment by Mortensen et al. typically moved their eyes from the first to the second object about 200 ms before speech onset and were therefore usually looking at the second object when they articulated and monitored the name of the first object. Mortensen et al. suggested that, compared with unrelated distractors, the related distractors might interfere with the recognition of the produced target names or with the matching of the produced with the intended names. This manifested itself in longer gazes at the second object. This proposal is in line with evidence in the word recognition literature demonstrating that the processing of spoken words can be affected by the presence of related primes (e.g., Dufour & Peereman, 2003; McQueen & Sereno, 2005).

In the present experiment, the responses to the picture and the tone may have been affected by the distractors in ways similar to the responses to the two pictures in the study by Mortensen et al. (in press): The participants began to articulate the picture name as soon as they had retrieved the phonological form and at least the first part of the phonetic code (A. S. Meyer, Roelofs, & Levelt, 2003), but since speech monitoring is a capacity demanding process (Hartsuiker, Corley, & Martensen, 2005; Oomen & Postma,

2002), they could respond to the tone only after this process had been completed as well. Therefore, the picture naming latencies showed only the facilitatory effect of the related distractors arising during the selection of the target phonemes. This effect propagated to the tone discrimination latencies but was offset by an opposing effect of the related distractors on the time required for speech monitoring. The net effect of these two influences was that the tone discrimination latencies were slightly longer in the related condition than in the unrelated condition. By contrast, related distractor pictures facilitated phoneme selection but did not hinder the speech monitoring processes more than unrelated distractor pictures did. Therefore, the facilitatory effect arising during phoneme selection carried forward to the tone discrimination latencies. This difference in the effects of related pictures and words might arise because written words are directly linked to the phonological input representations involved in monitoring (e.g., Frost, 1998; Harm & Seidenberg, 2004; Rastle & Brysbaert, 2006), whereas the visual representations of pictures are linked to these representations only via several layers of conceptual and lexical output representations (e.g., Roelofs, 2005; Roelofs et al., 1996).

This view predicts that the facilitatory effect of related distractor words on the picture naming latencies will carry forward to the tone discrimination latencies if the effect of the distractors on speech monitoring is eliminated. Experiment 3 tested this prediction. As in Experiment 2, the target pictures were combined with words (now called primes, following the conventions in the literature), but instead of remaining in view for 850 ms, they were presented for only 50 ms and were preceded and followed by pattern masks. Several studies (Ferrand, Segui, & Grainger, 1996; Schiller, 1998, 1999) have shown that masked phonologically related prime words facilitate phoneme selection relative to unrelated primes. Therefore, we expected the picture naming latencies to be shorter after related than unrelated prime words. Under the chosen timing and masking conditions, few participants should notice the primes, and hardly anyone should be able to identify them. Therefore, the primes should be unlikely to affect the participants' speech monitoring processes and we should be able to observe their effects on phoneme selection, uncontaminated by any effects arising during speech monitoring. If phoneme selection requires processing capacity, the expected facilitatory effect of the related primes should carry forward to the tone discrimination latencies.

Experiment 3

Method

Participants. The experiment was carried out with 48 participants.

Materials and design. The same materials and design were used as in Experiment 2, but masked primes were used instead of clearly visible distractor words. The prime words were preceded and followed by up to four percent signs so that all prime strings were 11 characters long (e.g., %%hamster%%; %%%dog%%%; see also Ferrand et al., 1996; Schiller, 1998). They were shown in lower case in 12-point Courier type font. The forward and backward mask consisted of 13 hash marks, which completely covered the prime string. All visual stimuli were shown in the center of the screen in white on a black background.

Procedure. The participants were instructed and trained in the same way as in the preceding experiments, except that no distractors or primes were mentioned. They were asked to name the pictures and categorize the tones as quickly and accurately as possible.

At the beginning of each trial, a fixation cross was shown for 500 ms. It was followed by a forward mask shown for 500 ms, a prime shown for 50 ms, a backward mask shown for 17 ms, and finally the target picture, which was shown for 850 ms. As in the preceding experiments, one of three tones was presented 50, 150, or 900 ms after picture onset. The trial ended as soon as the voice key was activated and a button-press response was registered but not earlier than 2 s after picture onset. The maximal trial duration was 3.5 s. The intertrial interval was 500 ms.

After participants completed the six experimental blocks, they were asked whether they had noticed the primes. Most participants reported that they had noticed some sort of flickering before the picture appeared. None reported that they could read the words. To assess the visibility of the primes, we conducted a two-alternative, forced-choice recognition test. The primes were presented under the same masking conditions as in the preceding experimental blocks, but instead of a picture, two words were presented slightly to the left and right of the fixation point. One of them was the prime, and the other was a foil (one of the primes used on other trials). The primes appeared equally often on the left and right side of the screen. The participants were asked to press the left or right button on a two-button panel to indicate the position of the prime. Their overall accuracy (52%) was at chance level, $t(40) < 1$.

Results and Discussion

The results of Experiment 3 are summarized in Table 3. The analysis of the rates of picture naming errors yielded no significant effects. The analysis of the rates of tone discrimination errors yielded a significant effect of SOA, $F_1(2, 94) = 6.42$; $F_2(2, 142) = 6.27$, both ps < .01. Participants made more errors at the SOA of

Table 3
Results of Experiment 3

Condition	Picture naming response		Tone discrimination response	
	Latency	Error rate	Latency	Error rate
SOA = 50 ms				
Identical	682 (83)	1.4 (0.02)	1,030 (156)	4.9 (0.05)
Related	718 (80)	1.4 (0.02)	1,091 (173)	5.5 (0.05)
Unrelated	748 (113)	1.5 (0.03)	1,131 (184)	5.8 (0.07)
SOA = 150 ms				
Identical	689 (88)	1.6 (0.03)	949 (156)	6.0 (0.08)
Related	736 (74)	1.7 (0.03)	988 (156)	6.5 (0.06)
Unrelated	772 (114)	1.6 (0.03)	1,060 (159)	6.0 (0.07)
SOA = 900 ms				
Identical	653 (88)	1.8 (0.02)	605 (99)	5.3 (0.06)
Related	706 (88)	2.0 (0.03)	626 (106)	3.5 (0.04)
Unrelated	734 (127)	1.5 (0.03)	647 (159)	3.7 (0.05)

Note. Mean picture naming and tone discrimination latencies (in ms, based on participant means, with standard deviations in parentheses) and error rates (%, means and standard deviations) for each SOA and prime type. SOA = stimulus onset asynchrony.

150 ms (6.2%) than at the SOAs of 50 ms (5.4%) and 900 ms (4.2%).

Trials on which a picture naming or tone discrimination error had occurred were eliminated from the latency analyses. Further trials were excluded because the picture naming latency was longer than 2,000 ms or shorter than 300 ms (0.4% of the trials) or because the tone discrimination latency was longer than 3,000 ms (0.3% of the trials). Also, 0.8% of the trials were excluded for technical reasons.

For the remaining latencies, we obtained a main effect of SOA, $F_1(2, 94) = 410.04$; $F_2(2, 142) = 439.29$, both $ps < .001$, and a significant interaction between response type and SOA, $F_1(2, 94) = 364.74$; $F_2(2, 142) = 532.03$, both $ps < .001$. The tone discrimination latencies decreased across SOAs, from 1,084 ms at the SOA of 50 ms to 999 ms at the SOA of 150 ms and 626 ms at the SOA of 900 ms, $F_1(2, 94) = 707.95$; $F_2(2, 142) = 1,199.76$, both $ps < .001$. The picture naming latencies did not decrease across SOAs but were longest at the SOA of 150 ms (732 ms), intermediate at the SOA of 50 ms (716 ms), and shortest at the SOA of 900 ms (698 ms). The difference between the three SOAs was significant, $F_1(2, 94) = 16.48$; $F_2(2,142) = 15.02$, both $ps < .001$, for the simple effect.

As Figure 3 shows, the pictures were named faster in the identical and the phonologically related than in the unrelated condition. Importantly, this facilitation effect carried forward to the tone discrimination latencies. At the two shortest SOAs, the primes had a stronger effect on the tone discrimination than on the picture naming latencies, whereas the reverse held at the SOA of 900 ms. In the ANOVA, we obtained a significant main effect of prime type, $F_1(2, 94) = 14.35$; $F_2(2, 142) = 26.59$, both $ps < .001$, and significant interactions between prime type and response type, $F_1(2, 94) = 29.23$; $F_2(2, 142) = 12.03$, both $ps < .001$, and between prime type, response type, and SOA, $F_1(4, 188) = 4.08$, $p < .01$; $F_2(4, 284) = 6.23$, $ps < .001$.

ANOVAs including only the identical and unrelated conditions yielded a significant main effect of prime type, $F_1(1, 47) = 16.38$, $p < .001$; $F_2(1, 71) = 57.05$, $p < .01$, and significant interactions of prime type and response type, $F_1(1, 47) = 51.26$; $F_2(2, 142) = 27.55$, both $ps < .001$, and of prime type, response type, and SOA, $F_1(2, 94) = 7.68$; $F_2(2, 142) = 11.32$, $ps < .001$.

ANOVAs including only the phonologically related and unrelated conditions yielded a main effect of prime type, $F_1(1, 47) = 5.27$, $p < .05$; $F_2(1, 71) = 9.28$, $p < .01$. The interaction between prime type and response type approached significance by participants, $F_1(1, 47) = 3.78$, $p < .06$; $F_2(1,71) = 1.63$, reflecting the fact that the effect of phonological relatedness was slightly stronger for the tone discrimination than for the picture naming latencies (45 ms vs. 31 ms, respectively). Analyses of the simple effects showed that the effect was significant for each response type: picture naming latencies, $F_1(1, 47) = 5.70$, $p < .05$; $F_2(1,71) = 8.70$, $p < .01$; tone discrimination latencies, $F_1(1, 47) = 4.50$, $p < .05$; $F_2(1,71) = 7.50$, $p < .01$.

In sum, the phonologically related primes facilitated target naming, and this effect propagated to the tone discrimination latencies. This indicates that the process affected by the primes was a capacity demanding process. The implications of the results will be considered below.
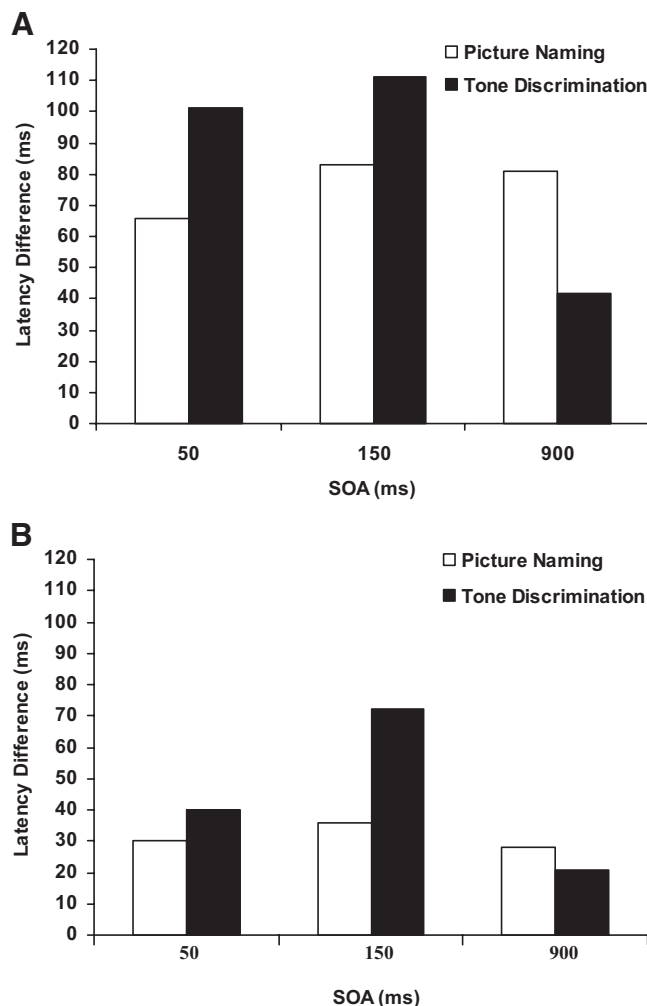

Figure 3. Results of Experiment 3. Effects of identical primes (A) and phonologically related primes (B) compared with unrelated primes on picture naming and tone discrimination latencies.

General Discussion

The goal of the present research was to investigate the capacity demands of one component of lexical access: the selection of the phonemes constituting object names. To this end, three dual-task experiments were carried out in which picture naming (Task 1) was combined with tone discrimination (Task 2). In the critical conditions, the pictures were accompanied by phonologically related or unrelated distractors. Experiment 1 involved distractor pictures, whereas the following experiments involved distractor words, which were either clearly visible (Experiment 2) or masked (Experiment 3). Related distractor pictures facilitated the naming of the targets relative to the unrelated distractors, and this effect carried forward to the tone discrimination latencies. The same pattern was observed for related and unrelated masked distractor words. By contrast, clearly visible distractor words facilitated target naming, but this effect did not carry forward to the tone discrimination latencies.

A plausible account for the finding that distractor pictures and words used in the first two experiments differed in their effect

would be that the facilitatory effects of phonologically related distractor pictures and words had different origins. The related pictures primarily facilitated a capacity demanding process, whereas the related words facilitated a later process that could run in parallel with the selection of the tone response. These two processes could be the selection of the phonological word form and the selection of the individual phonemes, respectively. However, the result of Experiment 3, which demonstrates that masked words had the same effect as pictures, argues against such a proposal. Instead we propose that related words and pictures facilitated the same component of lexical access, but that the related distractor words had the additional effect of slowing down speech monitoring processes that occurred after speech onset but before the response to the tone. The picture naming latencies showed the facilitatory effect of the related distractor words and pictures arising during phoneme selection. This effect carried forward to the tone discrimination latencies, but when distractor words were presented, the effect was counteracted by the opposing effect arising during speech monitoring.

Our hypothesis that the clearly visible distractor words affected the participants' speech monitoring processes was based on results obtained by Mortensen et al. (in press). Additional evidence that is consistent with our proposal comes from a recent picture naming study by Slevc and Ferreira (2006). In this study, participants prepared to name pictures and on about a third of the trials carried out a simple picture naming task. On the remaining trials, a spoken or written word was presented approximately 300 ms after picture onset. This was either the name of the picture or another word that was either unrelated to the name of the picture or related to it in meaning, phonological form, or both meaning and form. The participants were instructed to say the name of the picture only if the word was the picture name but to remain silent otherwise. Therefore, they had to compare the picture name they planned with the written or spoken word. Analyses of the proportions of correctly stopped responses and of the stop latencies showed that this was more difficult when the picture name and the spoken or written word were phonologically related than when they were unrelated. No semantic relatedness effect was found. As Slevc and Ferreira pointed out, these results demonstrate that the speech monitor is particularly sensitive to phonological variables (see also Nooteboom, 2005).

A related study was conducted by Hartsuiker, Pickering, and De Jong (2005). On most trials of the authors' experiments, the participants performed a standard picture naming task, but on some trials, the picture seen at trial onset was replaced after 300 ms by a different picture, which had to be named instead of the original picture. The names of the first and second pictures were semantically or phonologically either related or unrelated. The participants found it more difficult to interrupt their response to the first picture when the two pictures were semantically related than when they were unrelated, but there was no phonological relatedness effect. These results are compatible with our conclusion that the phonologically related distractor pictures did not slow down speech monitoring more than the unrelated distractor pictures did.

One may ask whether the monitoring effects observed in these studies and in our own study were induced by specific features of the experimental situation, such as the presence of written or spoken words, and whether the results yield any useful information about speech monitoring in general. In our view, speech monitor-

ing is always task-specific, as speakers can monitor their speech more or less carefully and can pay attention to different features of their speech (e.g., Hartsuiker, 2006). It is possible that the presence of distractor words induces monitoring or response selection processes that do not occur when words are produced in the absence of distractors (for further discussion see Finkbeiner & Caramazza, 2006a, 2006b; La Heij, Kuipers, & Starreveld, 2006). Nevertheless, the results reflect on properties of the speech monitoring processes that have also been attested to in other contexts, namely that monitoring requires cognitive capacity (Oomen & Postma, 2002) and that it is sensitive to phonological variables (e.g., Nooteboom, 2005; Roelofs, 2004, 2005).

The main goal of the present research was not to study speech monitoring but to study the capacity demands of lexical access. Effects arising during monitoring were discussed because they appeared to conceal other effects. A number of recent studies (Cleland, Gaskell, Quinlan, & Tamminen, 2006; McCann, Remington, & Van Selst, 2000) have demonstrated that the dual-task paradigm can be used to study linguistic processing. Our results confirm this, but they also show that the interpretation of differences in the sensitivity of Task 1 and Task 2 performance to a particular experimental variable is not entirely straightforward. This is because Task 2 performance can be affected by influences taking their effect after the first response has been initiated. A similar point was made by Ferreira and Pashler (2002), who found that the semantic effect in their picture–word interference experiment was more pronounced for the tone discrimination than for the picture naming latencies. They suggested that this pattern could be due to postresponse monitoring processes, which took more time in the semantically related than in the unrelated condition (see also Welford, 1952). The present data highlight the interpretive problem more clearly because they show that one experimental variable, phonological relatedness, can have opposing effects on two components of a linguistic task, and this can conceal the capacity demands of both components.

The most important findings of the present study are that (a) when monitoring effects were minimized, phonologically related words and pictures facilitated target naming relative to unrelated stimuli, and (b) this effect carried forward to the tone discrimination latencies. As explained above, we assume that the primary locus of the phonological effect was the selection of the target phonemes. Ferreira and Pashler (2002) established that lemma selection and word form selection were capacity demanding processes. Our results imply that the same is true for phoneme selection. An interesting topic for future research is whether these processes are equal in capacity demands or whether some processes are more demanding than others.

References

Blackmer, E. R., & Mitton, J. L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition, 39,* 173–194.

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.

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review, 89,* 1–47.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14,* 177–208.

Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: Evidence from the "tip-of-the-tongue" phenomenon. *Cognition, 64,* 309–343.

Clark, H. H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology, 37,* 201–242.

Cleland, A. A., Gaskell, M. G., Quinlan, P. T., & Tamminen, J. (2006). Frequency effects in spoken and visual word recognition: Evidence from dual-task methodologies. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 104–119.

Cutting, J. C., & Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 318–344.

Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychonomic Bulletin & Review, 10,* 111–117.

Damian, M. F., & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 345–361.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93,* 283–321.

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review, 104,* 123–147.

Dell, G. S., & Gordon, J. K. (2003). Neighbors in the lexicon: Friends or foes? In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 9–37). New York: Mouton.

Dufour, S., & Peereman, R. (2003). Inhibitory priming effects in auditory word recognition: When the target's competitors conflict with the prime word. *Cognition, 88,* B33–B44.

Ferrand, L., Segui, J., & Grainger, J. (1996). Masked priming of word and picture naming: The role of syllabic units. *Journal of Memory and Language, 35,* 708–723.

Ferreira, V. S., & Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 1187–1199.

Finkbeiner, M., & Caramazza, A. (2006a). Lexical selection is not a competitive process: A reply to La Heij et al. (2006). *Cortex, 42,* 1032–1035.

Finkbeiner, M., & Caramazza, A. (2006b). Now you see it, now you don't: On turning semantic interference into facilitation in a Stroop-like task. *Cortex, 42,* 790–796.

Foygel, D., & Dell, G. S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language, 43,* 182–216.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123,* 71–99.

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech.* New York: Academic Press.

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition, 82,* B1–B14.

Griffin, Z. M. (2004). Why look? Reasons for eye movements related to language production. In J. M. Henderson & F. Ferreira (Eds.*), The interface of language, vision, and action: What we can learn from free-viewing eye tracking* (pp. 213–247). New York: Psychology Press.

Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language, 38,* 313–338.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111,* 662–720.

Hartsuiker, R. J. (2006). Studies on verbal self-monitoring: The perceptual loop model and beyond. In A. S. Meyer, L. R. Wheeldon, & A. Krott (Eds.), *Automaticity in language processing* (pp. 93–121). Hove, United Kingdom: Psychology Press.

Hartsuiker, R. J., Corley, M., & Martensen, H. (2005). The lexical bias effect is modulated by context, but the standard monitoring account doesn't fly: Belated reply to Baars, Motley, and MacKay (1975). *Journal of Memory and Language, 52,* 58–70.

Hartsuiker, R. J., Pickering, M. J., & De Jong, N. (2005). Semantic and phonological context effects in speech error repair. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 921–932.

Hazeltine, E., Ruthruff, E., & Remington, E. W. (2006). The role of input and output modality pairings in dual-task performance: Evidence for content-dependent central interference. *Cognitive Psychology, 52,* 291–345.

Johnson, C. J., Paivio, A., & Clark, J. M. (1996). Cognitive components of picture naming. *Psychological Bulletin, 120,* 113–139.

Johnston, J. C., McCann, R. S., & Remington, R. W. (1995). Chronometric evidence for two types of attention. *Psychological Science, 6,* 365–369.

Kahneman, D. (1973). *Attention and effort.* New York: Prentice–Hall.

Korvorst, M., Roelofs, A., & Levelt, W. J. M. (2007). Telling time from analog and digital clocks: A multiple-route account. *Experimental Psychology, 54,* 187–191.

Kubose, T., Bock, K., Dell, G. S., Garnsey, S. M., Kramer, A. F., & Mayhugh, J. (2006). The effects of speech production and speech comprehension on simulated driving performance. *Applied Cognitive Psychology, 20,* 43–63.

La Heij, W., Kuipers, J.-R., & Starreveld, A. (2006). In defense of the lexical-competition account of picture-word interference: A comment on Finkbeiner and Caramazza (2006). *Cortex, 42,* 1028–1031.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioural and Brain Sciences, 22,* 1–75.

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word, 15,* 19–44.

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 503–535.

McCann, R. S., & Johnston, J. C. (1992). Locus of the single-channel bottleneck in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 471–484.

McCann, R. S., Remington, R. W., & Van Selst, M. (2000). A dual-task investigation of automaticity in visual word processing. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1352–1370.

McLeod, P. (1977). Parallel processing and the psychological refractory period. *Acta Psychologica, 41,* 381–391.

McQueen, J. M., & Sereno, J. (2005). Cleaving automatic processes from strategic biases in phonological priming. *Memory & Cognition, 33,* 1185–1209.

Meyer, A. S., & Belke, E. (2007). Word form retrieval in language production. In G. Gaskell (Ed.), *Oxford handbook of psycholinguistics* (pp. 471–487). Oxford: Oxford University Press.

Meyer, A. S., Belke, E., Häcker, C., & Mortensen, L. (2007). Use of word length information in utterance planning. *Journal of Memory and Language, 57,* 210–231.

Meyer, A. S., & Damian, M. F. (2007). Activation of distractor names in the picture-picture interference paradigm. *Memory & Cognition, 35,* 494–503.

Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects

in picture naming: The role of a response criterion. *Journal of Memory and Language, 47,* 131–147.

Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 1146–1160.

Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition, 66,* B25–B33.

Meyer, A. S., & van der Meulen, F. F. (2000). Phonological priming effects on speech onset latencies and viewing times in object naming. *Psychonomic Bulletin and Review, 7,* 314–319.

Meyer, D. E., & Kieras, D. E. (1997a). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review, 104,* 3–65.

Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of psychological refractory period phenomena. *Psychological Review, 104,* 749–791.

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.

Mortensen, L., Meyer, A. S., & Humphreys, G. W. (in press). Speech planning during multiple-object naming: Effects of aging. *Quarterly Journal of Experimental Psychology.*

Navarrete, E., & 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.

Navon, D., & Gopher, D. (1980). Task difficulty, resources, and dual-task performance. In R. S. Nickerson (Eds.), *Attention and Performance* (Vol. 8, pp. 297–315). Hillsdale, NJ: Erlbaum.

Navon, D., & Miller, J. (2002). Queuing or sharing? A critical evaluation of the single-bottleneck notion. *Cognitive Psychology, 44,* 193–251.

Nooteboom, S. G. (2005). Listening to oneself: Monitoring speech production. In R. Hartsuiker, Y. Bastiaanse, A. Postma, & F. Wijnen (Eds.), *Phonological encoding and monitoring in normal and pathological speech* (pp. 167–186). Hove, United Kingdom: Psychology Press.

Oomen, C. C. E., & Postma, A. (2002). Limitations in processing resources and speech monitoring. *Language and Cognitive Processes, 17,* 163–184.

Pashler, H. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 358–377.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116,* 220–244.

Pashler, H. (1998). *The psychology of attention.* Cambridge, MA: MIT Press.

Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition, 77,* 97–131.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107,* 460–499.

Rapp, B., & Goldrick, M. (2006). Speaking words: Contributions of cognitive neuropsychological research. *Cognitive Neuropsychology, 23,* 39–73.

Rastle, K., & Brysbaert, M. (2006). Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychology, 53,* 97–145.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42,* 107–142.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition, 64,* 249–284.

Roelofs, A. (2002). Spoken language planning and the initiation of articulation. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 55A,* 465–483.

Roelofs, A. (2004). Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: Comment on Rapp and Goldrick (2000). *Psychological Review, 111,* 561–572.

Roelofs, A. (2005). Spoken word planning, comprehending, and self-monitoring: Evaluation of WEAVER++. In R. J. Hartsuiker, R. Bastiaanse, & F. Wijnen (Eds.), *Phonological encoding and monitoring in normal and pathological speech* (pp. 42–63). Hove, United Kingdom: Psychology Press.

Roelofs, A. (2007). Attention and gaze control in picture naming, word reading, and word categorizing. *Journal of Memory and Language, 57,* 232–251.

Roelofs, A., Meyer, A. S., & Levelt, W. J. M. (1996). Interaction between semantic and orthographic factors in conceptually driven naming: Comment on Starreveld and La Heij (1995). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(1), 246–251.

Schiller, N. O. (1998). The effect of visually masked syllable primes on the naming latencies of words and pictures. *Journal of Memory and Language, 39,* 484–507.

Schiller, N. O. (1999). Masked syllable priming of English nouns. *Brain and Language, 68,* 300–305.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29,* 86–102.

Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychological Science, 12,* 101–108.

Slevc, L. R., & Ferreira, V. S. (2006). Halting in single-word production: A test of the perceptual-loop theory of speech monitoring. *Journal of Memory and Language, 54,* 515–540.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174–215.

Starreveld, P. (2000). On the interpretation of onsets of auditory context effects in word production. *Journal of Memory and Language, 42,* 497–525.

Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 686–698.

Starreveld, P., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 896–918.

Strayer, D. L., & Drews, F. A. (2007). Cell-phone-induced driver distraction. *Current Directions in Psychological Science, 16,* 128–131.

Telford, C. W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology, 14,* 1–36.

Tombu, M., & Jolicoeur, P. (2003). A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 3–18.

Tombu, M., & Jolicoeur, P. (2005). Testing the predictions of the central capacity sharing model. *Journal of Experiment Psychology: Human Perception and Performance, 31,* 790–802.

Vigliocco, G., & Hartsuiker, J. R. (2002). The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin, 128,* 442–472.

Vitevitch, M. S. (2002). The influence of phonological similarity neighbourhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 735–747.

Welford, A. T. (1952). The "psychological refractory period" and the timing of high speed performance: A review and a theory. *British Journal of Psychology, 43,* 2–19.

Welford, A. T. (1959). Evidence of a single-channel decision mechanism limiting performance in a serial reaction task. *Quarterly Journal of Experimental Psychology, 11,* 193–210.

Wheeldon, L. R., Meyer, A. S., & van der Meulen, M. M. (2007). Eye gaze patterns in anticipation errors. In R. van Gompel, M. H. Fisher, W. S. Murray, & L. R. Hill (Eds.), *Eye movements: A window to mind and brain* (pp. 501–515). Amsterdam: Elsevier.

Wilshire, C. E., & Saffran, E. M. (2005). Contrasting effects of phonological priming in aphasic word production. *Cognition, 95,* 31–71.

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design* (3rd ed.). New York: McGraw–Hill.

Appendix

Targets Followed by Related and Unrelated Distractors

Begin-related items

bed–bell–nut, bell–bed–nun, nun–nut–bell, nut–nun–bed

bucket–button–rabbit, button–bucket–radish, rabbit–radish–bucket, radish–rabbit–button

camel–candle–pencil, candle–camel–pepper, pencil–pepper–candle, pepper–pencil–camel

card–car–dog, car–card–doll, dog–doll–card, doll–dog–car

hammer–hamster–chain, hamster–hammer–chair, chain–chair–hammer, chair–chain–hamster

hand–hat–pin, hat–hand–pig, pig–pin–hat, pin–pig–hand

lamp–lamb–plane, lamb–lamp–plate, plane–plate–lamp, plate–plane–lamb

lemon–letter–doctor, letter–lemon–dolphin, dolphin–doctor–letter, doctor–dolphin–lemon

can–cat–harp, cat–can–heart, harp–heart–can, heart–harp–cat

End-related items

bear–pear–cart, pear–bear–dart, cart–dart–bear, dart–cart–pear

bone–throne–clock, throne–bone–sock, clock–sock–bone, sock–clock–throne

book–hook–ring, hook–book–king, king–ring–hook, ring–king–book

train–brain–scale, brain–train–whale, whale–scale–brain, scale–whale–train

cake–snake–spoon, snake–cake–moon, moon–spoon–snake, spoon–moon–cake

bread–head–nail, head–bread–snail, snail–nail–head, nail–snail–bread

house–mouse–clown, mouse–house–crown, clown–crown–house, crown–clown–mouse

plug–jug–vest, jug–plug–nest, nest–vest–jug, vest–nest–plug

box–fox–gun, fox–box–sun, gun–sun–fox, sun–gun–box