Journal of Experimental Psychology:
Human Perception and Performance
2008, Vol. 34, No. 6, 1580–1598

0096-1523/08/$12.00 DOI: 10.1037/a0012476

Attention, Gaze Shifting, and Dual-Task Interference From Phonological Encoding in Spoken Word Planning

Ardi Roelofs
Nijmegen Institute for Cognition and Information and F. C. Donders Centre for Cognitive Neuroimaging

Controversy exists about whether dual-task interference from word planning reflects structural bottleneck or attentional control factors. Here, participants named pictures whose names could or could not be phonologically prepared, and they manually responded to arrows presented away from (Experiment 1), or superimposed onto, the pictures (Experiments 2 and 3); or they responded to tones (Experiment 4). Pictures and arrows/tones were presented at stimulus onset asynchronies of 0, 300, and 1,000 ms. Earlier research showed that vocal responding hampers auditory perception, which predicts earlier shifts of attention to the tones than to the arrows. Word planning yielded dual-task interference. Phonological preparation reduced the latencies of picture naming and gaze shifting. The preparation benefit was propagated into the latencies of the manual responses to the arrows but not to the tones. The malleability of the interference supports the attentional control account. This conclusion was corroborated by computer simulations showing that an extension of WEAVER++ (A. Roelofs, 2003) with assumptions about the attentional control of tasks quantitatively accounts for the latencies of vocal responding, gaze shifting, and manual responding.

Keywords: attention, computational modeling, executive control, eye tracking, speech production

It is often assumed that the extent to which a task can be performed without hindering other simultaneously performed tasks indexes the extent that the task is practiced and carried out by separate, independent, functional processing mechanisms (e.g., Allport, 1980, 1989; D. E. Meyer & Kieras, 1997a, 1997b; Navon & Gopher, 1979; Pashler, 1994; Sperling & Dosher, 1986; Tombu & Jolicoeur, 2003). Spoken word production is a specialized human activity, and it is one of our most highly practiced everyday psychomotor skills. Some researchers believe, therefore, that most aspects of spoken word production need not slow the performance of a concurrent unrelated task (e.g., Levelt, Roelofs, & Meyer, 1999). Seemingly in contrast to this assumption, recent studies have shown that there are detrimental consequences of speaking for other highly practiced activities, such as driving a car (Kubose et al., 2006; Levy, Pashler, & Boer, 2006). It is unclear, though, whether the negative effects from speaking on concurrent driving arise from conceptualizing what to convey or actually planning the words. However, Ferreira and Pashler (2002) obtained evidence that word planning itself, in particular the stages of lemma retrieval and morphological encoding engaged in object naming, may hinder the performance of a concurrently performed, unrelated manual task, whereas only the stage of phonological encoding does not cause dual-task interference.

The results of Ferreira and Pashler (2002) are somewhat surprising in the light of the finding that phonological encoding in object naming affects gaze shifting and the performance of another object-naming task (e.g., A. S. Meyer & Van der Meulen, 2000; A. S. Meyer, Roelofs, & Levelt, 2003). This raises the question of whether dual-task interference from stages of word planning other than conceptualizing is caused by structural (Ferreira & Pashler, 2002) or attentional control factors (D. E. Meyer & Kieras, 1997a, 1997b; Roelofs, 2007). The aim of the experiments reported in the present article was to clarify this issue by examining the relationship among dual-task interference, phonological encoding in vocal responding, gaze shifting, and manual responding.

I begin by briefly describing the study of Ferreira and Pashler (2002) and the eye-tracking studies of A. S. Meyer and colleagues (A. S. Meyer & Van der Meulen, 2000; A. S. Meyer et al., 2003). Next, I argue that the findings of Ferreira and Pashler do not need to imply that a central response-selection mechanism is shared between linguistic and nonlinguistic tasks, as Ferreira and Pashler assumed, but that they can also be explained in terms of attentional control. I then present four dual-task experiments that examined the relationship among dual-task interference, phonological encoding in vocal responding, gaze shifting, and manual responding. The experiments manipulated the requirement of gaze shifting and the attention demands of the secondary manual task. The attention requirements of the second task should matter under the attentional control account but not under the structural bottleneck account. Next, the utility of the attentional control account is demonstrated through computer simulations of the key findings. The account quantitatively fits the data on vocal responding, gaze shifting, and manual responding. Theoretical consequences are discussed in a General Discussion section.

Ardi Roelofs, Nijmegen Institute for Cognition and Information, Radboud University Nijmegen, The Netherlands, and F. C. Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands.

The preparation of the article was supported by a Vici grant from the Netherlands Organization for Scientific Research. I am indebted to Bicoor Bolla-Bong, Marc van Ekeren, and Marianne Severens for their help in running the experiments and to Ben Maassen and the members of the Division of Action, Intention, and Motor Control of the Nijmegen Institute for Cognition and Information for helpful discussion.

Correspondence concerning this article should be sent to Ardi Roelofs, Nijmegen Institute for Cognition and Information, Radboud University Nijmegen, Spinoza Building B.01.08, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands. E-mail: A.Roelofs@nici.ru.nl

Background to the Present Study

Some models of spoken word production divide the word-planning process into conceptualizing, lemma retrieval (i.e., response selection), word-form encoding (i.e., response programming), and articulation, with word-form encoding further divided into morphological encoding, phonological encoding, and phonetic encoding (e.g., Levelt et al., 1999). According to the WEAVER++ model of spoken word production (e.g., Levelt et al., 1999; Roelofs, 1992, 1997, 2003), the naming of a perceived entity (e.g., a baby) involves the retrieval from memory of the corresponding concept (i.e., BABY(X)), the word's lemma (i.e., a representation specifying that the word *baby* is a noun and, for languages such as Dutch, takes nonneuter grammatical gender), one or more morphemes (i.e., <baby>), phonemes (i.e., the speech sounds /b/, /eI/, and /i/), and syllable motor programs (i.e., [beI] and [bi]). In addition to the selection of phonemes, phonological encoding includes the syllabification of the phonemes and the assignment of a stress pattern across the syllables (e.g., /eI/ is made the nucleus of the first syllable, which is assigned lexical stress). Some researchers assume that conceptualizing is based on central processing mechanisms that are shared with other nonlinguistic tasks, whereas lemma retrieval, word-form encoding, and articulation are achieved by dedicated decentralized mechanisms that are functionally distinct from other nonlinguistic processes (e.g., Levelt et al., 1999; Roelofs, 2003).

Seemingly opposing this assumption, Ferreira and Pashler (2002) obtained evidence in seminal experiments that the word-planning stages of lemma retrieval and morphological encoding engaged in object naming hinder the performance of a concurrently performed, unrelated manual task, whereas only the stage of phonological encoding does not cause dual-task interference. Dual-task research has a long history (e.g., Telford, 1931; Welford, 1952). Ferreira and Pashler (2002) conducted two experiments using the psychological refractory period (PRP) paradigm, a widely used dual-task procedure (for reviews, see D. E. Meyer & Kieras, 1997a; Pashler, 1994). On each trial in a PRP experiment, two stimuli are presented, each requiring a quick and accurate response. The time between the two task stimuli is the stimulus onset asynchrony (SOA), which typically ranges from 0 to 1 s. Instructions to the participants typically state that one response (the "Task 1 response") should be given before the other (the "Task 2 response"). Response times are measured to determine the extent that Task 1 delays the performance of Task 2. The delay is called *dual-task interference*.

In one of the experiments of Ferreira and Pashler (2002, Experiment 2), participants had to name the picture of picture–word combinations (Task 1) and manually indicate the pitch of a tone presented for about 300 ms (Task 2). The SOAs between picture–word stimulus and tone were 50, 150, and 900 ms. The written distractor words were semantically related (e.g., pictured baby, distractor *mother*), phonologically related (e.g., distractor *bagel*), or unrelated to the picture names. Compared to the unrelated distractor words, the semantically related words increased picture-naming latencies and the phonologically related words reduced the naming latencies. Earlier research has suggested that the semantic interference arises in lemma retrieval, whereas the phonological facilitation arises in phonological encoding (cf. Levelt et al., 1999).

Ferreira and Pashler (2002) observed that the semantic interference, but not the phonological facilitation, was propagated into the manual response latencies. That is, the manual response latencies were longer in the semantically related than unrelated condition, and shorter in the phonologically related than unrelated condition. Moreover, in another experiment by Ferreira and Pashler (2002, Experiment 1), effects of picture name frequency (probably arising in morphological encoding, cf. Levelt et al., 1999) were transmitted forward into the manual responses. According to Ferreira and Pashler, these results suggest that lemma retrieval and morphological encoding in word planning draw on central nonlinguistic processing mechanisms that are shared with other nonlinguistic tasks, whereas phonological encoding is achieved by dedicated linguistic mechanisms. In particular, Ferreira and Pashler proposed that lemma and morpheme selection in speaking are achieved by a central response-selection mechanism. When this central mechanism is busy with one task (i.e., picture naming), any other task requiring this mechanism (i.e., the tone discrimination task) must passively wait until the mechanism is done, causing dual-task interference. Because phonological encoding does not use this central mechanism, it does not delay any other task, which explains the absence of dual-task interference.

The results obtained by Ferreira and Pashler (2002) using vocal–manual task combinations are interesting in the light of the finding that phonological encoding in picture naming delays the performance of another picture-naming task. For example, A. S. Meyer and Van der Meulen (2000) asked participants to name two spatially separated pictured objects while hearing spoken distractor words that were phonologically related or unrelated to the first picture name. Naming latencies and eye movements were recorded. It was observed that the participants looked longer at the first picture before making a gaze shift to the second picture in the phonologically unrelated than in the related condition. The onset of naming the second picture was likewise delayed. Similarly, A. S. Meyer et al. (2003) observed that participants looked longer at first pictured objects with disyllabic than with monosyllabic names, and fixating and naming the second picture was correspondingly postponed. Similar effects of phonological length on gaze shifts were obtained by Korvorst, Roelofs, and Levelt (2006) and Levelt and Meyer (2000). The effects of phonological priming and phonological length suggest that shifts of gaze from one object to another are initiated only after the phonological form of the object name has been encoded sufficiently. This suggests that phonological encoding in naming cannot happen simultaneously with another naming task or that speakers prefer not to overlap phonological encoding for two naming responses.

According to Ferreira and Pashler (2002), "Such effects only show, however, that eye movements performed in the service of a linguistic task are affected by linguistic variables and do not show that an independent, unrelated nonlinguistic task must be directly affected by processing effects within the linguistic processing system" (p. 1188). Ferreira and Pashler's own findings, described earlier, suggest that phonological encoding does not delay the performance of such an unrelated task. This raises the question why gaze shifts in double-picture naming are triggered by phonological encoding rather than by an earlier stage, such as conceptualizing, lemma retrieval, or morphological encoding. According to Ferreira and Pashler, lemma retrieval and morphological encoding draw on a central response-selection mechanism. This makes the question about the moment of gaze shifting important, because some researchers have argued that gaze shifts are generally initiated before response selection (D. E. Meyer & Kieras, 1997a, 1997b; Sanders, 1998; Sanders & Lamers, 2002; Van Duren & Sanders, 1995). This does not correspond to what A. S. Meyer et al. (2003) and A. S. Meyer and Van der Meulen (2000) observed.

Whereas A. S. Meyer and colleagues found that gaze shifts occur after response selection, Sanders and colleagues (Sanders, 1998; Sanders & Lamers, 2002; Van Duren & Sanders, 1995) provided empirical evidence that gaze shifts may occur before response selection. In one of the experiments conducted by Van Duren and Sanders (1995), participants had to press one of four keys in response to the digits 1, 2, 3, or 4 presented on the left side of a computer screen depending on a go/no-go symbol presented on the right side of the screen. Perceiving the go/no-go signal required a shift of gaze away from the digits. The participants were told that the order in which the digits were assigned to response keys was from left to right (a compatible stimulus-response mapping) or from right to left (an incompatible stimulus-response mapping). Evidence suggests that stimulus-response compatibility affects response selection (e.g., Sanders, 1998, for a review of the literature). Van Duren and Sanders (1995) observed that the onset latency of the saccade from the digit to the go/no-go symbol was unaffected by stimulus-response compatibility. This finding suggests that the gaze shift was initiated before response selection.

In line with Sanders and colleagues, D. E. Meyer and Kieras (1997a, 1997b) hypothesized that gaze shifts are generally initiated before response selection. Such early gaze shifting was assumed for the strategic response-deferment (SRD) model of dual-task performance implemented in the executive-process interactive control (EPIC) architecture (D. E. Meyer & Kieras, 1997a). In simulations of vocal–manual dual-task performance using EPIC-SRD, the eyes were instructed to move to the spatial location of the visual stimulus for a second manual task when the perceptual processing of a visual stimulus for a first vocal task had progressed far enough. This assumption of early gaze shifting was speculative, because evidence on the actual eye movements was not available from the simulated experiments. Early gaze shifting was not observed by A. S. Meyer et al. (2003) and A. S. Meyer and Van der Meulen (2000). Why, then, do gaze shifts occur so late during the word-planning process, thereby apparently delaying the second naming response more than when gaze shifts had occurred earlier? Below, I argue that the differing gaze patterns observed by A. S. Meyer et al. (2003), A. S. Meyer and Van der Meulen (2000), and Van Duren and Sanders (1995) reflect different executive strategies used in different circumstances.

To summarize, A. S. Meyer et al. (2003) found propagation of a phonological manipulation of picture naming in secondary vocal responses. This finding is surprising for two reasons. First, Ferreira and Pashler (2002) did not observe such phonological propagation with secondary manual responses. Second, the work of Sanders and colleagues (Sanders, 1998; Sanders & Lamers, 2002; Van Duren & Sanders, 1995) and D. E. Meyer and Kieras (1997a, 1997b) suggests that gaze shifts occur before response selection (and hence before phonological encoding). These discrepancies may indicate that individuals shift attention and gaze to one task or another differentially depending on the specific demands raised by the task situation, as I argue next.

An Attentional Control Account

Elsewhere (Roelofs, 2007), I proposed an account of dual-task interference from word planning in terms of the attentional control processes that coordinate task-level processes (e.g., Allport, 1980, 1989; D. E. Meyer & Kieras, 1997a, 1997b; Monsell, 1996). Attentional control concerns "the computational mechanisms by which attentional engagement is established, coordinated, maintained, interrupted, and redirected, both in spatial and nonspatial terms, in the preparation and control of action" (Allport, 1989, pp. 662–663). The account assumes that participants strategically set a criterion for when the shift between two tasks should occur (cf. D. E. Meyer & Kieras, 1997a, 1997b), whereby gaze shifts reflect task shifts. The position of the task-shift criterion within the Task 1 process serves to maintain acceptable levels of speed and accuracy (e.g., Sperling & Dosher, 1986), to minimize resource consumption and to avoid crosstalk between tasks (e.g., Allport, 1980, 1989; Levelt & Meyer, 2000; A. S. Meyer & Van der Meulen, 2000; Zelinsky & Murphy, 2000), and to satisfy instructions about task priorities (e.g., D. E. Meyer & Kieras, 1997a, 1997b). It is possible that the difference in experimental results between Ferreira and Pashler (2002) and Sanders and colleagues (Sanders, 1998; Sanders & Lamers, 2002; Van Duren & Sanders, 1995), on the one hand, and A. S. Meyer et al. (2003) and A. S. Meyer and Van der Meulen (2000), on the other, is related to two important concerns for criterion setting in the attentional control of action, namely the minimization of resource consumption and the avoidance of interference.

As concerns the minimization of resource consumption, a major difference between studies is whether or not Tasks 1 and 2 used the same effector system, leading to differences in the need for response buffering. In the experiments of Van Duren and Sanders (1995), participants had to give only a single response based on both the left and right stimuli (respectively, the digit and go/no-go signal). In the experiments of Ferreira and Pashler (2002), the Task 1 response was vocal and the Task 2 response was manual. In contrast, in the experiments of A. S. Meyer et al. (2003) and A. S. Meyer and Van der Meulen (2000), participants had to vocally name two visual stimuli. Articulation is a slow process. Pronouncing a word can easily take half a second or more. Consequently, at zero SOA, Task 2 planning may be completed well before articulation of the Task 1 response has been finished. Therefore, the vocal response for Task 2 needs to be buffered for a relatively long time. By adopting a serial strategy (i.e., starting perception of the Task 2 stimulus when Task 1 planning is done), the use of buffering resources can be limited (Levelt & Meyer, 2000).

As concerns the avoidance of interference, in case of two vocal responses, planning the response for Task 2 may interfere with planning the response for Task 1. Shifting gaze away from an object while planning its name may lead to interference from seeing other objects whose names may inadvertently be activated (A. S. Meyer & Van der Meulen, 2000; Zelinsky & Murphy, 2000). The interference hypothesis is supported by evidence suggesting that context pictures activate their names (e.g., Morsella & Miozzo, 2002; Navarrete & Costa, 2005). Moreover, it is supported by so-called environmental intrusion errors, whereby irrelevant information about what speakers are looking at is inadvertently included in an utterance (Harley, 1984). With vocal Task 1 responses and manual Task 2 responses (Ferreira & Pashler, 2002), Task 2 response planning should not interfere much with phonological encoding for the Task 1 responses, whereas there may be interference when both responses are vocal. Thus, by adopting a serial strategy, the planning of the first vocal response is protected against interference from planning the second vocal response.

Aims of the Present Study

The purpose of the experiments reported in the present article was twofold. First, the experiments aimed to clarify the role of gaze shifts in a PRP situation with vocal and manual tasks and to test the generality of the claim that gaze shifts happen before response selection (D. E. Meyer & Kieras, 1997a, 1997b; Van Duren & Sanders, 1995). Both dual-task investigations and research on eye movements have a long history in experimental psychology (e.g., Woodworth, 1938; Woodworth & Schlosberg, 1954). However, although "People make eye movements almost continuously . . . they have been almost completely neglected in dual-task-research" (Pashler, Carrier, & Hoffman, 1993, p. 54). Furthermore, "No PRP study has yet examined eye movements carefully during multiple-task performance" (D. E. Meyer and Kieras (1997a, p. 52). To my knowledge, the situation has not been improved since these observations were made. The research reported in the present article compared PRP performance using visual displays that require gaze shifts between the Task 1 and Task 2 stimuli (Experiment 1) with displays containing exactly the same materials but requiring no gaze shifts (Experiment 2).

Second, the reported Experiments 1–4 aimed to further examine dual-task interference from phonological encoding in picture naming. Roelofs (2007) reported dual-task interference from picture naming on Task 2 manual responses. However, this study used zero SOA only, and it did not include the crucial phonological manipulation of picture naming. The research reported in the present article explicitly manipulated phonological encoding in picture naming. It investigated whether the absence of dual-task interference from phonological encoding observed by Ferreira and Pashler (2002) is a general finding or whether by using a slightly different phonological manipulation and manual task, dual-task interference from phonological encoding can be obtained. If phonological encoding draws on decentralized specialized linguistic processing, it should do so in all word production tasks requiring phonological encoding. If phonological encoding yields dual-task interference in some task situations but not in others, the interference would be more readily explained in terms of attentional control (e.g., Allport, 1980, 1989; D. E. Meyer & Kieras, 1997a, 1997b) than in terms of a central response-selection bottleneck in verbal action (Ferreira & Pashler, 2002; Pashler, 1994).

Plan of the Present Study

In the present experiments, Task 1 required a vocal response and Task 2 required a manual response, as in the experiments of Ferreira and Pashler (2002). A Task 2 with manual responses was chosen to avoid the need for vocal response buffering and to minimize verbal interference between the planning of Task 1 and Task 2 responses. In Experiment 1, participants were presented with pictures displayed on the left side of a computer screen and left- or right-pointing arrows displayed on the right side of the screen. The picture and the arrow were presented simultaneously on the screen (SOA = 0 ms) or the arrow was presented 300 or 1,000 ms after picture onset. The participants' tasks were to name the picture (Task 1) and to indicate the direction in which the arrow was pointing by pressing a left or right button (Task 2). Eye movements were recorded to determine the onset of the shift of gaze between the picture and the arrow. In Experiments 2 and 3, the arrow was superimposed onto the pictures. Eye movements were not recorded, because gaze shifts from one side of the computer screen to the other were not required. In Experiment 2, the arrow remained on the screen until manual response onset, whereas the arrow was presented for 300 ms in Experiment 3. In Experiment 4, the arrow discrimination task was replaced by a tone discrimination task. Instead of presenting an arrow, a tone with a low or high pitch was presented over headphones. The duration of the tone was 300 ms (cf. Ferreira & Pashler, 2002). Participants pressed a left button in response to the low tone and a right button in response to the high tone. The primary task and the SOAs were the same as those in Experiments 1–3. Figure 1 illustrates the experimental displays.

Earlier behavioral research has shown that tone discrimination (high/low) may be hampered when another auditory stimulus is presented shortly after tone offset (e.g., Massaro, 1970, 1972). Moreover, brain imaging studies have demonstrated that during self-produced speech, the auditory cortex suppresses its response to acoustic signals. For example, brain responses to tones are weaker during speech production compared with tones alone (Houde, Nagarajan, Sekihara, & Merzenich, 2002). Single-cell recordings in monkeys have revealed that activity in the auditory cortex is inhibited by vocalization (Müller-Preuss, Newman, & Jürgens, 1980; Müller-Preuss & Ploog, 1981). The vocalization-induced suppression already begins several hundred milliseconds

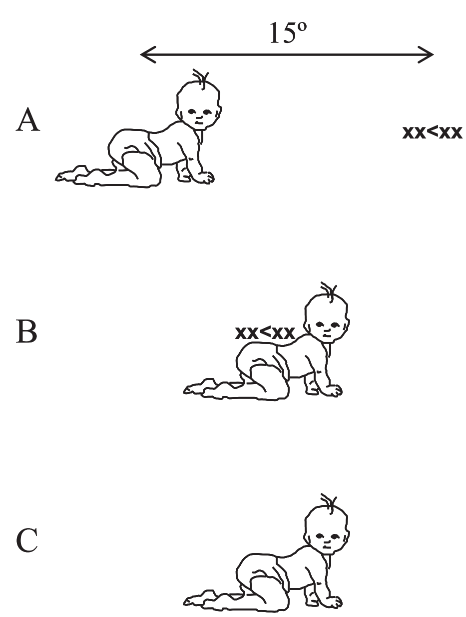

Figure 1. Illustration of the visual stimulus displays used in Experiment 1 (A) and Experiments 2 and 3 (B). In Experiment 4 (C), tones instead of arrows were presented.

prior to the onset of vocalization (Eliades & Wang, 2003).[1] Based on these findings, one expects that in a PRP experiment with vocal Task 1 responses and Task 2 tone stimuli (i.e., Ferreira & Pashler, 2002, and the present Experiment 4), planning and articulating the speech sounds of the Task 1 responses may hamper the auditory recognition of the Task 2 tones. Consequently, the Task 2 process may need to be protected against interference from the Task 1 process (cf. Allport, 1989; Roelofs, 2007). Under a strategic account, attentional engagement may therefore shift earlier with Task 2 tone discrimination (i.e., Ferreira & Pashler, 2002, and the present Experiment 4) than with Task 2 arrow discrimination (i.e., the present Experiments 1–3). As a result, phonological planning may cause dual-task interference on manual responding to the arrows (Experiments 1–3) but not on manual responding to the tones (Experiment 4), even when the arrows are presented as long as the tones (i.e., 300 ms, Experiment 3). In contrast, under a central response-selection bottleneck account (Ferreira & Pashler, 2002), the Task 2 stimulus modality (i.e., arrow vs. tone) should be irrelevant.

If phonological encoding in picture naming delays the performance of the manual task, manipulations of phonological encoding should be propagated into the manual response latencies. The duration of phonological encoding was manipulated in the present experiments by using a version of the form-preparation paradigm developed by A. S. Meyer (1990). The preparation paradigm has been described in depth in various other places, and I refer to these publications for an extensive discussion and motivation of it (see especially A. S. Meyer, 1990, and Roelofs, 2004). In the preparation paradigm, vocal responses (e.g., to pictures) are produced in two types of sets, called homogeneous and heterogeneous sets. In a homogeneous set, the response words share part of their form, for example the first syllable, as in “baby,” “bagel,” and “basin.” In the heterogeneous sets, the response words are unrelated in word form. Regrouping the responses from the homogeneous sets creates the heterogeneous sets.

Typically, response latencies are shorter in homogeneous than in heterogeneous sets, henceforth called the *preparation benefit* (cf. D. E. Meyer & Kieras, 1997a). The magnitude of the preparation benefit increases with the number of shared word-initial phonemes. Moreover, the magnitude of the preparation benefit depends on morpheme structure and other abstract linguistic variables. For example, the preparation benefit is larger when a shared syllable makes up a morpheme in the response words (e.g., the syllable *in* of *input*) than when the same syllable does not make up a morpheme (e.g., the syllable *in* of *insect*) and the benefit is larger for low-frequency morphemes compared with high-frequency morphemes. These findings suggest that preparation benefits are due to word-form encoding processes.

Several of the key findings in the literature on the preparation benefits obtained with A. S. Meyer’s (1990) paradigm have been simulated by WEAVER++ (e.g., Roelofs, 1997, 2002). According to the model, shorter latencies in homogeneous sets than in heterogeneous sets are obtained when participants prepare and buffer a partial word-form representation for the response words before the identity of the actual response on a trial is known. The encoding process is then temporarily suspended, and it is resumed when information about the remainder of the word form becomes available after response selection. For example, if the set of response words consists of “baby,” “bagel,” “basin,” the phonological encoder can construct the first phonological syllable before response selection. In the heterogeneous condition (“baby,” “melon,” etc.), nothing can be prepared. Therefore, response latencies will be shorter in homogeneous than in heterogeneous sets.

To summarize, the primary aims of Experiments 1–4 were (a) to clarify the role of gaze shifts in the PRP paradigm with vocal and manual tasks and to test whether the claim that gaze shifts occur before response selection is universally true, and (b) to investigate whether the absence of dual-task interference from phonological encoding is a general finding or whether by using a different phonological manipulation (homogeneous vs. heterogeneous response sets) and manual task (arrow vs. tone discrimination), dual-task interference from phonological encoding in speaking can be obtained.

Experiment 1

In the first experiment, the participants’ tasks were to name the pictured item and to manually indicate the direction in which the arrow was pointing. The arrows were presented 15° to the right of the pictures, demanding a gaze shift from picture to arrow. The latencies of the vocal responses, gaze shifts, and manual responses were recorded. To ensure that the arrows were fixated and to minimize the chance that participants could identify the direction of the arrows by their peripheral vision, the arrows were presented in small font and they were flanked by two Xs on each side, yielding XX < XX and XX > XX as stimuli.

The picture and the arrow were presented simultaneously on the screen (SOA = 0 ms), or the arrow was presented 300 or 1,000 ms after picture onset. Using the materials of the present study, Roelofs (2004) obtained mean picture-naming latencies of about 525 ms. Thus, on average, picture naming is likely to be completed before the presentation of the arrow at SOA = 1,000 ms but not at the SOA of 300 ms. Thus, if phonological encoding yields dual-task interference, this is expected to happen at the short SOAs (i.e., 0 and 300 ms) but not at the long one (i.e., 1,000 ms). That is, the preparation benefit should be propagated into the manual response latencies at the short SOAs but not at the long one.

Method

Participants. Each experiment was carried out with a different group of 18 paid participants, who were students at Radboud University Nijmegen. All participants were young adults and native speakers of Dutch. None of the participants took part in more than one experiment.

Materials and Design. The stimuli consisted of the nine pictures used in the begin-overlap conditions of Experiments 2 and 3 of Roelofs (2004). All pictures had disyllabic names, which are listed in Appendix A. The pictures were line drawings of simple objects, which were selected from the picture gallery of the Max Planck Institute for Psycholinguistics in Nijmegen. The pictures were digitized and scaled to fit into a virtual frame of 10 cm × 10 cm. On average, the pictures subtended visual angles of 8.7° horizontally and 8.7° vertically at a viewing distance of 66 cm. The

[1] Roelofs, Özdemir, and Levelt (2007) observed that speech planning may affect concurrent speech perception. Moreover, although planning may hamper auditory recognition, a sharing of phonemes between planned speech and heard speech helped recognize the phonemes in the auditory signal.

arrow stimuli XX < XX and XX > XX were presented in 28-point uppercase Arial font, subtending 0.9° vertically and 3.5° horizontally. The horizontal distance between the middle of the picture stimuli and the arrow stimuli was 15°.

The pictures were grouped into three response sets of three stimuli each. Each set was tested in a separate block of trials. The grouping was such that, in three sets (the homogeneous sets), the response words shared their first syllable; and, in the remaining sets (the heterogeneous sets), they were unrelated in form. The shared first syllables were /be/, /wa/, and /le/. Following A. S. Meyer (1990), the independent variable of homogeneous versus heterogeneous sets is called *context*. Each participant was tested once on each set. Each of the pictures in a set was tested four times within a block of trials. The order of testing the pictures was random, except that immediate repetitions of stimuli were excluded. A different order was used for each block and each participant. The order of the sets was counterbalanced across participants. Half the participants were first tested on the homogeneous sets and then on the heterogeneous ones; and, for the other half of the participants, the order of homogeneous and heterogeneous sets was reversed. As indicated, the SOAs between picture and arrow were 0, 300, and 1,000 ms. Trials were blocked by SOA. The order of the SOAs was counterbalanced across participants.

Apparatus. Materials were presented on a 39-cm ViewSonic 17PS screen. Eye movements were measured using an SMI EyeLink-HiSpeed 2D headband-mounted eye-tracking system (SensoMotoric Instruments GmbH, Teltow, Germany). The eye tracker was controlled by a Pentium 90 MHz computer. The experiment was run under the Nijmegen Experiment Setup with a Nijmegen Experiment Setup button box on a Pentium 400 MHz computer. The participants' vocal responses were registered by a Sennheiser ME400 microphone. Vocal response latencies were measured using an electronic voice key. Manual responses were registered using a panel with left and right push buttons. Vocal and manual response latencies were measured from picture and arrow onset, respectively.

Procedure. The participants were tested individually. They were seated in front of the computer monitor, the push-button panel, and the microphone. The distance between participant and screen was approximately 66 cm. Participants were given written instructions telling them how their eyes would be monitored and what the task was. The experimenter also orally described the eye-tracking equipment and restated the instructions. The participants were told that they had to name the picture presented on the left side of a computer screen and manually respond by pressing a left or right button in response to the arrows, XX < XX or XX > XX, presented on the right side of the screen. The participants were asked to respond as quickly as possible without making mistakes. To familiarize them with the pictures and their names, the participants received a booklet showing them all pictures used in the experiment together with the expected names. The instructions to the participants stressed that the vocal responses should have earlier onsets than the manual responses (cf. D. E. Meyer & Kieras, 1997a, 1997b). A manual response completed a trial by blanking the screen. This procedure reinforced the instruction that vocal responses should start before the manual responses.

When a participant had read the instructions and had studied the picture booklet, the headband of the eye-tracking system was placed on the participant's head and the system was calibrated and validated. For pupil-to-gaze calibration, a grid of 3×3 positions had been defined. During a calibration trial, a fixation target appeared once, in random order, in each of these positions for 1 s. Participants were asked to fixate upon each target until the next target appeared. After the calibration trial, the estimated positions of the participant's fixations and the distances from the fixation targets were displayed to the experimenter. Calibration was considered adequate if there was at least one fixation within 1.5° of each fixation target. When calibration was inadequate, the procedure was repeated, sometimes after adjusting the eye cameras. Successful calibration was followed by a pupil-to-gaze validation trial. For the participants, this trial did not differ from the calibration trial, but the data collected during the validation trial were used to estimate the participants' gaze positions, and the error (i.e., the distance between the estimated gaze position and the target position) was measured. Validation was considered completed if the average error was below 1.0° and the worst error below 1.5°. Depending on the result of the validation trial, the calibration and validation trials were repeated or testing began.

After successful calibration and validation, the 216 experimental trials were presented. Before each block of trials testing one set, the pictures of the set were presented on the screen together with their written names. The prior presentation of the pictures and their names provided the participants with foreknowledge about the vocal responses required in the subsequent block of trials. After the participants indicated that they were ready, the first trial of a block began. The structure of a trial was as follows. A trial started by the presentation of the left (picture) and right (arrow) stimuli with the appropriate SOA. The stimuli remained on the screen until the participant pushed the button in response to the arrow. The picture was presented in white color and the arrow was presented in yellow. The background of the screen was black. Before the start of the next trial, there was a blank interval of 1.5 s. The position of the left and right eyes was determined every 4 ms. Drift correction occurred automatically after every 8 trials.

Analyses. To analyze the speakers' gaze shifts, their eye fixations were classified as falling within or on the outer contours of the picture or the arrow. Although viewing was binocular and the positions of both eyes were tracked, only the position of the right eye was analyzed. The latency of a gaze shift was defined as the time interval between the beginning of the first fixation on the picture stimulus and the end of the last fixation before the first saccade was initiated to the arrow.

A naming response was considered to be invalid when it included a speech error, when a wrong word was produced, or when the voice key was triggered incorrectly. A manual response was considered to be invalid when the wrong button was pressed. Technical errors of the eye tracker were virtually nonexistent (< 0.1%). Trials with either a vocal, ocular, or manual error were removed from the analyses of the naming latencies, gaze durations, and manual response latencies. The vocal response latencies, gaze shift latencies, manual response latencies, and errors were submitted to repeated-measure analyses of variance (ANOVAs) with SOA and context as within-participants factors. The analyses were performed both by participants, denoted by F_1, and by items, denoted by F_2 (Clark, 1973). The conventional significance level of $\alpha = .05$ was adopted. When significance is reached in both the analyses by participants and by items, the effect is conventionally significant. In addition, *minF′* (Raaijmakers, Schrijnemakers, & Gremmen, 1999) was computed when both the F_1 and the F_2 reached significance. For post hoc comparisons, a Bonferroni corrected α of .017 was used, both in the comparisons by participants (t_1) and by items (t_2).

Finding main effects of context (i.e., phonological preparation) on the manual responses in Experiments 1–3, but not in Experiment 4, would demonstrate the malleability of the propagation and support the attentional control account. In addition, the statistical analyses tested the specific prediction of propagation effects at SOA = 0 ms and SOA = 300 ms, but not at SOA = 1,000 ms. This explicit prediction allows for direct tests of the propagation effects at each SOA, making testing for an overall interaction of context and SOA superfluous. For the context effect in picture naming, the situation is different, because the effect is predicted not to be dependent on SOA. Testing for an interaction of context and SOA is the most effective way of assessing this.

Results and Discussion

The upper panel of Figure 2 shows for each context and SOA condition the mean latencies for the vocal responses (Task 1), gaze shifts, and manual responses (Task 2). The lower panel of Figure 2 shows the mean error percentages for the vocal and manual responses for each condition.

Task 1 vocal responses. The statistical analysis of the vocal response latencies yielded main effects of context (revealing a preparation benefit of 25 ms, on average), $F_1(1, 17) = 21.66$, $p = .001$, $F_2(1, 8) = 43.91$, $p = .001$, $minF'(1, 25) = 14.50$, $p < .001$, and SOA, $F_1(2, 34) = 12.22$, $p = .001$, $F_2(2, 16) = 49.92$, $p = .001$, $minF'(2, 47) = 9.82$, $p < .001$. The effect of context did not depend on the SOA, $F_1(2, 34) = 1.38$, $p = .27$, $F_2(2, 16) = 3.82$, $p = .04$. Thus, a preparation benefit was obtained, which was constant across SOAs. Post hoc comparisons revealed that the vocal response latencies were longer at zero SOA than at SOA = 1,000 ms, $t_1(17) = 3.46$, $p = .002$, $t_2(8) = 6.48$, $p = .001$, and longer at SOA = 300 ms than at SOA = 1,000 ms, $t_1(17) = 5.04$, $p = .001$, $t_2(8) = 10.39$, $p = .001$, whereas the latencies did not differ between zero SOA and SOA = 300 ms, $t_1(17) = 1.46$, $p = .16$, $t_2(8) = 3.15$, $p = .01$. There were no effects on the errors (all $ps > .21$).

Gaze shifts. The statistical analysis of the gaze durations yielded main effects of context (39 ms, on average), $F_1(1, 17) = 8.97$, $p = .008$, $F_2(1, 8) = 79.36$, $p = .001$, $minF'(1, 21) = 8.06$, $p < .01$, and SOA, $F_1(2, 34) = 20.20$, $p = .001$, $F_2(2, 16) = 583.52$, $p = .001$, $minF'(2, 36) = 19.52$, $p < .001$. The effect of context did not depend on the SOA, $F_1(2, 34) = 0.06$, $p = .95$, $F_2(2, 16) = 0.28$, $p = .76$. Thus, the preparation benefit was reflected in the gaze shifts at all SOAs. Post hoc comparisons revealed that the gaze shift latencies were longer at SOA = 300 ms than at SOA = 0 ms, $t_1(17) = 2.22$, $p = .02$, $t_2(8) = 12.46$, $p = .001$, longer at SOA = 1,000 ms than at SOA = 300 ms, $t_1(17) = 4.26$, $p = .001$, $t_2(8) = 19.69$, $p = .001$, and longer at SOA = 1,000 ms than at SOA = 0 ms, $t_1(17) = 5.28$, $p = .001$, $t_2(8) = 33.04$, $p = .001$.

Task 2 manual responses. The statistical analysis of the manual response latencies yielded main effects of context, $F_1(1, 17) = 6.41$, $p = .021$, $F_2(1, 8) = 108.06$, $p = .001$, $minF'(1, 19) = 6.05$, $p = .02$, and SOA, $F_1(2, 34) = 180.58$, $p = .001$, $F_2(2, 16) = 3250.70$, $p = .001$, $minF'(2, 38) = 171.08$, $p < .001$. The main effect of context reveals that the effect of phonological preparation was propagated into the manual response latencies, which is the critical finding. Planned comparisons revealed that there was propagation of the preparation benefits into the manual response latencies at SOA = 0 ms, $t_1(17) = 2.08$, $p = .03$, $t_2(8) = 9.95$, $p = .001$, and at SOA = 300 ms, $t_1(17) = 1.75$, $p = .05$, $t_2(8) = 3.41$, $p = .004$, but there was no propagation at SOA = 1,000 ms, $t_1(17) = 0.14$, $p = .89$, $t_2(8) = 0.17$, $p = .87$. Thus, phonological encoding yielded dual-task interference at the short SOAs but not at the long one. Post hoc comparisons revealed that the manual response latencies were longer at SOA = 0 ms than at SOA = 300 ms, $t_1(17) = 8.55$, $p = .001$, $t_2(8) = 37.26$, $p = .001$, longer at SOA = 300 ms than at SOA = 1,000 ms, $t_1(17) = 10.97$, $p = .001$, $t_2(8) = 38.14$, $p = .001$, and longer at SOA = 0 ms than at SOA = 1,000 ms, $t_1(17) = 17.94$, $p = .001$, $t_2(8) = 96.39$, $p = .001$. There were no effects on the errors, all $ps > .34$.

In PRP experiments, the slope of the PRP curve (i.e., the tangent of the Task 2 response time by SOA curve) at short SOAs often approaches −1 (e.g., Pashler, 1994), indicating that for each unit of time the SOA decreases, the Task 2 response latency correspondingly increases. A slope of −1 is predicted if Task 1 processing precludes progress on Task 2 to the same extent at the short SOAs. The slope of the PRP curve at the short SOAs in the present experiment was −0.9, which is close to −1, as should be the case if progress on Task 2 depended on a gaze shift that allowed processing of the Task 2 stimulus.

The robustness of the effects reflected in the mean latencies was assessed by examining the latency distributions of the vocal responses, gaze shifts, and manual responses in the homogeneous and heterogeneous contexts for each SOA (cf. Luce, 1986). To obtain the latency distributions, the rank-ordered latencies for each participant were divided into deciles (10% quantiles) and mean latencies were computed for each decile separately for the vocal responses, gaze shifts, and manual responses in the homogeneous and heterogeneous contexts at each SOA. By averaging these

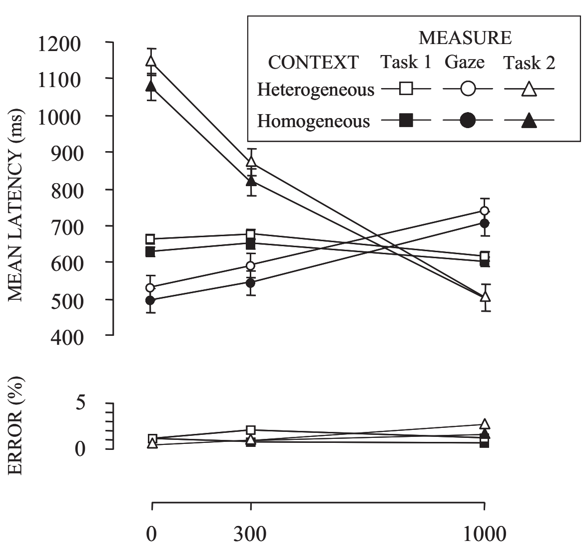

Figure 2. Mean latencies and error percentages per stimulus onset asynchrony (SOA) and context in Experiment 1. The upper panel shows the mean latencies for the vocal responses (Task 1), gaze shifts, and manual responses (Task 2). The error bars indicate the within-participant 95% confidence intervals. The lower panel shows the percentages of erroneous vocal and manual responses.

means across participants, Vincentized cumulative distribution curves were obtained (Ratcliff, 1979). Vincentizing the latency data across individual participants provides a way of averaging data while preserving the shapes of the individual distributions. Figure 3 shows the distributional plots for Experiment 1.

The upper panels of Figure 3 show that, except for the shortest latencies, the preparation benefit existed throughout the entire latency range for the vocal responses at all three SOAs. The middle panels show that the preparation benefit was also present across the entire latency range for the gaze shifts, again at all three SOAs. The lower panels of Figure 3 show that the manual responses exhibited the preparation benefit throughout the entire latency range at the short SOAs, whereas the preparation benefit was absent across the latency range for the long SOA. The latter was confirmed by a lack of statistical interaction of context (homogeneous, heterogeneous) and decile (1–10) for the manual responses at SOA = 1,000 ms, $F(9, 153) = 0.07$, $p = 1.0$. Thus, the latency distributions indicate that the effects observed for the mean latencies are robust. Manual responses were delayed and reflected the preparation benefit at the short SOAs but not at the long one, demonstrating dual-task interference from phonological encoding.

The upper panels of Figure 3 show that there is some tendency for preparation benefits to increase with increasing vocal latency. The dependence on absolute latency may suggest that anticipatory phonological preparation in the present experiment started after stimulus onset and happened in parallel with the normal word-planning process. Such a preparatory strategy has the consequence that its likelihood of being completed soon enough to facilitate word planning increases with the word-planning latency (cf. D. E. Meyer & Kieras, 1997a), as empirically observed. Moreover, inspection of the conditional accuracy curves (which specify accuracy as a function of response latency, e.g., Luce, 1986) revealed that more errors were made in the heterogeneous than in the homogeneous condition for the fastest responses. For example, for the 10% fastest responses, the error percentages in the heterogeneous and homogeneous conditions were 3.7 versus 0.0, 7.4 versus

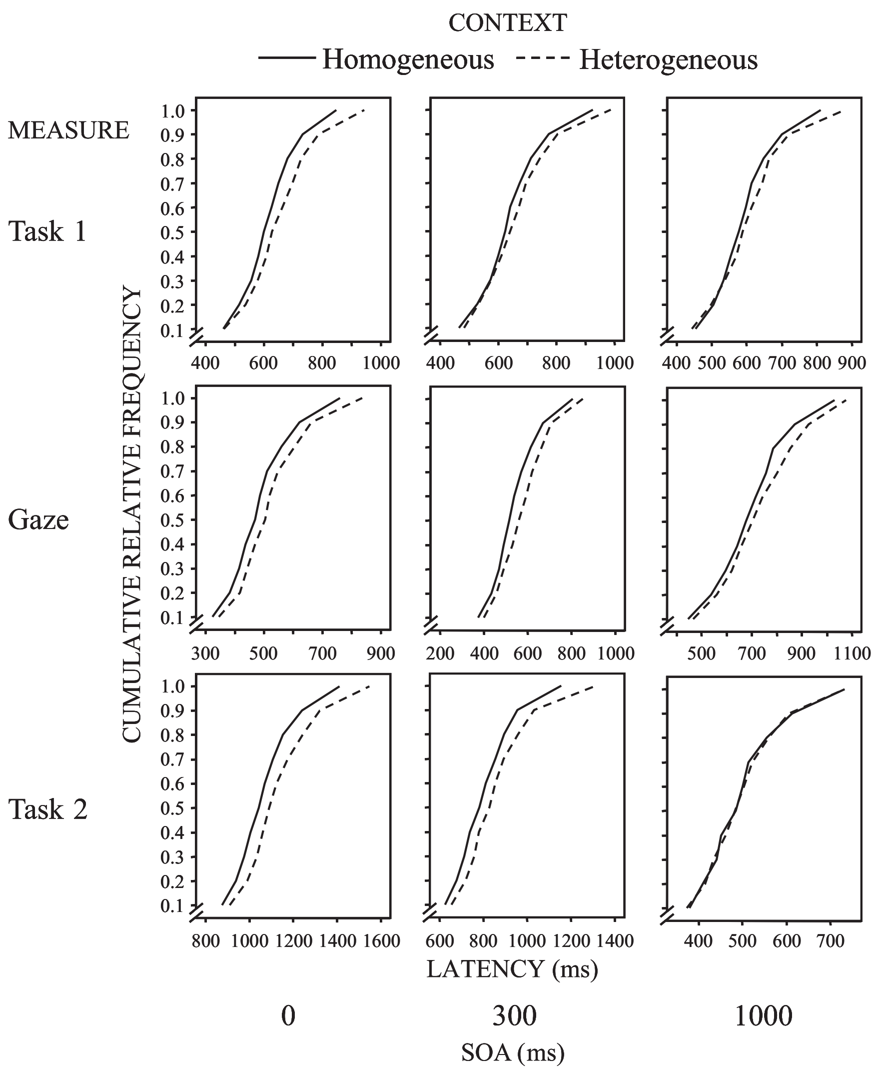

Figure 3. Vincentized cumulative distribution curves for the vocal responses (Task 1), gaze shifts, and manual responses (Task 2) in the homogeneous and heterogeneous contexts for each stimulus onset asynchrony (SOA) in Experiment 1.

3.7, and 9.2 versus 1.8 at the SOAs of 0, 300, and 1,000 ms, respectively. Thus, the absence of a preparation effect for the fastest responses may, at least partly, be due to a speed-accuracy tradeoff in the heterogeneous condition.

Most important for now, the results of the experiment show that phonological encoding in picture naming delayed the performance of an unrelated manual task. This suggests that the absence of dual-task interference observed by Ferreira and Pashler (2002) is not a general finding. Instead, by using another phonological manipulation and manual task, dual-task interference from phonological encoding can be obtained. The present results indicate that phonological encoding may yield dual-task interference in some experimental situations. This finding is more in line with an account of the dual-task interference in terms of attentional control (e.g., Allport, 1980, 1989; D. E. Meyer & Kieras, 1997a, 1997b) than in terms of a structural central response-selection bottleneck (Ferreira & Pashler, 2002; Pashler, 1994). If phonological encoding draws on a central response-selection mechanism, as the present data suggest following the logic of Ferreira and Pashler, phonological encoding should do so in all word production tasks requiring phonological encoding. However, Ferreira and Pashler argued on the basis of their own data that phonological encoding does not draw on a central response-selection mechanism. This suggests that the notion of an immutable, structural central response-selection bottleneck does not account well for the findings on phonological encoding.

The phonological effect in the present experiment was reflected in the latencies of the gaze shifts from picture to arrow. This finding replicates the eye-tracking results of A. S. Meyer et al. (2003) and A. S. Meyer and Van der Meulen (2000). However, whereas the latter studies used two vocal tasks, the dual-task interference from phonological encoding in the present experiment was obtained using a vocal–manual task combination. This suggests that dual-task interference from phonological encoding may be obtained even when there is no need to avoid response buffering (Levelt & Meyer, 2000). It cannot be excluded that visually attended arrow stimuli activate the phonological forms of their names, *left* and *right* (e.g., Morsella & Miozzo, 2002; Navarrete & Costa, 2005). Therefore, it is possible that gaze shifts were initiated only after phonological encoding had proceeded sufficiently to prevent interference from the phonological form activated by the arrow.

In the present experiment, the participants began to articulate the picture name after they shifted gaze to the arrows, except for the SOA of 1,000 ms. The overall difference in onset latency between the vocal responses and the gaze shifts was 132 ms at zero SOA (which corresponds to the SOA used by A. S. Meyer et al., 2003, and A. S. Meyer & Van der Meulen, 2000). It has been estimated that word-form encoding in picture naming takes about 350 ms if the total naming latency is about 600 ms (Indefrey & Levelt, 2004). The mean naming latency at zero SOA in the present experiment was 645 ms. The process of word-form encoding can be divided into morphological encoding, phonological encoding, and phonetic encoding. Estimates for the durations of these component processes are 80, 125, and 145 ms, respectively (Indefrey & Levelt, 2004). The observed vocal-to-gaze lag of 132 ms (roughly corresponding to the duration of phonetic encoding, which follows phonological encoding) independently confirms that the gaze shifts were triggered around the completion of the phonological encoding process, which explains why the preparation benefit was reflected in the gaze shift latencies. Sanders and colleagues (Sanders, 1998; Sanders & Lamers, 2002; Van Duren & Sanders, 1995) assumed that gaze shifts are based on perceptual processing and are initiated before response selection. A similar assumption was used by D. E. Meyer and Kieras (1997a, 1997b) in simulations of dual-task performance using EPIC-SRD. In contrast, the present findings indicate that gaze shifts in a PRP experiment with vocal and manual tasks may happen after response selection. Thus, it is not universally true that gaze shifts happen before response selection. Instead, the early gaze shifting observed by Van Duren and Sanders (1995) seems to be a strategy used in certain limited circumstances, such as the special case of only a single response.

The present results show that the phonological effect was propagated into the gaze shift latencies, not only at zero SOA, but also at the longer SOAs. Moreover, the gaze shift latencies increased with SOA. In particular, the gaze shifts were initiated about 54 ms later at SOA = 300 ms than at SOA = 0 ms, and they were initiated about 209 ms later at SOA = 1,000 ms than at SOA = 0 ms. The data suggest that the participants temporarily postponed the gaze shifts at the nonzero SOAs. Perhaps this is because participants try to avoid moving their gaze into empty space. Note that the propagation of the phonological effect forward into the gaze shift latencies at the nonzero SOAs suggests that the postponement of the gaze shifts was dependent on phonological encoding. If the gaze shifts had simply been initiated a constant time interval after picture onset or before arrow onset, the phonological effect should not have been present in the gaze shift latencies.

The vocal response latencies were longer at zero SOA than at SOA = 1,000 ms, and longer at SOA = 300 ms than at SOA = 1,000 ms, whereas the latencies did not differ between zero SOA and SOA = 300 ms. Perhaps, the presence of the Task 2 stimulus on the screen during Task 1 processes distracted Task 1 performance somewhat (cf. Yantis, 2000), leading to a slight delay in vocal responding. Alternatively, the increased vocal latencies at the short SOAs may reflect dual-task overhead costs (Navon & Gopher, 1979; Tombu & Jolicoeur, 2003). Overhead costs may relate to the active maintenance of two goals in working memory or maintaining preparations for two tasks. Once Task 1 has been completed, its maintenance is no longer needed and the overhead cost disappears. This suggestion is in line with some earlier studies, which showed that participants are sometimes faster at performing a given task alone than at performing it as the first of two tasks (e.g., Gottsdanker, Broadbent, & Van Sant, 1963). I further discuss the finding of increased vocal response latencies at the short SOAs in the General Discussion section.

To conclude, manual responses were delayed and reflected the advance phonological encoding in picture naming throughout the entire latency range at the short SOAs, demonstrating dual-task interference from phonological encoding. The absence of dual-task interference from phonological encoding observed by Ferreira and Pashler (2002) and the presence of the interference in the present experiment are more in line with an account of the interference in terms of attentional control than a structural response-selection bottleneck. The dual-task interference was obtained using a vocal–manual task combination. Thus, the interference can be obtained even when there is no need to avoid response buffering, contrary to the claim of Levelt and Meyer (2000). Anticipatory phonological encoding reduced picture naming and gaze shifting latencies at all SOAs, even though gaze shifts were postponed at the nonzero

SOAs. This suggests that the timing of gaze shifts in the PRP procedure depends on phonological encoding in picture naming, indicating that it is not universally true that gaze shifts occur before response selection (D. E. Meyer & Kieras, 1997a, 1997b; Sanders, 1998).

Experiment 2

The experimental situation of Experiment 1 differed from the situation in the experiments of Ferreira and Pashler (2002). In the present Experiment 1, the Task 1 and Task 2 stimuli were both visual and they were located in different spatial positions. Thus, task performance required gaze shifts. To examine whether the findings of Experiment 1 can be replicated using an experimental situation that is closer to that of Ferreira and Pashler, the second experiment used displays that were identical to those of Experiment 1 except that the arrows were now centered on the pictures. The tasks for the participants were to name the picture and to respond to the arrow as in Experiment 1. The latencies of the vocal responses and manual responses were again recorded. The arrows appeared in the middle of the pictures. Given the absence of a visual angle between the middle of the picture and arrow, the notion of gaze shift was undefined (i.e., eye fixations on the arrow were still within the contours of the picture). Therefore, eye movements were not recorded.

Method

The method was the same as that of Experiment 1, except that the pictures and arrows were superimposed. Gaze shifts were not required.

Results and Discussion

The upper panel of Figure 4 shows for each context and SOA condition the mean latencies for the vocal responses (Task 1) and manual responses (Task 2). The lower panel of Figure 4 shows for each condition the mean error percentages for the vocal and manual responses.

Task 1 vocal responses. The statistical analysis of the vocal response latencies yielded main effects of context (revealing a preparation benefit of 27 ms, on average), $F_1(1, 17) = 9.08$, $p = .008$, $F_2(1, 8) = 43.21$, $p = .001$, $minF'(1, 23) = 7.50$, $p = .01$, and SOA, $F_1(2, 34) = 8.16$, $p = .001$, $F_2(2, 16) = 47.36$, $p = .001$, $minF'(2, 44) = 6.96$, $p = .002$. The effect of context did not depend on the SOA, $F_1(2, 34) = 0.07$, $p = .93$, $F_2(2, 16) = 0.07$, $p = .93$. Thus, as in Experiment 1, a preparation benefit was obtained, which was constant across SOAs. Post hoc comparisons revealed that the vocal response latencies were longer at zero SOA than at SOA = 1,000 ms, $t_1(17) = 4.80$, $p = .001$, $t_2(8) = 7.19$, $p = .001$, and longer at SOA = 300 ms than at SOA = 1,000 ms, $t_1(17) = 3.22$, $p = .003$, $t_2(8) = 10.34$, $p = .001$, whereas the latencies did not differ between zero SOA and SOA = 300 ms, $t_1(17) = 0.49$, $p = .63$, $t_2(8) = 1.29$, $p = .23$. There were no effects on the errors, all $ps > .23$.

Task 2 manual responses. The statistical analysis of the manual response latencies yielded main effects of context, $F_1(1, 17) = 6.79$, $p = .018$, $F_2(1, 8) = 30.32$, $p = .001$, $minF'(1, 23) = 5.55$, $p = .03$, and SOA, $F_1(2, 34) = 133.28$, $p = .001$, $F_2(2, 16) = 3168.14$, $p = .001$, $minF'(2, 37) = 127.90$, $p < .001$. The main

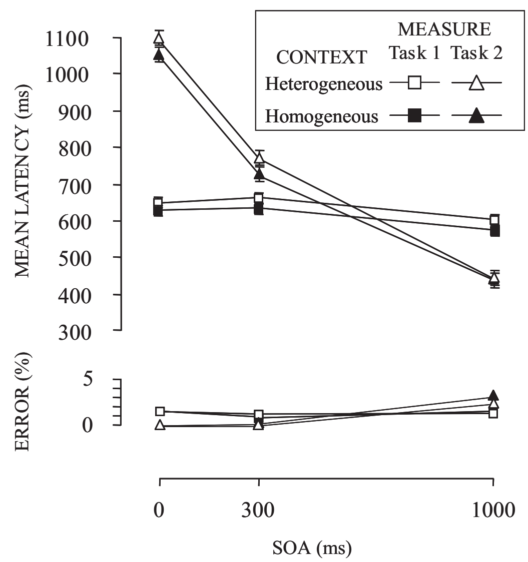

Figure 4. Mean latencies and error percentages per stimulus onset asynchrony (SOA) and context in Experiment 2. The upper panel shows the mean latencies for the vocal responses (Task 1) and manual responses (Task 2). The error bars indicate the within-participant 95% confidence intervals. The lower panel shows the percentages of erroneous responses.

effect of context reveals that the effect of phonological preparation was propagated into the manual response latencies, which is the critical finding. Planned comparisons revealed that there was propagation of the preparation benefits at SOA = 0 ms, $t_1(17) = 2.17$, $p = .02$, $t_2(8) = 3.34$, $p = .005$, and at SOA = 300 ms, $t_1(17) = 2.12$, $p = .03$, $t_2(8) = 3.44$, $p = .005$, but there was no propagation at SOA = 1,000 ms, $t_1(17) = 0.8$, $p = .43$, $t_2(8) = 1.25$, $p = .25$. Post hoc comparisons showed that the manual response latencies were longer at SOA = 0 ms than at SOA = 300 ms, $t_1(17) = 13.63$, $p = .001$, $t_2(8) = 39.04$, $p = .001$, longer at SOA = 300 ms than at SOA = 1,000 ms, $t_1(17) = 6.41$, $p = .001$, $t_2(8) = 38.94$, $p = .001$, and longer at SOA = 0 ms than at SOA = 1,000 ms, $t_1(17) = 15.55$, $p = .001$, $t_2(8) = 83.22$, $p = .001$. There were no effects on the errors (all $ps > .05$), except for a main effect of SOA ($ps < .001$). Thus, as in Experiment 1, phonological encoding yielded dual-task interference at the short SOAs but not at the long one.

The slope of the PRP curve in Experiment 1 approached −1, indicating that progress on Task 2 depended on the gaze shifts allowing processing of the Task 2 stimulus. The slope of the curve in the present experiment was −1.1, suggesting that, again, processing of the Task 2 stimulus was postponed until Task 1 had progressed far enough. Interestingly, the slope of the PRP curve was slightly steeper than −1. Thus, the decrease in PRP effect by increasing the SOA was larger than the magnitude of the SOA. In terms of a task-shift criterion model, this may suggest that the criterion for shifting to Task 2 was set slightly more conservatively at zero SOA than at SOA = 300 ms, perhaps because out-of-order responses are more likely as the SOA decreases (cf. D. E. Meyer & Kieras, 1997b).

To conclude, the results of the experiment again show that phonological encoding in picture naming can delay the performance of an unrelated manual task. The experiment replicates the findings of Experiment 1, now with superimposed pictures and arrows. The replication indicates that the dual-task interference from phonological encoding in Experiment 1 is not due to the use of spatially separated stimuli requiring a gaze shift.

As is evident from Figures 2 and 4, and from the similarity in slope of the PRP curves, there is a striking resemblance between the patterns of results for the vocal and manual responses in Experiments 1 and 2, even though gaze shifts had to be made between the task stimuli in Experiment 1 but not in Experiment 2. Why are the patterns of results for the vocal and manual responses so similar between the two experiments? Given that both task stimuli were visual, one would expect that gaze shifts created a perceptual bottleneck (i.e., processing of the arrow was impossible as long as the picture was fixated) in Experiment 1, but not in Experiment 2. The arrows were small compared to the pictures. The arrows, $<$ and $>$, subtended less than 1° of visual angle horizontally and vertically, whereas the angle was about 9° for the pictures. Moreover, the arrows were flanked by two Xs on each side, making them difficult to discriminate. It is possible that the perceptual spatial resolution required for identifying the pictured object was insufficient for discriminating the direction of the arrow. This may have created a perceptual bottleneck in Experiment 2, even though gaze shifts were not required. An alternative or additional possibility is that the participants in the two experiments used similar process-scheduling strategies (i.e., comparable perceptual strategies and similar criteria for shifting between tasks). An account along these lines is proposed later and its utility to account for the present findings is tested through computer simulations.

Difference With Ferreira and Pashler (2002)

Whereas Ferreira and Pashler (2002) obtained no dual-task interference from phonological encoding in picture naming on an unrelated manual task, such interference was obtained in the present Experiments 1 and 2. According to the attentional process-scheduling account (Roelofs, 2007), the difference in dual-task interference between studies reflects a difference in the task-shift criterion. But why would there be a difference in the task-shift criterion between the experiment of Ferreira and Pashler and the present ones? It is possible that task-shift criteria differed because of methodological differences between experiments.

The phonological manipulation of Ferreira and Pashler (2002) was different from the manipulation in the present experiments. It was a block-to-block manipulation in the present Experiments 1 and 2, allowing phonological planning ahead of time, whereas Ferreira and Pashler randomly presented related or unrelated written distractors superimposed onto the pictures. Before trial onset, the participants in the experiments of Ferreira and Pashler did not know whether they would see related or unrelated distractors. Therefore, they did not know whether they were going to be getting a relatively easy trial or not. Lupker, Brown, and Colombo (1997) argued that such design differences may affect criterion setting. Whereas D. E. Meyer and Kieras (1997a, 1997b), A. S. Meyer et al. (2003), and Roelofs (2007) assumed a task-shift criterion based on the accumulation of information, Lupker et al. (1997) proposed a criterion based on time. According to the time-based view, the participants of Ferreira and Pashler (2002) set a time criterion that would handle most trials, because of the mixing of trial types. As a result, the engagement of Task 2 would start at about the same time for all trials regardless of the nature of the Task 1 stimulus, producing no phonological encoding effect on Task 2 latencies, as empirically observed. In contrast, in the present Experiments 1 and 2, participants could set a different time criterion in different blocks (i.e., a later one for the heterogeneous blocks than for the homogeneous blocks). That difference would then transfer to the gaze shift latency and the Task 2 latencies. As a result, one would expect to see the observed difference in dual-task interference from phonological encoding between studies.

However, a time criterion fails to account for the observation by Ferreira and Pashler (2002) that the effect of semantically related versus unrelated distractors was propagated into the latencies of the responses to the tones. If participants set a fixed time criterion that handles most trials so that phonological effects are not propagated into the manual responses, the semantic effects should not be propagated either, unlike what Ferreira and Pashler observed. Thus, it is unlikely that the difference in dual-task interference from phonological encoding between studies is due to the mixing or blocking of trial types.

Another difference between the experiments of Ferreira and Pashler (2002) and the present experiments concerns the modality of the Task 2 stimuli. Whereas the arrow discrimination task in the present experiments involved manual responses to visual stimuli, the experiments of Ferreira and Pashler involved a tone discrimination task. It may be that participants feel more pressed to shift to a manual task with tones than with arrows. If so, they may set the task-shift criterion earlier in the situation with a secondary tone discrimination task (Ferreira & Pashler) than an arrow discrimination task (the present Experiments 1 and 2). The tones were presented for 285 ms in the study of Ferreira and Pashler, whereas the arrows remained on the screen until manual responding in the present Experiments 1 and 2. It may be that participants shifted earlier to the tones (Ferreira & Pashler) than to the arrows (present Experiments 1 and 2) because of the limited duration of the tones. This was tested in Experiment 3.

Experiment 3

Experiment 3 examined the possibility that participants shifted attention earlier to the tones (Ferreira & Pashler, 2002) than to the arrows (present Experiments 1 and 2) because of the limited duration of the tones. In the experiment, the presentation duration of the arrows was set to 300 ms, which corresponds to the duration of the tones in the study of Ferreira and Pashler.

Method

The method was the same as that of Experiment 2, except that the arrows were now presented for only 300 ms.

Results and Discussion

The upper panel of Figure 5 shows for each context and SOA condition the mean latencies for the vocal responses (Task 1) and manual responses (Task 2). The lower panel of Figure 5 shows for each condition the mean error percentages for the vocal and manual responses.

Task 1 vocal responses. The statistical analysis of the vocal response latencies yielded main effects of context (revealing a preparation benefit of 27 ms, on average), $F_1(1, 17) = 7.32$, $p=$

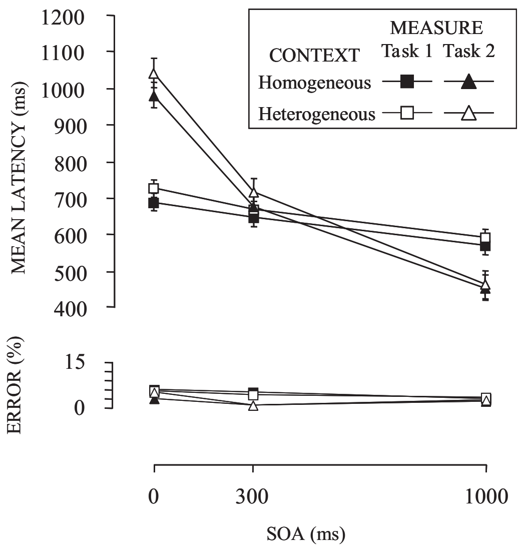

Figure 5. Mean latencies and error percentages per stimulus onset asynchrony (SOA) and context in Experiment 3. The upper panel shows the mean latencies for the vocal responses (Task 1) and manual responses (Task 2). The error bars indicate the within-participant 95% confidence intervals. The lower panel shows the percentages of erroneous responses.

.015, $F_2(1, 8) = 35.40$, $p < .001$, $minF'(1, 23) = 6.07$, $p = .02$, and SOA, $F_1(2, 34) = 20.11$, $p < .001$, $F_2(2, 16) = 96.67$, $p < .001$, $minF'(2, 45) = 16.65$, $p < .001$. The effect of context did not depend on the SOA, $F_1(2, 34) = 0.34$, $p = .71$, $F_2(2, 16) = 2.10$, $p = .16$. Thus, as in Experiments 1 and 2, a preparation benefit was obtained, which was constant across SOAs. Post hoc comparisons revealed that the vocal response latencies were longer at zero SOA than at SOA = 1,000 ms, $t_1(17) = 6.83$, $p < .001$, $t_2(8) = 11.23$, $p < .001$, longer at SOA = 300 ms than at SOA = 1,000 ms, $t_1(17) = 4.07$, $p = .001$, $t_2(8) = 16.37$, $p < .001$, and at zero SOA and SOA = 300 ms, the difference was nonsignificant, $t_1(17) = 2.10$, $p = .05$, $t_2(8) = 4.87$, $p = .001$. There were no effects on the errors (all *p*s > .60), except for a main effect of SOA (*p*s < .04).

Task 2 manual responses. The statistical analysis of the manual response latencies yielded main effects of context, $F_1(1, 17) = 5.05$, $p = .038$, $F_2(1, 8) = 27.70$, $p = .001$, $minF'(1, 22) = 4.27$, $p = .05$, and SOA, $F_1(2, 34) = 152.79$, $p < .001$, $F_2(2, 16) = 1305.95$, $p < .001$, $minF'(2, 41) = 136.79$, $p < .001$. The main effect of context reveals that the effect of phonological preparation was propagated into the manual response latencies, which is the critical finding. Planned comparisons revealed that there was propagation of the preparation benefits at SOA = 0 ms, $t_1(17) = 2.10$, $p = .05$, $t_2(8) = 3.87$, $p = .005$, but not at SOA = 300 ms, $t_1(17) = 1.33$, $p = .20$, $t_2(8) = 3.26$, $p = .01$, and not at SOA = 1,000 ms, $t_1(17) = 0.7$, $p = .47$, $t_2(8) = 2.79$, $p = .02$. Whereas in the preceding experiments, the phonological effect was carried forward to the Task 2 latencies at the SOAs of 0 and 300 ms, propagation now only occurred at zero SOA. However, numerically, there was a 41-ms effect at SOA = 300 ms in the present experiment, which may not have reached statistical significance because of greater noise. Post hoc comparisons showed that the manual response latencies were longer at SOA = 0 ms than at SOA = 300 ms, $t_1(17) = 11.16$, $p < .001$, $t_2(8) = 24.63$, $p < .001$, longer at SOA = 300 ms than at SOA = 1,000 ms, $t_1(17) = 7.92$, $p < .001$, $t_2(8) = 28.39$, $p < .001$, and longer at SOA = 0 ms than at SOA = 1,000 ms, $t_1(17) = 15.22$, $p < .001$, $t_2(8) = 50.62$, $p < .001$. There were no effects on the errors (all *p*s > .12), except for a main effect of SOA (*p*s < .008). Thus, as in Experiments 1 and 2, phonological encoding yielded dual-task interference at the short SOA.

The experiment replicates the findings of Experiment 2, now with limited presentation duration of the arrows. The replication indicates that the dual-task interference from phonological encoding in Experiment 2 is not due to the presentation of the arrow until manual response onset. The slope of the PRP curve in Experiment 2 was −1.1, and it was −1.05 in the present experiment. Thus, in both Experiments 2 and 3, the slope approached −1, which implies that Task 1 processing precluded progress on Task 2 to the same extent in both experiments. However, there was also a difference. Whereas the total time for both tasks at SOA = 0 ms almost equaled the sum of the time to complete each task by itself in Experiment 2 (only a 5 ms difference), the total time was much less (i.e., 156 ms) than the sum in Experiment 3. In Experiment 2, the average manual response latency at SOA = 0 ms was 1,079 ms and the sum of the vocal response latency (at SOA = 0 ms) and the manual response latency (at SOA = 1,000 ms, an estimate of the baseline latency) was 1,074 ms. However, in Experiment 3, the manual response latency at SOA = 0 ms was 1,168 ms and the sum of the vocal response latency (at SOA = 0 ms) and the manual response latency (at SOA = 1,000 ms) was 1,012 ms. In terms of attentional control, this difference between experiments suggests that the Task 2 suspension point was located later during the course of Task 2 in Experiment 3 than in Experiment 2 (cf. D. E. Meyer & Kieras, 1997a, 1997b). This difference presumably arose because the arrow had to be processed immediately because of its limited presentation duration in Experiment 3 but not in Experiment 2.

Experiment 4

The results of Experiment 3 suggest that limiting the Task 2 stimulus duration is not promoting an early disengagement from Task 1. Another possibility is that participants shifted earlier to the tones (Ferreira & Pashler, 2002) than to the arrows (present Experiments 1–3) because producing speech hampers tone discrimination. Brain imaging studies have demonstrated that during self-produced speech, the auditory cortex suppresses its response to acoustic signals, including tones (Houde et al., 2002). Similarly, vocalization by monkeys yields suppression of activity in auditory cortex (Müller-Preuss et al., 1980; Müller-Preuss & Ploog, 1981), which already begins several hundred milliseconds prior to the onset of vocalization (Eliades & Wang, 2003). Based on these findings, one expects that in a PRP experiment with vocal Task 1 responses and auditory Task 2 stimuli, the planning and articulation of the speech sounds for Task 1 may hinder the auditory recognition of the Task 2 stimuli. Thus, with tones as Task 2 stimuli, Task 2 processing may need to be protected against interference from Task 1 processing (cf. Allport, 1989; Roelofs, 2007). This might be achieved through attentional enhancement of the processing of the tone (cf. Benedict et al., 2002; Crottaz-Herbette & Menon, 2006). The greater attention demands of tone discrimination than arrow discrimination in the context of vocal responding may explain why attentional engagement shifts earlier between tasks with Task 2 tone discrimination (Ferreira & Pashler,

2002) than with Task 2 arrow discrimination (Experiments 1–3). If the participants of Ferreira and Pashler had set the task-shift criterion before the onset of the Task 1 phonological encoding process and the participants in Experiments 1–3 had set the criterion after phonological encoding, then the difference in dual-task interference from phonological encoding between studies is readily explained.

Experiment 4 tested the Task 2 modality hypothesis by replacing the arrow discrimination task by a tone discrimination task. Instead of presenting a left- or right-pointing arrow, a tone with a low or high pitch was presented over headphones, and participants pressed a left or right button to indicate the pitch of the tones. The primary task and the SOAs were the same as those in Experiments 1–3. If the difference in dual-task interference between studies is due to a difference in secondary task (tone vs. arrow discrimination), the phonological preparation effect for the vocal responses should not be propagated into the manual response latencies for the tone task at any SOA in the present experiment. Such a finding would replicate the results of Ferreira and Pashler (2002).

Method

The method was the same as that of Experiment 3, except that tones instead of arrows were presented. On each trials, a tone with a low or high pitch (600 or 1200 Hz, respectively) was presented over headphones for 300 ms. Participants pressed a left button in response to the low tone and a right button in response to the high tone.

Results and Discussion

The upper panel of Figure 6 shows for each context and SOA condition the mean latencies for the vocal responses (Task 1) and manual responses (Task 2). The lower panel of Figure 6 shows for each condition the mean error percentages for the vocal and manual responses.

Task 1 vocal responses. The statistical analysis of the vocal response latencies yielded main effects of context (revealing a preparation benefit of 38 ms, on average), $F_1(1, 17) = 6.71$, $p = .019$, $F_2(1, 8) = 41.20$, $p = .001$, $minF'(1, 22) = 5.77$, $p = .03$, and SOA, $F_1(2, 34) = 13.71$, $p = .001$, $F_2(2, 16) = 1271.81$, $p = .001$, $minF'(2, 35) = 13.56$, $p < .001$. The effect of context did not depend on the SOA, $F_1(2, 34) = 0.15$, $p = .87$, $F_2(2, 16) = 0.94$, $p = .41$. Thus, as in the other experiments, a preparation benefit was obtained, which was constant across SOAs. Post hoc comparisons showed that the vocal response latencies were longer at zero SOA than at SOA = 1,000 ms, $t_1(17) = 5.44$, $p = .001$, $t_2(8) = 33.18$, $p = .001$, and longer at SOA = 300 ms than at SOA = 1,000 ms, $t_1(17) = 4.38$, $p = .001$, $t_2(8) = 45.51$, $p = .001$. The difference between zero SOA and SOA = 300 ms was not significant in the analysis by participants, $t_1(17) = 1.6$, $p = .07$, and significant by items, $t_2(8) = 17.46$, $p = .001$. There were no effects on the errors, all $ps > .10$.

Task 2 manual responses. The statistical analysis of the manual response latencies yielded no main effect of context, $F_1(1, 17) = 0.46$, $p = .51$, $F_2(1, 8) = 1.44$, $p = .27$, but there was an effect of SOA, $F_1(2, 34) = 79.03$, $p = .001$, $F_2(2, 16) = 3073.44$, $p = .001$, $minF'(2, 36) = 77.05$, $p < .001$. Planned comparisons revealed that there was no propagation of the preparation benefits at any SOA: SOA = 0 ms, $t_1(17) = 0.73$, $p = .48$, $t_2(8) = 0.83$, $p = .43$; SOA = 300 ms, $t_1(17) = 0.0$, $p = .99$, $t_2(8) = 0.51$, $p = .63$; SOA = 1,000 ms, $t_1(17) = 0.65$, $p = .53$, $t_2(8) = 1.29$, $p = .23$. Thus, unlike the first three experiments, phonological encoding yielded no dual-task interference at any of the SOAs. Post hoc comparisons revealed that the manual response latencies were longer at SOA = 0 ms than at SOA = 300 ms, $t_1(17) = 5.49$, $p = .001$, $t_2(8) = 34.84$, $p = .001$, longer at SOA = 300 ms than at SOA = 1,000 ms, $t_1(17) = 7.03$, $p = .001$, $t_2(8) = 45.48$, $p = .001$, and longer at SOA = 0 ms than at SOA = 1,000 ms, $t_1(17) = 12.71$, $p = .001$, $t_2(8) = 75.56$, $p = .001$. There were no effects on the errors (all $ps > .05$), except for a main effect of SOA ($ps < .006$).

The slope of the PRP curve in Experiments 1–3 approached −1, indicating that Task 1 processing precluded progress on Task 2 to the same extent in the three experiments. The slope of the PRP curve in the present experiment was only 0.73, suggesting that Task 2 processing was resumed earlier during Task 1 processing in the present experiment than in the previous two experiments. This finding agrees with the observation that phonological encoding did not cause dual-task interference in the present experiment whereas it did in the other experiments. Moreover, whereas the percentage of vocal response errors was less than 5% in the first two experiments, it was around 10% in the present experiment, in line with the assumption that attention disengaged earlier in the present experiment than in the other ones.

To conclude, the present results show that the phonological preparation effect for the vocal responses is not propagated into the manual responses to the tones, unlike what was observed for the manual responses to the arrows in Experiments 1–3. The present finding for the tone discrimination task replicates the results of Ferreira and Pashler (2002). Earlier research suggested that vocal responding may hamper auditory recognition. The present results suggest that the moment of shifting between tasks depends on the

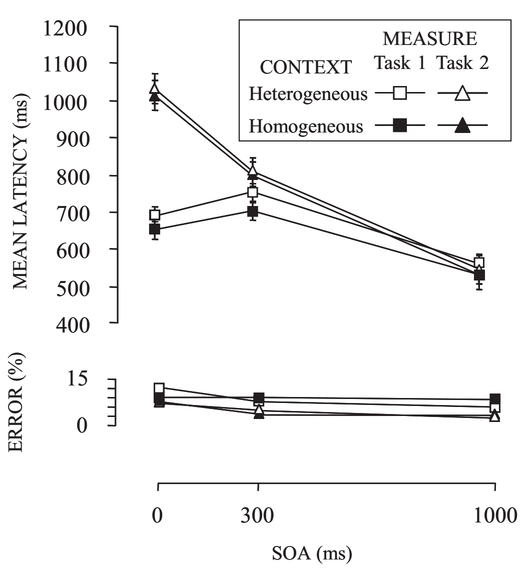

Figure 6. Mean latencies and error percentages per stimulus onset asynchrony (SOA) and context in Experiment 4. The upper panel shows the mean latencies for the vocal responses (Task 1) and manual responses (Task 2). The error bars indicate the within-participant 95% confidence intervals. The lower panel shows the percentages of erroneous responses.

attention demands of the secondary task, which is in line with the strategic but not with the structural account.

Computer Simulations

In this section, I demonstrate that the findings on phonological encoding in the literature (i.e., Ferreira & Pashler, 2002; A. S. Meyer et al., 2003; A. S. Meyer & Van der Meulen, 2000) and those from the present experiments may be explained by a simple model of attentional control in the coordination of vocal responding, gaze shifting, and manual responding. The model is an attempt at cumulative modeling (cf. Logan, 2004; Roelofs, 2005), combining assumptions of the EPIC-SRD model of dual-task performance (D. E. Meyer & Kieras, 1997a, 1997b) and the WEAVER++ model of word planning (Levelt et al., 1999; Roelofs, 1992, 1997, 2003). Both EPIC-SRD and WEAVER++ account for a wide variety of findings in their domains of application. Figure 7 illustrates the proposed model for Experiment 1 (cf. Roelofs, 2007). The figure illustrates how, according to the model, the attentional control process coordinates the multiple threads of processing in phonological preparation, vocal responding, gaze shifting, and manual responding.

The model assumes that to maintain acceptable levels of speed and accuracy, to minimize resource consumption and crosstalk between tasks, and to satisfy instructions about task priorities, participants set a criterion for when the shift between the vocal and manual tasks should occur (cf. D. E. Meyer & Kieras, 1997a, 1997b; Sperling & Dosher, 1986). In EPIC-SRD, reaching the task-shift criterion is called the occurrence of the "Task 1 unlocking event" (D. E. Meyer & Kieras, 1997b, p. 753), which unlocks Task 2. The point at which Task 2 processing is strategically suspended is called the "Task 2 lockout point." Presumably, the position of the task-shift criterion within the Task 1 process is determined on the basis of the initial trials of an experiment, when participants become familiar with the experimental situation, and the criterion stays more or less constant throughout the experiment (cf. Treisman & Williams, 1984). At the beginning of each trial, the attentional control process enables both tasks, engages on Task 1 and temporarily suspends Task 2, instructs the oculomotor system to direct gaze towards the Task 1 stimulus, and maintains engagement on Task 1 and monitors performance until the task process reaches the task-shift criterion. Moreover, in homogeneous sets, the phonological encoder is instructed to prepare the syllable that is shared by the responses in a set. Also during the planning of the target word, a saccade to the arrow is prepared. The oculomotor system may use low-spatial frequency information to determine the position of the arrow or the position may be retrieved from memory. When the task-shift criterion is reached during the course of Task 1, attention disengages from Task 1 and shifts to Task 2, which is then resumed, directly followed by a signal to the saccadic control system to execute the prepared saccade to the Task 2 stimulus. EPIC-SRD may accommodate this proposal about gaze shifting by assuming that the planning of the saccade to the Task 2 stimulus is issued early, whereas the signal to execute the prepared saccade (i.e., to move the eyes) depends on the Task 1 unlocking event.

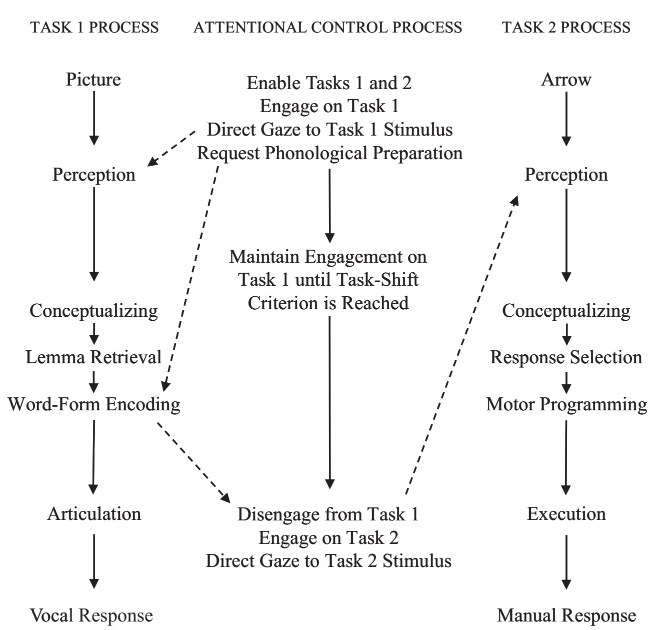

Figure 7. A model of attentional control in the coordination of vocal responding, gaze shifting, and manual responding. The model combines assumptions of the EPIC-SRD model of dual-task performance (D. E. Meyer & Kieras, 1997a, 1997b) and the WEAVER++ model of spoken word production (Levelt et al., 1999; Roelofs, 1992, 1997, 2003, 2007).

D. E. Meyer and Kieras (1997b) noted the similarity between strategic process-scheduling models of dual-task performance and classic signal-detection theory (e.g., Tanner & Swets, 1954). According to this theory, participants performing a signal-detection task set their decision criterion strategically to achieve preferred frequencies of hits for signals and correct rejections for noise (see Treisman & Williams, 1984, for a theory of how participants set and maintain the criterion and adjust it to take account of relevant events). Similarly, according to a strategic process-scheduling model of dual-task performance, participants set their task-shift criterion to achieve preferred levels of speed and accuracy, depending on prevailing experimental circumstances. A related proposal has been made for the initiation of responding in single-task performance by A. S. Meyer et al. (2003). Sperling and Dosher (1986) present a thorough comparison of criterion setting in dual-task performance and signal detection.

According to the WEAVER++ model (Roelofs, 1992, 1997, 2003), picture naming involves picture perception, conceptualizing, lemma retrieval, word-form encoding, and articulation. Words are planned by spreading activation through a lexical network and application of condition-action production rules. In securing task-relevant control, the model uses goal symbols, whose presence constitutes one of the conditions for the firing of a production rule. When a goal symbol is placed in working memory, the attention of the system is focused on those production rules that include this goal among their conditions. The EPIC-SRD model of dual-task performance advanced by D. E. Meyer and Kieras (1997a, 1997b) also uses goal-factored production rules to achieve executive control. EPIC's executive production rules coordinate cognitive processes across the tasks in a dual-task situation. The processes illustrated in Figure 7 may be computationally implemented by combining WEAVER++'s word planning and attentional control processes and EPIC's task coordination processes. In particular, WEAVER++ may be extended by including executive production rules for dual-task goal setting and task shifting and for directing gaze (cf. Roelofs, 2003, 2007).

The proposed model of attentional control may accommodate the difference in results between earlier studies (e.g., between Ferreira & Pashler, 2002, and A. S. Meyer et al., 2003, and A. S. Meyer & Van der Meulen, 2000) in terms of differences in task-shift criterion. Presumably, the task-shift criterion was set earlier in Experiment 2 of Ferreira and Pashler than in the experiments of A. S. Meyer et al. (2003) and A. S. Meyer and Van der Meulen (2000). In the latter experiments, two vocal responses were required on each trial. A late task-shift criterion avoided interference in phonological encoding and response buffering. Such a late task-shift criterion was not required in the study of Ferreira and Pashler, where the second response was a manual one. However, the present Experiments 1–3 with secondary manual responses indicate that dual-task interference from phonological encoding may be obtained even with a difference in response modality between Task 1 and Task 2. This suggests that participants sometimes play safe by setting a late task-shift criterion even when it is not strictly required. However, if Task 1 responding hampers the perception of the Task 2 stimulus, as may be the case with Task 1 vocal responding and Task 2 auditory stimuli (Experiment 4 and the experiments of Ferreira & Pashler), the task-shift criterion is set earlier during the Task 1 process than otherwise would have been the case (Experiments 1–3).

Computer simulations of Experiments 1 and 2 using the proposed model demonstrate the utility of the theoretical approach. The aim of the simulations was to compute the consequences of the type of task scheduling depicted in Figure 7. In the simulations, details of component processes were unimportant, but only the latencies of the processes and their interdependencies mattered (cf. Schweickert, 1980). For example, the Task 2 process was represented in the simulations by a single constant latency. First the model was applied to Experiment 1. The vocal latencies were obtained by running word-form encoding in WEAVER++ and adding a constant residual latency for the processes of perception, conceptualizing, lemma retrieval, and the initiation of articulation (cf. Luce, 1986), yielding a mean vocal latency of 640 ms. The WEAVER++ parameters were identical to those of Roelofs (1997). To account for the longer latency of vocal responding at zero SOA and SOA = 300 ms than at SOA = 1,000 ms, a dual-task overhead cost was assumed, which affected the central process of conceptualizing (shared between tasks, see Figure 7), slowing Task 1 processing (set to 40 ms). The task-shift criterion was reached by the completion of phonological encoding, set to 130 ms before articulation onset. The eye-tracking data suggest that gaze shifting at the nonzero SOAs was temporarily postponed relative to the completion of phonological encoding, depending on the SOA (the postponement was set to 0.2 × SOA). The mean saccade duration was set to 70 ms. Finally, the mean manual response latency was set to 500 ms. The values of the six parameters were informally estimated from the real data. Note that the number of parameters is much less than the number of data points (i.e., 18), making the fit of the model to the data nontrivial.

Figure 8 shows the simulation results together with the data obtained in Experiment 1. The fit between model and data was good (R^2 = .99, meaning that the model accounts for 99% of the systematic variance in the empirical mean latencies). Next, the model was applied to Experiment 2, now without the gaze shifts. The parameter values were exactly the same as those in the simulations of Experiment 1 (i.e., the fit is parameter free). The fit between model and data was again good (R^2 = .99). The good model fits demonstrate the utility of the theoretical approach. To account for the absence of dual-task interference from phonological preparation in Experiment 4, it has to be assumed that the task-shift criterion was set before the onset of phonological encoding.

The proposed model of attentional control in task coordination assumes that the postponement of Task 2 responses is strategically determined by executive control processes rather than due to an immutable central bottleneck in word planning, as assumed by Ferreira and Pashler (2002). According to the proposed model, all stages of spoken word production, except for conceptualizing, are carried out by separate, dedicated processing mechanisms (Roelofs, 2003). Unlike what Ferreira and Pashler propose, lemma retrieval and morphological encoding do not draw on a central

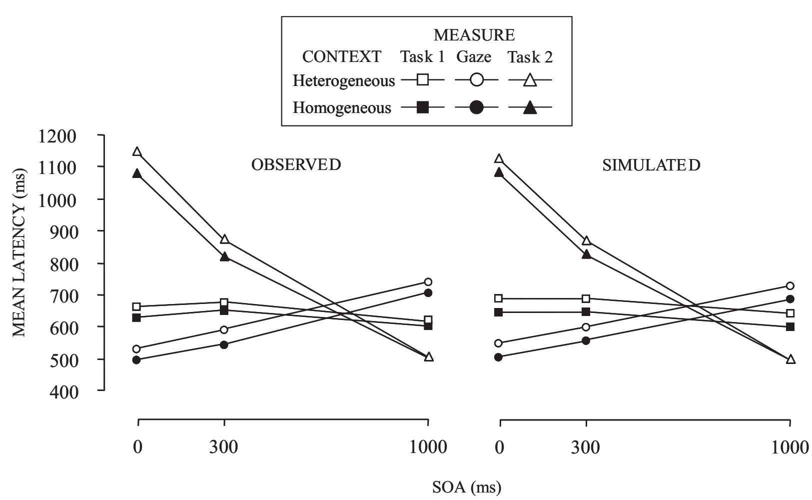

Figure 8. Simulation of the coordination of vocal responding (Task 1), gaze shifting, and manual responding (Task 2) in Experiment 1 using the attentional control model. The left-hand panel shows the real latency data of Experiment 1 and the right-hand panel shows the results of the computer simulations. SOA = stimulus onset asynchrony.

response-selection mechanism. The proposed attentional control process does not contain any mechanism that is needed for performing the individual subordinate task processes. As Logan, Schachar, and Tannock (2000) put it:

> *Subordinate processes* do the basic computations involved in performing a task. They are part of the chain of processes that lead from stimulus to response, taking input from the stimulus or stimulus-driven processes, and giving output to response systems or the processes that drive them. *Executive processes* are outside the chain, but they act on it; they control the subordinate processes, enabling them and directing them, turning them on and off (pp. 653–654, original italics).

Moreover, "the task of controlling other processes is just as specialised a function as that of recognising phonological patterns or rotating images" (Monsell, 1996, pp. 101–102). Figure 7 lists the specialized component attentional processes that are assumed for the control of tasks in the present experiments.

According to the proposed model of attentional control, if the task-shift criterion is located before the onset of phonological encoding in vocal responding, manipulations of phonological encoding will not be propagated into manual response latencies, as observed by Ferreira and Pashler (2002) and in the present Experiment 4. However, if the task-shift criterion is located after the completion of phonological encoding in vocal responding, manipulations of phonological encoding will be propagated into manual response latencies, as observed in the present Experiments 1–3. Moreover, the model predicts that if the durations of lemma retrieval and morphological encoding are manipulated, resulting latency effects will propagate into manual response latencies, as observed by Ferreira and Pashler. To conclude, differences in task-shift criterion may explain differences in dual-task interference from phonological encoding between studies.

According to the model, attention is deployed to one task or another differentially depending on the specific demands raised by the relationship between the tasks. This does not mean that the relation between word planning, on the one hand, and attention and gaze shifting, on the other, is completely flexible. In WEAVER++, the activation flow from concepts to phonological forms is limited unless attentional enhancements are involved to boost the activation. Thus, attention should be sustained until the word has been planned far enough (Roelofs, 2003), which typically comprises some aspects of form encoding (Roelofs, 2007, 2008a, 2008b).

General Discussion

The research reported in the present article examined dual-task interference from phonological encoding in object naming. Ferreira and Pashler (2002) obtained evidence that phonological encoding in picture naming may happen simultaneously with the performance of an unrelated manual task, whereas A. S. Meyer et al. (2003) and A. S. Meyer and Van der Meulen (2000) obtained evidence from eye tracking that phonological encoding in picture naming delays another naming task. This suggests that the dual-task interference from phonological encoding may have a strategic origin. To clarify the role of gaze shifts in dual-task performance and to further examine whether phonological planning may yield dual-task interference with secondary manual responses, four PRP experiments were run. Participants named pictures in blocks of trials in which part of the names could or could not be phonologically encoded in advance, and the participants manually indicated the direction of an arrow presented 15° away from the pictures (Experiment 1) or centered on the pictures (Experiments 2 and 3), or they responded manually to tones (Experiment 4). The SOAs between picture and arrow or tone were 0, 300, and 1,000 ms. Earlier research has shown that speech production may hamper auditory perception, which predicts earlier shifts of attention from the vocal task to the manual task with tones than with arrows. Advance phonological encoding reduced picture naming and gaze shifting latencies at all SOAs. Gaze shifts were dependent on phonological encoding even when they were postponed at the nonzero SOAs. Manual responses to the arrows were delayed and reflected the preparation benefit from anticipatory phonological encoding at the short SOAs (i.e., 0 and 300 ms) but not at the long one (i.e., SOA = 1,000 ms), demonstrating dual-task interference from phonological planning. Manual responses to the tones were also delayed at the short SOAs but did not reflect the preparation benefit from anticipatory phonological encoding at any SOA. These results suggest that phonological planning may yield dual-task interference depending on the experimental situation, suggesting that the interference has a strategic origin.

The results from Experiments 1–3 indicate that dual-task interference from phonological planning may be obtained even when there is no need for the avoidance of vocal response buffering (Levelt & Meyer, 2000). Moreover, the results show that gaze shifts may happen after response selection, indicating that the claim by Sanders and colleagues (Sanders, 1998; Sanders & Lamers, 2002; Van Duren & Sanders, 1995) that gaze shifts occur before response selection is not universally true. Importantly, the results of Experiments 1–4 indicate that phonological planning may yield dual-task interference in some task situations (Experiments 1–3) but not in other situations (Experiment 4). These findings are more in line with an account of the dual-task interference in terms of attentional control (e.g., Allport, 1989; D. E. Meyer & Kieras, 1997a, 1997b) than in terms of a structural central response-selection bottleneck in speaking (Ferreira & Pashler, 2002; Pashler, 1994). If phonological encoding draws on a central response-selection mechanism, as the data of the present Experiments 1–3 suggest following the logic of Ferreira and Pashler (2002), phonological encoding should do so in all word production tasks requiring phonological planning. However, Ferreira and Pashler argued on the basis of their own data that phonological encoding does not draw on a central response-selection mechanism. This suggests that the notion of an immutable, structural central response-selection bottleneck does not account well for the findings on phonological planning. The reported computer simulations demonstrated that a simple model of attentional control in the coordination of vocal responding, gaze shifting, and manual responding quantitatively accounts for the present results.

It was suggested that the reason the shift criterion was set earlier for the tone task than for the arrow task is that speech planning may hamper tone processing. Consequently, the tone task may need to be protected against interference from speech planning. This might be achieved through attentional enhancement of the processing of the tone (cf. Benedict et al., 2002; Crottaz-Herbette & Menon, 2006). Interestingly, one of the brain areas implicated in auditory attention, the anterior cingulate cortex (ACC) on the medial surface of the human brain, is also involved in high-order control of ocular, vocal, and manual responses (e.g., Paus, 2001; Pickard & Strick, 1996; Roelofs, 2008a; Roelofs, Van Turennout, & Coles, 2006; Turken & Swick, 1999), although the involvement of part of the ACC is independent of response modality (Barch et al., 2001). Also, the ACC is involved in the suppression of activity in auditory cortex during vocalization by monkeys (Müller-Preuss et al., 1980; Müller-Preuss & Ploog, 1981). Extensive connectivity between the ACC and lateral prefrontal cortex (implicated in working memory) allows the ACC to access information about current goals and task demands (Paus, 2001). Thus, it would seem that the ACC is in a good position to mediate the kind of attentional coordination of vocal responding, auditory processing, gaze shifting, and manual responding investigated in the present study. This may be further examined in future functional brain imaging studies.

Whereas a central response-selection bottleneck model cannot easily account for the finding that phonological encoding in verbal action may cause dual-task interference in some circumstances but not in others, perhaps graded central-capacity sharing models can (e.g., Kahneman, 1973; Sperling & Dosher, 1986; Tombu & Jolicoeur, 2003). According to a graded central-capacity sharing model, central capacity is limited and allocated in a graded fashion to the Task 1 and Task 2 processes. Capacity may concern a single mental resource (e.g., Kahneman, 1973) or multiple resources (e.g., Sperling & Dosher, 1986; Tombu & Jolicoeur, 2003). The proportion of the total capacity that is allocated to Tasks 1 and 2 is strategically determined. Dual-task interference arises because capacity is shared between tasks at short SOAs. Because a graded central-capacity sharing model assumes that the allocation of central capacity is strategically determined, it may account for the observation that phonological encoding may cause dual-task interference in some circumstances but not in others. For example, to account for the dual-task interference in the present Experiments 1–3, the model may assume that most of the central capacity is allocated to Task 1 word-planning processes, including phonological encoding, and therefore not much capacity is available for the Task 2 processes. The absence of dual-task interference from phonological encoding in the experiments of Ferreira and Pashler (2002) and in Experiment 4 would then suggest that the participants in these experiments did not allocate much capacity to phonological encoding, thereby leaving most capacity for Task 2 processes.

A critical feature of a graded central-capacity sharing model is that it can explain why Task 1 response times sometimes increase as SOA decreases (Tombu & Jolicoeur, 2003). For example, if 90% of the central capacity is allocated to Task 1 and the remainder to Task 2 when both tasks require central processes, this will increase Task 1 response latencies at short SOAs compared to long ones (when 100% of the capacity may be allocated to Task 1). If the participants of Ferreira and Pashler (2002) allocated little capacity to phonological encoding and most capacity to Task 2 processes (explaining the absence of dual-task interference), picture-naming latencies should increase as SOA decreases. However, Ferreira and Pashler obtained no effect at all of SOA on the Task 1 response latencies, contrary to this prediction. Moreover, dual-task interference from phonological encoding was absent in the present Experiment 4 (suggesting that the participants allocated little capacity to phonological encoding and most capacity to Task 2 processes); but vocal latencies did not really differ between zero SOA and SOA = 300 ms. If anything, they were numerically longest at SOA = 300 ms. This is contrary to the prediction of a graded central-capacity sharing model that Task 1 response times should increase as SOA decreases (Tombu & Jolicoeur, 2003). Furthermore, if the participants in the present Experiments 1–3 allocated most capacity to phonological encoding and little capacity to Task 2 processes (explaining the dual-task interference from phonological encoding), picture-naming latencies should not vary with SOA. However, in the present Experiments 1–3, picture-naming latencies were longer at the short SOAs than at the long one, contrary to this prediction. To conclude, graded central-capacity sharing cannot easily account for the results of Ferreira and Pashler and the present findings.

In the experiments of Ferreira and Pashler (2002), the semantic effect of distractors in picture naming propagated into the manual responses to the tones even at the SOA of 900 ms, although picture-naming latencies were, on average, shorter than 900 ms. According to Ferreira and Pashler, such late propagation effects might be due to postresponse monitoring processes that are subject to the central bottleneck (e.g., Welford, 1952). Cook and Meyer (2006) presented an account of dual-task interference from spoken word planning in terms of self-monitoring, an important attentional control function. According to this view, dual-task interference from picture naming consists of some mixture of speed-up from phonological facilitation and slow-down from self-monitoring. In case of Ferreira and Pashler's experiments and the present Experiment 4, it may be that these two traded off about equally in some complex way, yielding no observed effect. To trade off equally, self-monitoring should be more difficult in the homogeneous than heterogeneous blocks of trials. In the present Experiments 1–3, there might have been little self-monitoring, causing the phonological facilitation to propagate into the Task 2 responses.

A problem with the self-monitoring account is that there is no independent evidence that self-monitoring is more difficult in homogeneous than heterogeneous contexts. Moreover, although it seems plausible to assume that tone discrimination makes self-monitoring more difficult, real evidence for this assumption is lacking. A final problem is that the self-monitoring mechanism is not specified, which makes the explanation less explicit than the account presented in Figure 7. More research is necessary to further develop the self-monitoring explanation and to determine whether it is correct. It is important to note, however, that the self-monitoring account attributes the dual-task interference from word planning to attentional control factors rather than a structural response-selection bottleneck, which agrees with the main conclusion drawn from Experiments 1–4.

To summarize, the present Experiments 1–3 show that phonological encoding in picture naming may delay performance of an unrelated, manual task. The dual-task interference from phonological planning was obtained using a vocal–manual task combination. Thus, the interference may be obtained even when there is no need to avoid response buffering. The shift of gaze between task stimuli depended on phonological encoding in word planning, indicating that gaze shifts in a PRP experiment may happen after response selection. Moreover, gaze shifts were dependent on the completion of phonological encoding even when they were temporarily postponed at the nonzero SOAs. The results of Experiment 4 show that whether dual-task interference from phonological encoding is obtained depends on the nature of the unrelated, manual task. Computer simulations showed that a simple extension of WEAVER++ (Roelofs, 1992, 1997, 2003) with assumptions about attentional control in the coordination of vocal responding, gaze shifting, and manual responding quantitatively accounts for the latency results. It was argued that central response-selection bottleneck and capacity-sharing models cannot easily account for the findings.

References

Allport, D. A. (1980). Attention and performance. In G. Claxton (Ed.), *Cognitive psychology: New directions* (pp. 112–153). London: Routledge & Kegan Paul.

Allport, D. A. (1989). Visual attention. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 631–682). Cambridge, MA: MIT Press.

Barch, D. M., Braver, T. S., Akbudak, E., Conturo, T., Ollinger, J., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: Effects of response modality and processing domain. *Cerebral Cortex, 11,* 837–848.

Benedict, R. H. B., Shucard, D. W., Santa Maria, M. P., Shucard, J. L., Abara, J. P., Coad, M. L., et al. (2002). Covert auditory attention generates activation in the rostral/dorsal anterior cingulate cortex. *Journal of Cognitive Neuroscience, 14,* 637–645.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12,* 335–359.

Cook, A. E. C., & Meyer, A. S. (2006). *Capacity demands of phonological encoding: A dual-task study.* Paper presented at the Architectures and Mechanisms of Language Processing conference, Nijmegen, the Netherlands.

Crottaz-Herbette, S., & Menon, V. (2006). Where and when the anterior cingulate cortex modulates attentional response: Combined fMRI and ERP evidence. *Journal of Cognitive Neuroscience, 18,* 766–780.

Eliades, S. J., & Wang, X. (2003). Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *Journal of Neurophysiology, 89,* 2194–2207.

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

Gottsdanker, R., Broadbent, L., & Van Sant, C. (1963). Reaction time to single and to first signals. *Journal of Experimental Psychology, 66,* 163–167.

Harley, T. A. (1984). A critique of top-down independent levels models of speech production: Evidence from nonplan-internal speech errors. *Cognitive Science, 8,* 191–219.

Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: An MEG study. *Journal of Cognitive Neuroscience, 14,* 1125–1138.

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92,* 101–144.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice Hall.

Korvorst, M., Roelofs, A., & Levelt, W. J. M. (2006). Incrementality in naming and reading complex numerals: Evidence from eyetracking. *Quarterly Journal of Experimental Psychology, 59,* 296–311.

Kubose, T. 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.

Levelt, W. J. M., & Meyer, A. S. (2000). Word for word: Multiple lexical access in speech production. *European Journal of Cognitive Psychology, 12,* 433–452.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–38.

Levy, J., Pashler, H., & Boer, E. (2006). Central interference in driving: Is there any stopping the psychological refractory period? *Psychological Science, 17,* 228–235.

Logan, G. D. (2004). Cumulative progress in formal theories of attention. *Annual Review of Psychology, 55,* 207–234.

Logan, G. D., Schachar, R. J., & Tannock, R. (2000). Executive control problems in childhood pathology: Stop signal studies of attention deficit hyperactivity disorder. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 653–677). Cambridge, MA: MIT Press.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization.* New York: Oxford University Press.

Lupker, S. J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 570–590.

Massaro, D. W. (1970). Preperceptual auditory images. *Journal of Experimental Psychology, 85,* 411–417.

Massaro, D. W. (1972). Preperceptual images, processing time, and perceptual units in auditory perception. *Psychological Review, 79,* 124–145.

Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory & Language, 29,* 524–545.

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.

Meyer, A. S., & Van der Meulen, F. F. (2000). Phonological priming of picture viewing and picture naming. *Psychonomic Bulletin & 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.

Monsell, S. (1996). Control of mental processes. In V. Bruce (Ed.), *Unresolved mysteries of the mind: Tutorial essays in cognition* (pp. 93–148). Hove, England: Erlbaum.

Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28,* 555–563.

Müller-Preuss, P., Newman, J. D., & Jürgens, U. (1980). Anatomical and physiological evidence for a relationship between the ‘cingular’ vocalization area and the auditory cortex in the squirrel monkey. *Brain Research, 202,* 307–315.

Müller-Preuss, P., & Ploog, D. (1981). Inhibition of auditory cortical neurons during phonation. *Brain Research, 215,* 61–76.

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. (1979). On the economy of the human processing system. *Psychological Review, 86,* 214–255.

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

Pashler, H., Carrier, M., & Hoffman, J. (1993). Saccadic eye movements and dual-task interference. *The Quarterly Journal of Experimental Psychology, 46A,* 51–82.

Paus, T. (2001). Primate anterior cingulate cortex: Where motor control, drive and cognition interface. *Nature Reviews Neuroscience, 2,* 417–424.

Pickard, N., & Strick, P. L. (1996). Motor areas of the medial wall: A review of their location and functional activation. *Cerebral Cortex, 6,* 342–353.

Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with “The language-as-fixed-effect fallacy”: Common misconceptions and alternative solutions. *Journal of Memory and Language, 41,* 416–426.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86,* 446–461.

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, 55(A),* 465–483.

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review, 110,* 88–125.

Roelofs, A. (2004). Seriality of phonological encoding in naming objects and reading their names. *Memory & Cognition, 32,* 212–222.

Roelofs, A. (2005). From Popper to Lakatos: A case for cumulative computational modeling. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 313–330). Hillsdale, NJ: Erlbaum.

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. (2008a). Attention to spoken word planning: Chronometric and neuroimaging evidence. *Language and Linguistics Compass, 2,* 389–405.

Roelofs, A. (2008b). Tracing attention and the activation flow in spoken word planning using eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 353–368.

Roelofs, A., Özdemir, R., & Levelt, W. J. M. (2007). Influences of spoken word planning on speech recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 900–913.

Roelofs, A., Van Turennout, M., & Coles, M. G. H. (2006). Anterior cingulate cortex activity can be independent of response conflict in Stroop-like tasks. *Proceedings of the National Academy of Sciences USA, 103,* 13884–13889.

Sanders, A. F. (1998). *Elements of human performance: Reaction processes and attention in human skill.* London: Erlbaum.

Sanders, A. F., & Lamers, J. M. (2002). The Eriksen flanker effect revisited. *Acta Psychologica, 109,* 41–56.

Schweickert, R. (1980). Critical-path scheduling of mental processes in a dual task. *Science, 209,* 704–706.

Sperling, G. A., & Dosher, B. A. (1986). Strategy and optimization in human information processing. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 1. Sensory processes and perception* (pp. 2.1–2.65). New York: Wiley.

Tanner, W. P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review, 61,* 401–409.

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.

Treisman, M., & Williams, T. C. (1984). A theory of criterion setting with an application to sequential dependencies. *Psychological Review, 91,* 68–111.

Turken, A. U., & Swick, D. (1999). Response selection in the human anterior cingulate cortex. *Nature Neuroscience, 10,* 920–924.

Van Duren, L. L., & Sanders, A. F. (1995). Signal processing during and across saccades. *Acta Psychologica, 89,* 121–147.

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.

Woodworth, R. S. (1938). *Experimental Psychology.* New York: Holt.

Woodworth, R. S., & Schlosberg, H. (1954). *Experimental Psychology.* New York: Holt.

Yantis, S. (2000). Goal-directed and stimulus-driven determinants of attentional control. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 73–103). Cambridge, MA: MIT Press.

Zelinsky, G. J., & Murphy, G. L. (2000). Synchronizing visual and language processing: An effect of object name length on eye movements. *Psychological Science, 11,* 125–131.

Appendix A

Table A1
Materials of the Experiments

Context	Vocal Response Set
Homogeneous	Set 1: baby, bezem, beker [baby, broom, beaker]
	Set 2: wapen, waaier, water [weapon, fan, water]
	Set 3: leraar, lepel, lelie [teacher, spoon, lily]
Heterogeneous	Set 4: baby, wapen, leraar
	Set 5: bezem, waaier, lepel
	Set 6: beker, water, lelie

Note. English translations of the Dutch picture names are given in brackets.

Received April 27, 2007
Revision received March 14, 2008
Accepted March 19, 2008 ■