Published in final edited form as:

Q J Exp Psychol (Hove). 2008 November ; 61(11): 1687–1709. doi:10.1080/17470210701647422.

Are phonological influences on lexical (mis)selection the result of a monitoring bias?

Els Severens,
Ghent University

Elie Ratinckx,
Ghent University

Victor S. Ferreira, and
University of California, San Diego

Robert J. Hartsuiker
Ghent University

Abstract

A monitoring bias account is often used to explain speech error patterns that seem to be the result of an interactive language production system, like phonological influences on lexical selection errors. A biased monitor is suggested to detect and covertly correct certain errors more often than others. For instance, this account predicts that errors which are phonologically similar to intended words are harder to detect than ones that are phonologically dissimilar. To test this, we tried to elicit phonological errors under the same conditions that show other kinds of lexical selection errors. In five experiments, we presented participants with high cloze probability sentence fragments followed by a picture that was either semantically related, a homophone of a semantically related word, or phonologically related to the (implicit) last word of the sentence. All experiments elicited semantic completions or homophones of semantic completions, but none elicited phonological completions. This finding is hard to reconcile with a monitoring bias account and is better explained with an interactive production system. Additionally, this finding constrains the amount of bottom-up information flow in interactive models.

Though people are extremely good at selecting the words that convey what they want to say, speech errors sometimes occur. As in many scientific domains (from particle physics to psychopathology), such breakdowns in a system can reveal critical properties of the normally functioning system. Competing theories of normally functioning word production differ in whether they assume the production system is interactive (e.g., Dell, 1986) or discrete (e.g., Levelt, Roelofs, & Meyer, 1999). Both classes of theories can explain certain error patterns (e.g., the lexical bias effect, see below), but to do so, discrete theories often need to assume an additional, production-external mechanism, namely a self-monitoring system that covertly “edits out” some errors more than others (see Hartsuiker, 2006, for a critical review of these accounts). The aim of this article is to arbitrate between these classes of theories with respect to one speech error finding, namely the phonological influence on lexical selection errors (Ferreira & Griffin, 2003).

Correspondence to: Els Severens.

address: Els Severens, Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium, Tel. 32-9-264 64 32, Fax. 32-9-264 64 96, e-mail: els.severens@ugent.be.

Many theories of language production agree that when speakers want to say something, they first need to activate semantic representations at the conceptual level. Then they have to select lexical-syntactic representations, often named lemmas (Kempen & Huijbers, 1983), which contain syntactic information like a noun's grammatical gender (but note that some theories do not make the lemma/word form distinction, e.g., Caramazza, 1997). The lemma in turn activates the word form. Word form encoding contains two separate processes, one which retrieves the phonological segments and one which retrieves the structure of the word. Discrete models assume that at each production level, a representation must be selected before the selection of the next representation is started at the next level. Furthermore, discrete models forbid feedback from lower levels to higher levels (Levelt et al., 1999). In contrast, interactive models assume that information feeds forward or *cascades* even before an item is selected on a specific level, and there is also feedback from lower levels to earlier levels (Dell et al., 1997). The assumption of discrete models that each step has to be completed before the next step begins is challenged by a range of data (see Vigliocco & Hartsuiker, 2002, for review). However, there is relatively little evidence that argues against the assumption of feedforward-processing only, and most of this evidence comes from speech error patterns (Baars, Motley, & MacKay, 1975; Dell & Reich, 1981; Hartsuiker, Corley, & Martensen, 2005; Nooteboom, 2005; Rapp & Goldrick, 2000; Stemberger, 2004; Vitevitch, 2002). As mentioned above, discrete models can account for such speech-error patterns when invoking a self-monitoring system. To illustrate both kinds of explanations, we will next discuss the *lexical bias effect*.

The lexical bias effect is the finding that phonological errors are more often real words than nonwords than is expected by chance (e.g., Baars, et al., 1975; Dell & Reich, 1981; Hamm, Junglas, & Bredenkamp, 2004; Hartsuiker, et al., 2005; Hartsuiker, Anton-Mendez, Roelstraete, & Costa, 2006; Nooteboom, 2005). Interactive models can explain this finding by using feedback. Suppose a phonological segment becomes wrongly activated (e.g., because of priming or because of noise in the system). In conjunction with the other segments of the word, it will then send feedback to the lexical level. This means that segments belonging to existing words, and in particular for phonological neighbors of the target word, become highly active, so they are relatively likely to be produced. Conversely, because nonwords do not have lexical representations, feedback will never converge upon a nonword representation, so that segments belonging to nonwords are less likely to be produced (e.g., Dell, 1986).

In contrast, discrete models explain the lexical bias effect by assuming an internal self-monitor that inspects internal speech for errors and that attempts to filter these errors out before they are pronounced. One criterion that the self-monitor can use is lexicality. Because nonword errors fail this criterion, they have a higher chance of being detected and hence of being covertly replaced with the correct word. Word errors, on the other hand, pass the lexicality criterion, and can thus more easily slip through. As a result, unrepresentative numbers of word errors will be overtly produced, even though there were chance-level numbers of each error prior to monitoring (Baars et al., 1975; Levelt et al., 1999; Roelofs, 2004; Nooteboom, 2005).

Similar "monitoring bias" accounts have been presented for other speech error patterns that appear to reveal interactivity. For example, the mixed error bias (the tendency for speech errors to be related in both form and meaning to intended words), can be explained by feedback from the word form to the lemma level (e.g., Rapp & Goldrick, 2000). But the mixed error bias could also be explained by a monitoring bias account, under the assumption that it is more difficult to detect a semantic error when it is phonologically similar to the target word than when it is phonologically different from the target word (Levelt et al., 1999; Roelofs, 2004). The purpose of the present study is to test whether a monitoring bias

account or an interactive account holds for another phonological effect on lexical selection, namely the effect that the partial activation of a particular word (e.g., *none*) triggers lexical selection errors when the target word (*priest*) is semantically related to a homophone of the primed word (*nun*; Ferreira & Griffin, 2003).

In particular, Ferreira and Griffin (2003) presented speakers with highly constraining sentence fragments, followed by a picture. Participants read the sentence fragments and then named each picture. The dependent variable was the frequency with which participants produced the sentence completion instead of the correct picture name. The sentence fragments were constructed so that on critical trials, the most obvious completion was semantically related to the picture, a homophone of the semantically related word, or an unrelated control word. For example, given a picture of a *priest*, a semantic competitor sentence was: *The woman went to the convent to become a...* and a homophone competitor sentence was: *I thought that there would still be some cookies left, but there were....* Speakers produced more sentence completions (i.e., *nun* or *none*) in sentences that primed the semantically related word or the homophone of the semantically related word compared to unrelated sentences.

Under the assumption that the word forms of the homophone are shared (Biedermann, Blanken, & Nickels, 2002; Cutting & Ferreira, 1999; Jescheniak, Meyer, & Levelt, 2003; Note, however, that there is also evidence for separate word forms for both meanings of the homophone Caramazza, Costa, Miozzo & Bi, 2001), interactive models explain the results found in the homophone condition by claiming that the competitor sentence activates the concept and the lemma of the homophone [*none*], and then the shared *nun/none* word form. This shared word form then feeds activation back to the lemma of the other meaning of the homophone [*nun*]. This other meaning of the homophone also receives some activation from the concept corresponding to the picture [*priest*], due to the semantic relation between these two (Figure 1). Because feedback from the word form and the activation from the picture converge on the lemma of the sentence completion (*nun*), it will be mistakenly selected more often than the lemma for an unrelated completion.

Ferreira and Griffin's data can also be explained by a discrete model if one assumes that it is harder for the monitor to detect that the homophone was wrongly selected compared to unrelated words. Assume that in the homophone condition, the production process selects the phonological representation of [*none*]. Before "none" (or "nun") is articulated, the monitor will attempt to determine if it was erroneously selected. Critically, the monitor may have difficulty detecting the error, essentially because the second meaning of the homophone [*nun*] and the picture [*priest*] are semantically related. One possibility is that a comprehension-based monitor may have access only to the phonology (and not the semantics) of the originally selected form [*nun/none*]. If so, then the monitor may determine that [*nun*] had been selected, which because of its semantic similarity to the target [*priest*] (Hartsuiker, Pickering, & De Jong, 2005), may be permitted by the monitor. In fact, the monitor's determination that [*nun*] had been selected may be promoted because [*nun*] will have additional activation at a conceptual and lemma level, due to activation spreading from [*priest*]. Furthermore, when comprehension mechanisms process "none," exhaustive access (Swinney, 1979; Tanenhaus, Leiman, & Seidenberg, 1979) may cause the "nun" form to become even more activated. All of these influences make it harder for the monitor to detect the error, under the assumption that the self-monitor checks whether the word recognized in comprehension matches an activated concept (Roelofs, 2004). When both words are unrelated, there is no match at the semantic level (as a result of the comprehension process or otherwise) and thus, more unrelated words are filtered out before they are erroneously pronounced.

This account, however, makes a further prediction. Consider a sentence context that primes a word that is phonologically related to the target, for example *For his birthday David got a huge ... [picture: priest]*. According to a monitoring bias account, the monitor should find it more difficult to detect the error *present* compared to an unrelated word. This is because *priest* is in the comprehension cohort of *present* as a result of the shared initial two phonemes, so that as a comprehension monitor processes “present,” the lexical and conceptual representations of “priest” will become activated (Zwitserslood, 1989). Thus, according to a comprehension-monitor account, cohort-primed activation should act just as exhaustive-access-primed activation does in the homophone case. This predicts that phonologically related completions should occur more frequently than unrelated completions.

There is of course a difference between the homophone condition and the phonological condition: In the homophone condition there is semantic activation of *nun* from the picture *priest*, but in the phonological condition there is no semantic activation of *present* from the picture *priest*; there is only the cohort-primed activation. However, our phonological case is still comparable to the explanation of the mixed error bias of Roelofs (2004). According to Roelofs (2004), the mixed error bias occurs because all else equal, the substitution *calf* for *cat* is more likely than *pig* for *cat*, because when the perceptual system processes *calf*, it will activate a cohort of similar sounding words including the target *cat*. This makes it more likely that the monitor will incorrectly accept *calf* as a correct production. Although in the mixed error bias there is a semantic relation between the error and the target, the same is true for phonologically unrelated semantic errors and these are less frequent. The monitoring bias account can therefore only explain the mixed error bias on the basis of the phonological similarity in the mixed errors.

Additionally, Slevc and Ferreira (2006) provided empirical support for the hypothesis that the internal monitor inspects the phonological level to check errors during production. In a picture naming study, participants were sometimes presented with the picture’s name or with phonologically related, semantically related, or unrelated words. These words were presented visually or auditory. If no word was presented or if the word was the same as the picture’s name, participants had to continue to name the picture, but if the words were different, participants had to stop naming. Slevc and Ferreira found that stopping the production of the word was sensitive to phonological similarity, but not to semantic similarity. These results support the hypothesis that at least under some conditions, the monitor checks accuracy by assessing phonological similarity. If monitoring bias accounts are correct that the internal self-monitor is responsible for the higher number of sentence completions in the homophone condition, then one would also expect more sentence completions when the picture and the word primed by the sentence are phonologically related.

Interactive models might also expect that there are more sentence completions when there is a phonological relation, but this depends on the extent to which interaction is limited or unbounded (Goldrick, 2006). To explain phonological completions in the present paradigm, one needs to assume a relatively unbounded amount of cascading, so that the competitor sentence in our earlier example will activate the *present* concept, and activation will flow through the lemma level and then to the word form level. Meanwhile, the picture activates the concept *priest* and again it activates the corresponding lemma, word form and then its phonological segments. The first two phonological segments are the same for both words, thus through feedback the word form *present* gets some activation from the segments activated by *priest*, which adds to activation this word form may have from the cascading process. For the segments of *present* to win the competition for selection, a further cascading step is necessary, namely from word form to segments (Figure 1). Note, however,

that this requires that the word form *present* maintains its activation (which cascaded to it from the concept and via the lemma) for a relatively long time, namely until the phonological segments for *priest* have been activated. Additionally, the activation has to cascade rather far (from concept to lemma to word form to segments). Based on neuropsychological data, simulation data and chronometric data, Goldrick (2006) concluded that interaction has to be limited. Taking into account that there are limitations on interactivity (Dell & O'Seaghdha, 1991; Goldrick, 2006), it is possible that in this paradigm phonological completions do not occur more often than unrelated completions.

To test for a phonological completion effect, we carried out five experiments. The first four experiments were conducted in Dutch. These experiments used phonological competitor sentences (*present-priest*) and semantic competitor sentences (*nun-priest*). The last experiment was carried out in English with phonological competitor sentences and homophone-of-semantic-competitor sentences (*none-priest*; from now we will refer to this as simply "homophone competitor sentences"). A discrete model with internal self-monitoring predicts more sentence completions in the phonological competitor sentences compared to the unrelated competitor sentence. In both the semantic and homophone competitor conditions we expect to replicate the findings of Ferreira and Griffin (2003), namely, that both types of related competitor sentences will elicit more sentence completions than unrelated sentences do.

Experiment 1

Method

Participants—Eighteen native Belgian Dutch speakers (17 females and 1 male) participated in the experiment. All were first year psychology students at Ghent University and participated in exchange for course credit, had normal or corrected-to-normal vision, and did not have any speaking disability.

Materials—One-hundred-ninety pictures were selected from a database of normed pictures (Severens, Van Lommel, Ratincx, & Hartsuiker, 2005). Each picture had a name agreement of at least 70%. For each picture we constructed two high cloze probability sentences. The last word of these sentences was either phonologically or semantically related to the picture. The mean length of the semantically related sentences was 8.9 words (range of 3 to 18 words) and of the phonologically related sentences 8.2 words (range of 3 to 16 words). Phonologically related words had the same first two phonemes as the pictures. This degree of overlap is at least as strong as the degree of overlap in which the mixed error effect has been found (Dell & Reich, 1981; Martin, Gagnon, Schwartz, Dell, & Saffran, 1997). Importantly, this shows that there is enough phonological overlap to activate the monitor. Care was taken that the phonologically related words were not semantically related and vice versa. Furthermore, the grammatical gender of the last word of the sentence and the picture were always the same (note that Marx, 1999 and Vigliocco, Vinson, Indefrey, Levelt, & Hellwig, 2004 found grammatical gender-influences on the frequency of semantic substitution errors). All the picture names and the phonologically related and semantically related words used in this experiment are shown in Appendix A.

In a pretest, 45 participants from the same population (who did not participate in any of the main experiments reported here) were asked to complete the sentence fragments. All the sentence fragments were finished with the word we expected in at least 75% of the trials. Mean sentence completion in the semantic related sentence fragments was 89.5% and in the phonologically related sentence fragments 90.6%. There was no significant difference in sentence completion between phonological and semantic competitor sentences ($t(378) = -1.39, p = 0.166$).

To create an unrelated condition, all the related sentence fragments were paired with another picture. Care was taken that none of the words were either semantically or phonologically related to the pictures. Additionally, there were also 190 filler sentences. These sentences were highly constrained, followed by a picture which corresponded to the last word of the sentence. The pictures used in the filler sentences were never used in one of the experimental conditions.

Procedure—The participants were seated in front of a computer screen and signed an informed consent sheet. First there was a practice session in which participants were familiarized with the experimental procedure. Trials started with the 500 ms presentation of a fixation cross (“+”) followed by the first word of the sentence fragment. Sentence fragments were presented word-by-word in the middle of the computer screen in red on a black background. Each word remained on the screen for 275 ms. The final word of the sentence was always replaced with a picture. The pictures were black-and-white line drawings. The participants were instructed to attentively read the sentences and not to complete it but to name the picture. The picture disappeared when the voice-key registered a response. The intertrial interval was 1500 ms. Stimulus presentation and response recording was controlled by E-prime (Psychology Software Tools, Inc., Pittsburgh, USA; www.pstnet.com/eprime). Oral responses were recorded with a mini-disc recorder.

Design and Analyses—Two lists were constructed that consisted of 95 phonologically related and 95 semantically related trials each. Sentences were never repeated within a participant, thus the remaining 95 phonological and 95 semantic sentences were used in the unrelated condition. Each sentence fragment that occurred in the related condition in one list occurred in the unrelated condition in the other list and vice versa. Each list also contained 190 filler sentences, in which the picture corresponded to the expected sentence completion.

The dependent measure was the percentage of trials per condition in which the participants completed the sentence instead of naming the picture. We carried out two 2×2 repeated measures ANOVAs with type of relation (semantic or phonological) and relatedness (related or unrelated) with subjects (F_1) and items (F_2) as random variables. Differences were considered significant if p values were less than .05. Arcsine-transformed proportions yielded the same results as reported here.

Results

Figure 2 shows the percentages of sentence completions speakers produced in the semantically related, phonologically related, and corresponding unrelated conditions. Both the main effects and the interaction were significant (type of relation: $F_1(1,17) = 23.54$, $MSE = .85$, $p < .0001$; $F_2(1,189) = 12.26$, $MSE = 17.18$, $p < .01$; relatedness: $F_1(1,17) = 26.87$, $MSE = 1.01$, $p < .0001$; $F_2(1,189) = 15.85$, $MSE = 18.08$, $p < .0001$; type of relation $\times$ relatedness: $F_1(1,17) = 31.15$, $MSE = .71$, $p < .0001$; $F_2(1,189) = 13.76$, $MSE = 17.05$, $p < .0001$). There were significantly more sentence completions in the semantically related condition (2.4%) than in the semantically unrelated condition (0.1%), $F_1(1,17) = 32.08$, $MSE = 1.54$, $p < .0001$ and $F_2(1,189) = 15.37$, $MSE = 33.83$, $p < .0001$. There was no reliable difference between the phonologically related condition (0.2%) and the unrelated condition (0.1%), $F_1 < 1$; $F_2 = 1.0$.

Discussion

There were significantly more sentence completions in the semantically related condition than in the unrelated condition. This is a replication of the findings of Ferreira and Griffin (2003). There was no difference in percentage of sentence completions between the phonologically related and unrelated condition. This finding is at odds with the prediction

that we derived from a monitoring bias account; if there is a biased monitor there should have been more sentence completions in the phonologically related condition due to the monitoring failing to filter out errors that are phonologically similar to the target. An interactive account would also predict that there are more sentence completions in the phonologically related condition, due to feedback from the phonological segments to the word-form, but as mentioned in the introduction, if interaction is limited it might be that there is no phonological effect. However, before we conclude that there is genuinely no effect of phonological relatedness in this paradigm, it is important to note that the percentage of sentence completions found here (2.4% in the semantically related condition) were rather low (cf. Ferreira & Griffin, 2003, Experiment 1 to 3: 19.5%, 18.4%, 15.3%). Because the percentage of sentence completions was low in our benchmark condition, it is conceivable that any effect in the phonological condition was too weak to surface.

There are two potential reasons for this low percentage of sentence completions. First, in the current experiment participants saw each picture four times. Possibly, this repetition makes picture naming so easy that only few completion errors occur. Second, only one-third of the sentences were filler sentences, so that in two-third of the sentences the picture name was not a good completion of the sentence. This may have led the participants to anticipate that the picture would not complete the sentence, and therefore to pay relatively little attention to the word primed by the sentence context.

In the next experiment, we therefore removed one unrelated condition (note that Experiment 1 had separate unrelated conditions for the semantic and phonological conditions) and replaced it by 95 filler sentences. Hence, half of the sentences were followed by the correct pictures and half of the sentences by incorrect pictures. Additionally, the pictures were only repeated three times.

Experiment 2

Method

Participants—Twenty first year psychology students from Ghent University (16 females and 4 males) participated in exchange for course credit. They all had normal or corrected-to-normal vision, and did not have any speaking disability; all were native Belgian Dutch speakers.

Materials—The same materials were used as in Experiment 1, with the exception that one unrelated condition was now replaced by filler sentences. Therefore, in half of the trials the expected sentence completions were the names of the corresponding pictures.

Procedure—The same procedure was used as in Experiment 1.

Design and Analyses—Two lists were constructed that consisted of 95 phonologically and 95 semantically related sentences. Half of the remaining phonologically and semantically related sentences were presented with an unrelated picture to create the control condition, which primed an unrelated word. This resulted in 95 control sentences. The sentences that were not used in one list as phonologically or semantically related sentences were used in the other list. Sentences were never repeated within a participant. Each list also contained 285 filler sentences, which ended in the corresponding pictures.

As in the previous experiment, the dependent measure was the percentage of trials per condition in which the participants completed the sentence rather than named the picture. Two repeated measures ANOVAs with Type of relation (semantic, phonological or unrelated) as a factor were carried out, one with participants (F_1) and one with items (F_2) as

a random variable. Arcsine-transformed proportions yielded the same pattern of significance as reported here.

Results

Figure 3 shows the percentages of sentence completions speakers produced in the semantically related, phonologically related, and unrelated conditions. The main effect of type of relation was significant, $F_1(2,38) = 23.81$, $MSE = 1.41$, $p < .0001$ and $F_2(2,378) = 25.97$, $MSE = .001$, $p < .0001$. There were 2.3% of sentence completions in the semantically related condition, 0.2% in the phonologically related condition and 0.0% in the unrelated condition. The difference between the semantically related and the unrelated condition was significant $F_1(1,19) = 26.58$, $MSE = 2.02$, $p < .0001$ and $F_2(1,189) = 29.26$, $MSE = .002$, $p < .0001$. There was no significant difference between the phonologically related and the unrelated condition ($F_1(1,19) = 3.35$, $MSE = .15$; $F_2(1,189) = 3.03$, $MSE = .0002$).

Discussion

Again the percentages of sentence completion differed between the semantically related and unrelated conditions but there was no difference between the phonologically related and unrelated conditions. These results replicate (with very similar numbers) those of Experiment 1. Thus, the results do not seem to be affected by a difference in the number of repetitions of each picture (3 vs. 4) or by the proportion of trials in which the picture name was identical to the sentence completion (1/2 vs. 1/3). Note, however, that the percentage of sentence completions in the semantically related condition (2.3%) is still very low (cf. Ferreira & Griffin, 2003, Experiment 1 to 3: 19.5%, 18.4%, 15.3%). It is thus possible that we did not find a difference between the phonologically related condition and the unrelated condition because the production system did not make any errors. According to a monitoring account at least, if there are no phonologically similar errors to filter out, there should be no difference between the phonologically similar and the unrelated conditions. Note however, that the results are at odds with an unbounded interactive account.

In the next experiment, we used time-pressure to attempt to elicit more sentence completions. Several studies have found that by increasing time pressure, participants make more errors during speech production (Dell, 1986; Vitkovitch & Humphreys, 1991; Oomen & Postma, 2001). Oomen and Postma (2001) suggested that the internal self-monitor works normally under time pressure as they showed that speech errors were interrupted and corrected more quickly in a fast-speech compared to a normal-speech condition. Thus, a monitoring bias account still predicts more phonological completions compared to unrelated completions. Similarly an unbounded interactive account would also predict more phonological completions.

Experiment 3

Method

Participants—Eighteen native Belgian Dutch speakers (12 females and 6 males) participated in the experiment. All were first year psychology students at Ghent University; they received course credit for their participation. They all had normal or corrected-to-normal vision, and did not have any speaking disability.

Materials—The same material was used as in Experiment 2.

Procedure—The same procedure was used as in the previous experiments, but now there was a deadline for naming the picture. Each picture was presented for 600 ms and participants were instructed to name the picture before it disappeared. Their response was

registered for a further 900 ms, during which a blank screen was presented. When the voice-key detected a response or when the 900 ms period ended, there was a blank screen for 1500 ms.

Design and Analyses—The same analyses were carried out as in Experiment 2. Arcsine-transformed proportions yielded the same pattern of significance as reported here.

Results

Figure 4 shows the percentages of sentence completions speakers produced in the semantically related, phonologically related, and unrelated conditions. Type of relation was significant, $F_1(2,34) = 35.09$, $MSE = 11.54$, $p < .0001$ and $F_2(2,378) = 36.77$, $MSE = 116.29$, $p < .0001$. There were 14.5% of sentence completions in the semantically related condition, 6.3% of sentence completions in the phonologically related condition, and 6.3% in the unrelated condition. The difference between the semantically related and the unrelated condition was significant, $F_1(1,17) = 44.59$, $MSE = 13.53$, $p < .0001$ and $F_2(1,189) = 41.23$, $MSE = 154.43$, $p < .0001$. There was no significant difference between the phonologically related and the unrelated condition, $F_1 < 1$ and $F_2 < 1$.

Discussion

The change in procedure used in this experiment (adding time pressure) was successful in increasing the percentage of sentence completions. However, there was still no difference in sentence completion rate between the phonologically related condition and the unrelated condition. Even though a sizable number of errors were observed in the phonologically similar and unrelated conditions, the lack of a difference between these two conditions suggests the monitor does not filter out fewer errors in the phonologically similar condition. Also, as before, this finding is at odds with an unbounded interactive account. Similar to the previous experiments, there was a reliable difference between the semantically related and the unrelated condition.

Name agreement of our pictures was relatively low (70% and higher). Therefore it is possible that we did not find a phonological effect because a number of the pictures simply did not activate the right phonology. To explore this, we reconducted the analyses with the pictures which had a name agreement of 90% or higher (124 pictures). Type of relation was significant, $F_1(2,34) = 17.04$, $MSE = 15.80$, $p < .0001$ and $F_2(2,246) = 17.49$, $MSE = 106.04$, $p < .0001$. There were 13.1% of sentence completions in the semantically related condition, 6.1% in the phonologically related condition and 6.7% in the unrelated condition. The difference between the semantically related and the unrelated condition was significant, $F_1(1,17) = 20.44$, $MSE = 17.81$, $p < .0001$ and $F_2(1,123) = 18.09$, $MSE = 138.69$, $p < .0001$. There was no significant difference between the phonologically related and the unrelated condition, $F_1 < 1$ and $F_2 < 1$. These results show that even if pictures with a high name agreement were used there is still no difference between the phonologically related and the unrelated condition. This shows that the lack of a difference can not be the result of activating the wrong phonology.

In contrast to the predictions of the monitoring bias account and of an unbounded interactive account, three experiments thus far observed no difference between the numbers of phonologically related and unrelated completions in this paradigm. As a control experiment, we asked the participants to first complete the sentence and then name the picture. This way, we can ascertain that under experimental circumstances that are almost identical to the previous experiments, the final word elicited by each sentence fragment was indeed phonologically (or semantically) related to the picture names.

Experiment 4

Method

Participants—Twenty first year psychology students of Ghent University (12 females and 8 males) participated in exchange for course credit. All were native speakers of Belgian Dutch, had normal or corrected-to-normal vision, and did not have any speaking disability.

Material—The same material was used as in Experiment 3.

Procedure—The sentence presentation was the same as in the previous experiments, but now the sentences were followed by dash marks that stayed on for 1500 ms. Participants were instructed to complete the sentence. When the voice-key registered a response the dash-marks disappeared from the screen. There was a blank screen for 2000 ms followed by a fixation point (“*”) for 500 ms and then the picture was shown for 600 ms, followed by another blank screen for 900 ms. Again participants were to name the picture before it disappeared; when the voice-key registered a response, the picture or the blank screen disappeared. Intertrial interval was 1500 ms.

Design and Analyses—As dependent measures we used the percentage of correct sentence completions and the percentage of pictures which were named erroneously with the last word of the sentence (only for the correctly completed sentences). The same analyses were carried out as in Experiment 3. As in previous experiments, arcsine-transformed proportions yielded the same pattern of significance as reported here.

Results

In the phonological, semantic and unrelated competitor sentence, there were respectively 87.3%, 87.1% and 87.1% correct completions; there was no significant effect of type of relation, $F_1 < 1$ and $F_2 < 1$.

Figure 5 shows the percentage of trials in which the pictures were named with the sentence completions in the semantically related, phonologically related, and unrelated conditions. There was a significant effect of type of relation $F_1(2,38) = 31.78$, $MSE = 5.47$, $p < .0001$ and $F_2(2,378) = 43.66$, $MSE = 37.83$, $p < .0001$. In the semantically related condition there were 5.2% of sentence completions, there were 0.0% of sentence completions in the phonologically related condition and 0.1% in the unrelated condition. The semantically related condition differed significantly from the unrelated condition $F_1(1,19) = 32.28$, $MSE = 15.81$, $p < .0001$ and $F_2(1,189) = 43.35$, $MSE = 111.90$, $p < .0001$. There was no significant difference between the phonologically related and the unrelated condition ($F_1(1,19) = 2.11$, $MSE = .11$; $F_2(1,189) = 2.00$, $MSE = 1.05$).

Discussion

The completion data showed that the sentence fragments were very successful in eliciting the expected completions and that the expected completion rate was very comparable in each condition (condition means varied from 87.0% to 87.3%). This demonstrates clearly that the lack of phonological completions in the naming of the pictures cannot be attributed to the sentence contexts priming other, unrelated words, from the ones we were expecting. Additionally, it shows that phonological sentences did not prime their completions more weakly than the semantic sentences did and this thus cannot cause the lack of phonological completions.

The rate of sentence completions, instead of naming the picture, was lower compared to Experiment 3, even though there was still a deadline. It is possible that the completion rate

diminished because there was now a longer time interval between reading the sentence and naming the picture. As a result, the activation level of the sentence completion's lexical representations may have decayed, so that these representations can exert less of an influence on picture naming. Additionally, the errors (sentence completions, instead of picture naming) in this experiment are a different kind of error compared to the ones in the previous experiments: they are perseverations instead of substitutions. Participants have already finished the sentence with the same word as they erroneously used for the picture. It may be that perseverations occur less often than substitutions. In a Dutch corpus of spontaneous speech errors only 4.5% of the whole word errors were perseverations, in contrast to 59.1% anticipations and 36.4% transpositions (Nooteboom, 1969).

Importantly, the picture naming data nicely replicate the previous experiments. There were more sentence completions in naming the pictures following a semantically related competitor sentence than following an unrelated competitor sentence. Again, there was no difference between naming the picture following the phonological competitor sentence and the unrelated competitor sentence.

The previous experiments were carried out in Dutch, complicating a direct comparison with Ferreira and Griffin's (2003) study, which was conducted in English. It has been suggested that there are cross-linguistic differences in patterns of speech errors (e.g., Del Viso, Igoa, & García-Albea, 1991; Hartsuiker, 2002; but see Hartsuiker et al., 2006). It is thus a theoretical possibility that phonological completions do not occur in Dutch but do occur in English. Additionally, there were slight procedural differences between Experiments 1-4 and Ferreira and Griffin's study. Furthermore, it is possible that we did not find phonological completions because having only two phonemes similar might simply not be enough. Experiment 5 therefore attempted to elicit phonological completions in English using exactly the same procedure as Ferreira and Griffin in their third experiment. Furthermore the phonological overlap between the picture and the last word was larger than two phonemes.

Because Dutch has a relatively transparent orthography and therefore few homophones, we used a semantic related benchmark condition in the Dutch experiments. In the English experiment we used a homophone condition, instead of a semantically related condition, to once more replicate the homophone effect.

Experiment 5

Method

Participants—Forty-eight undergraduates of the University of California, San Diego participated in exchange for course credit. All reported that they were native English speakers.

Materials—The same 24 black-and-white line drawings were used as in the Ferreira and Griffin study (2003). For each drawing, there was a homophone competitor sentence (with control sentences determined by re-assignment – see below). These were also the same as in the previous study. Additionally, we constructed 24 phonological competitor sentences. These sentences ended in a word that was phonologically similar to the picture. We selected phonologically similar words both by searching dictionaries for similar sounding words, and by presenting the names of the target pictures to 24 participants and asking them to come up with up to three words that sounded similar to the picture names. Note that this was done to have more phonological overlap than in the previous experiments. Then, sentences were constructed that ended in a large set of possible phonologically related words. These sentences were presented without the last word to a different group of 24 participants who were asked to complete the fragments. The final set of sentences was determined by jointly

considering how strong the phonological similarity was to the target picture and how well the designed sentence elicited the phonologically similar form. The picture names, the phonologically related and the homophone words used in this experiment can be seen in Appendix B.

To generate control sentences, sentences that elicited similar competitors (both homophone and phonologically similar ones) were re-assigned to different critical pictures that were not similar. This strategy controls for the possibility that some sentences will strongly elicit competitor intrusions irrespective of the similarity of the competitor to the target (because the same sentences are used both in similar and in control conditions). Finally, there were 32 filler sentences, which were highly constrained sentences followed by blank lines instead of a picture.

Procedure—The participants were seated in front of a computer screen and signed an informed consent sheet. First there was a practice session in which participants were familiarized with the experimental procedure. The experiment was administered with PsyScope 1.2.5 (Cohen, MacWhinney, Flatt, & Provost, 1993). Trials started with the 500 ms presentation of a fixation cross (“+”) followed by the first word of the sentence fragment. Sentence fragments were presented word-by-word in the middle of the computer screen. Each word stayed on for 275 ms. The final word of the sentence was never shown, but on critical trials, a picture appeared instead. The participants were instructed to name the picture. The picture disappeared when the voice-key registered a response. The intertrial interval was 2000 ms. The filler sentences were followed by blank lines instead of pictures, and the participants’ task was to complete the sentence as quickly as possible. This procedure was identical to that of Ferreira and Griffin (2003; Experiment 3).

Design and Analyses—We constructed four counterbalanced lists, so that each picture occurred once with each sentence type (homophone related, homophone unrelated, phonologically related and phonologically unrelated) across the lists. Each list contained six trials of each sentence type and 32 filler sentences.

The dependent measure was the percentage of trials per condition in which the participants completed the sentence rather than named the picture. We performed two 2×2 repeated measures ANOVAs with Type of relation (homophone or phonological) and Relatedness (related or unrelated) as factors. One analysis treated participants as a random factor (F_1) and one treated items as a random factor (F_2). Differences were considered significant if p values were smaller than .05. Arcsine-transformed proportions yielded the same pattern of significance as reported here.

Results

Figure 6 shows the percentage of sentence completions. There was a main effect of type of relation ($F_1(1,47) = 21.11$, $MSE = 49.96$, $p < .0001$; $F_2(1,23) = 6.18$, $MSE = 73.19$, $p < .05$) and of relatedness ($F_1(1,47) = 21.87$, $MSE = 41.34$, $p < .0001$; $F_2(1,23) = 5.43$, $MSE = 97.09$, $p < .05$), and the interaction was significant as well ($F_1(1,47) = 12.76$, $MSE = 904.23$, $p < .01$; $F_2(1,23) = 6.44$, $MSE = 70.17$, $p < .05$). There were significantly more sentence completions in the homophone related condition (10.1%) than in the homophone unrelated condition (1.4%), $F_1(1,47) = 18.01$, $MSE = 100.41$, $p < .01$ and $F_2(1,23) = 6.59$, $MSE = 137.32$, $p < .05$. There was no difference between the phonologically related and unrelated conditions (both 1.0% sentence completions), $F_1 < 1$ and $F_2 < 1$.

Discussion

The homophone conditions replicated the results of Ferreira and Griffin (2003). Significantly more sentences were erroneously completed in the related than the unrelated homophone condition. The homophone effect has now been found four times: Experiments 1, 2, and 3 of Ferreira and Griffin (2003) and Experiment 5 of the current study. In contrast, there again was no difference between the phonologically related and unrelated condition in the number of sentence completions, showing that the previous null effects were not the result of the Dutch language, of the procedure used in these experiments, or of a smaller phonological overlap.

General Discussion

In none of the five experiments reported here were there more phonological sentence completions than unrelated sentence completions. In contrast, clear effects were observed in the semantic competitor conditions (Experiments 1-4) and in the homophone of a semantic competitor condition (Experiment 5). Together these results show that the paradigm was able to elicit sentence completions in the semantic and homophone conditions, but not in the phonological condition.

The homophone effect alone can be explained by interactive models, by using feedback from the word form. Discrete models can only explain this effect by using a comprehension-based internal self-monitor, which inspects speech for errors. If the larger number of homophone completions than unrelated completions is caused by the internal self-monitor, one would expect that there are also more completions after phonological competitor sentences. A sentence completion in this case would activate the cohort in comprehension in which there is also the picture name (Roelofs, 2004), so that it would be harder for the monitor to detect that the sentence was completed instead of the picture named. Our experiments did not find more phonological sentence completions than unrelated completions, which is hard to reconcile with the monitoring bias account. Indeed, it would be very unsatisfactory if the internal monitor has to be invoked to explain the homophone completions but is then not used in the phonological condition, especially because the internal error-monitor has been shown to be sensitive to phonological information (Slevc & Ferreira, 2006). Furthermore, the effect of the cohort mechanism on monitoring specifically has been implicated in discrete accounts of mixed error effects; if such accounts hold, then phonologically similar intrusions should have appeared in the present task. Of course it can be argued that there are no intrusions in the phonological condition because these errors are simply not made by the production system and as a result there would be nothing to filter out by the monitor. However, Experiment 3 showed that with time-pressure there are intrusions in the phonological condition, and so by the monitoring account, a difference between the unrelated sentence completions and the phonological sentence condition should have been observed. The fact that we do not find this difference shows that the monitor does not filter fewer intrusions out in the phonological competitor condition even when there are such phonological intrusions.

An interactive language production theory might also predict that there should be more phonological completions than unrelated completions. In an interactive system the sentence will activate the concept, the lemma, the word form and the phonological segments of the last primed word. At the same time, the picture will activate its concept, lemma, word form, and phonological segments. The phonological segments of both the word and the picture will feed back to the word form of both the word and the picture. This will cause competition at the word form level and as a result we would expect errors. But it has been argued before that feedback is limited (Dell & O'Seaghdha, 1991; Goldrick, 2006) and thus it is possible that this feedback is not strong enough to cause errors.

Why then are there more sentence completions in the homophone competitor condition and not in the phonological competitor condition? Remember that in the homophone competitor condition, activation only has to flow to the word form to be able to cause competition. Similarly, to be able to cause mixed errors, information also has to flow only to the word form level. In the phonological competitor condition, activation has to go one step further to the phonological segments, and it is apparently too weak to cause any detectable competition. Furthermore, the word form of homophone alternatives is a single, fully overlapping representation. In contrast, phonologically similar words overlap only in a fraction of the segmental contents of those words. This also suggests that interactive effects should be weaker for phonologically similar words, compared to homophones. In line with this, Goldrick (2006) concludes that interaction in speech production is limited, but what the upper and lower limits on feedback are is still unresolved. Current data support the conclusion that feedback is limited; if feedback from the phonological segments to the word form was strong there would have been phonological errors in this paradigm.

The issue of whether feedback or self-monitoring is responsible for certain speech error patterns has received much attention in recent years. In their review of the production literature, Vigliocco and Hartsuiker (2002) rejected the assumption that activation only flows to the next level when processing at the current level is complete. They did so on the basis of a fairly large number of studies that have found evidence for cascading of activation. In contrast, Vigliocco and Hartsuiker could not reject the assumption of a unidirectional flow of activation, precisely because of the account of discrete models of speech error patterns that seem to argue for interactivity in terms of a monitoring bias. Note however that Vigliocco and Hartsuiker did, albeit tentatively, argue for feedback, on the basis of a number of theoretical arguments (e.g., they assume that there is continuity between development and processing; in contrast, Levelt, 1989 admits a role of feedback in the developing system, but he assumes that this has disappeared in the mature system).

Additionally, Hartsuiker (2006) considered the tenability of specific monitoring bias accounts, based on theoretical and empirical criteria. For example, if it is true that a given speech error pattern is caused by self-monitoring, one would expect it to be affected by manipulations that reduce the quality of monitoring, such as adding a secondary task. On the basis of this exercise, he concluded that a monitoring bias account cannot be excluded in the case of the lexical bias and mixed error bias.

The current study, however, is one of the first to directly empirically test the predictions of a monitoring bias account. The only other studies to directly address this are Baars et al. (1975) and Hartsuiker et al. (2005). The latter study concluded that monitoring biases can indeed interact with an underlying speech error pattern, but that the lexical bias effect results first and foremost from feedback (between phonological segments and word forms). It is thus possible that the completion errors in the homophone-of-semantic competitor conditions (Experiment 5 here; Ferreira & Griffin, 2003) are also the result of feedback, namely between word forms and lemmas. Note however, that this finding can also be explained by cascading alone. The sentence activates the concept [*none*], which in turn activates the lemma [*none*] and the shared word form [*none/nun*]. Meanwhile the picture activates the concept [*priest*], because of semantic similarity this concept also activates the lemma [*nun*], due to cascading the shared word form [*none/nun*] gets additional activation. In fact, a similar cascading mechanism could also explain the semantic errors made in Experiment 1-4. Importantly, this explanation only holds when lexical selection can take place at the word form level; if there is only lexical selection at the lemma level like Levelt et al. (1999) suggested, there has to be feedback to explain the homophone completions.

The current experiments did not elicit phonological sentence completions. However, this does not mean that it is completely impossible to elicit phonological errors in experimental settings. For instance, phonological errors have been found in the slip-task (e.g., Baars, Motley, & McKay, 1975; Motley, Camden, & Baars, 1982; Hartsuiker et al., 2005; Möller, Jansma, Rodriguez-Fornells, & Münte, 2006). In the slip-task participants have to silently read pairs of words; once in a while the pair (target pair) is followed by a cue that the last pair has to be pronounced. The preceding pairs are biasing pairs in which the first phoneme of the first word is the same as the first phoneme of the second word of the target pair and vice versa. These phonological errors can be explained by feedback of phonological segments to the word form. One important difference between the slip-task and the task that was used in the current experiments is that they elicit different kind of errors. The slip-task elicits segmental level errors and the current task elicits lexical selection errors.

Another paradigm that has been used to elicit phonological errors is the tongue twister paradigm (Wilshire, 1998, 1999; Goldrick, & Blumstein, 2006). In this task participants have to recite tongue twisters, like: *palm neck name pack*. Phonological errors in the tongue twister paradigm can be explained in the same way as the errors in the slip-task. Again this task elicits segmental level errors instead of lexical selection errors. Thus phonological errors can be elicited on a segmental level, but the current task taps into lexical selection; it thus appears that it is not possible to elicit phonological errors at this level. Note that in line with this, Rapp, Benzing, and Caramazza (1997) tested an aphasic patient who showed semantic errors; but did not show phonological errors, and they suggested that this patient had damage at a lexical level.

Both the slip-task and the tongue twister paradigm show that phonological errors can be elicited in an experimental setting. But crucially, the experiments here showed that phonological errors cannot be elicited in a setting in which homophone of semantically-related errors can be elicited - whereas a biased monitoring theory predicts that both types of errors should occur more often than unrelated errors. We therefore conclude that the homophone completions can only be explained by interactive models, instead of by discrete models which need to assume a biased internal self-monitor. Finally, the fact that we did not find phonological completions constrains the parameters controlling bottom-up spread of activation in an interactive production system.

Appendix A

Stimuli used in Experiments 1-4

Picture	Last word of Sentence	
	Phonologically Related	Semantically Related
bord (plate)	bos (wood)	glas (glass)
lamp (lamp)	ladder (ladder)	kaars (candle)
eenhoorn (unicorn)	eend (duck)	zeemeermin (mermaid)
ster (star)	stad (city)	maan (moon)
jongen (boy)	yoghurt (yogurt)	man (man)
orgel (organ)	origineel (original)	drumstel (drums)
stethoscoop (stethoscope)	stank (smell)	thermometer (thermometer)
mier (ant)	microfoon (microphone)	spin (spider)
kooi (cage)	koning (king)	stal (stall)
spuut (needle)	spier (muscle)	boor (drill)

Picture	Last word of Sentence	
	Phonologically Related	Semantically Related
slak (snail)	sleutel (key)	worm (worm)
boterham (sandwich)	boom (tree)	croissant (croissant)
worst (sausage)	wolk (cloud)	hamburger (hamburger)
hoed (hat)	hoek (corner)	pet (hat)
sigaret (cigarette)	citroen (lemon)	vredespijp (calumet)
paraplu (umbrella)	panter (panther)	regenjas (raincoat)
voet (foot)	voeding (food)	enkel (ankle)
baard (beard)	baas (boss)	snor (moustache)
schelp (shell)	schotel ([flying] saucer)	kwal (jelly fish)
trommel (drum)	traan (tear)	blokfluit (recorder)
schaar (scissors)	schoen (shoe)	lijm (glue)
vlot (raft)	vlies (membrane)	schip (ship)
verrekijker (binoculars)	vensterbank (window-sill)	bril (glasses)
springtouw (jump rope)	spel (game)	poppenhuis (doll house)
stofzuiger (vacuum)	straf (punishment)	bezem (broom)
tijger (tiger)	tijd (time)	leeuw (lion)
fontein (fountain)	formule (formula)	waterval (waterfall)
mand (basket)	matras (mattress)	doos (box)
soldaat (soldier)	sok (sock)	generaal (general)
vinger (finger)	vis (fish)	teen (toe)
hond (dog)	honger (hunger)	poes (cat)
fiets (bicycle)	file (file)	step (scooter)
masker (mask)	magazijn (storage)	kostuum (costume)
bed (bed)	beste (best)	kussen (pillow)
gieter (watering can)	gier (vulture)	tuinslang (garden hose)
schildpad (turtle)	schedel (skull)	kikker (frog)
zak (bag)	zanger (singer)	mand (basket)
bliksem (lightning)	bloei (blossoming)	donder (thunder)
duivel (devil)	duim (thumb)	engel (angel)
doos (box)	doorn (thorn)	kist (chest)
schommel (swing)	schat (treasure)	zandbak (sandpit)
radio (radio)	raaf (raven)	televisie (television)
dennenappel (pinecone)	demper (silencer)	kokosnoot (coconut)
mes (knife)	merk (brand)	bord (plate)
popcorn (popcorn)	pomp (pump)	lolly (lollipop)
beer (bear)	bedevaart (pilgrimage)	wolf (wolf)
kangoeroe (kangaroo)	kassa (cash desk)	panda (panda)
pizza (pizza)	piercing (piercing)	spaghetti (spaghetti)
roos (rose)	room (cream)	tulp (tulip)
ketting (chain)	kerk (church)	touw (rope)
kom (bowl)	koffie (coffee)	pan (pan)

Picture	Last word of Sentence	
	Phonologically Related	Semantically Related
hart (heart)	halfuur (half a hour)	kruis (cross)
sneeuwman (snowman)	snaar (string)	iglo (igloo)
tovenaar (wizard)	toren (tower)	heks (witch)
bal (ball)	bas (bass)	kegel (cone)
schilderij (painting)	schot (shot)	beeld (sculpture)
tandenborstel (toothbrush)	tabak (tobacco)	flosdraad (dental flos)
wiel (wheel)	weekend (weekend)	stuur (steering-wheel)
pleister (band aid)	plicht (duty)	zalf (cream)
zaag (saw)	zaak (business)	hamer (hammer)
brug (bridge)	brand (fire)	tunnel (tunnel)
riem (belt)	rivier (river)	broek (trousers)
jojo (yoyo)	jood (jew)	tol (spinning top)
racket (rocket)	ramp (disaster)	duikboot (submarine)
slang (snake)	slee (sledge)	kameleon (chameleon)
ananas (pineapple)	ander (other)	meloen (melon)
spaarvarken (piggybank)	spook (goast)	zakgeld (pocket money)
spiegel (mirror)	spar (spruce)	lippenstift (lipstick)
matroos (sailor)	magneet (magnet)	soldaat (soldier)
dolfijn (dolphin)	dochter (daughter)	haai (shark)
gitaar (guitar)	gynaecoloog (gynaecologist)	viool (violin)
eend (duck)	eencellige (unicellular)	zwaan (swan)
vlieg (fly)	vlam (flame)	mug (mosquito)
varken (pig)	vak (course)	schaap (sheep)
lepel (spoon)	lening (loan)	vork (fork)
auto (car)	auteur (author)	vrachtwagen (truck)
strik (bow)	staf (crosier)	knoop (button)
deur (door)	deuk (dent)	trap (stairs)
pan (pan)	partij (party)	schaal (dish)
citroen (lemon)	sigaar (cigar)	banaan (banana)
dokter (doctor)	dollar (dollar)	kapper (hairdresser)
trap (stairs)	tros (bunch)	lift (elevator)
dinosaurius (dinosaur)	dief (thief)	mammoet (mammoth)
fles (bottle)	flik (cop)	kan (can)
dak (roof)	dagboek (diary)	plafond (ceiling)
lamp (light bulb)	lasso (lasso)	schakelaar (switch)
poot (paw)	poort (gate)	snavel (beak)
slee (sled)	slak (snail)	koets (carriage)
diamant (diamond)	dierentuin (zoo)	parel (pearl)
kruiwagen (wheelbarrow)	kriek (cherry)	hark (shark)
zak (paper bag)	zalm (salmon)	tas (bag)
nest (nest)	net (net)	hol (burrow)

Picture	Last word of Sentence	
	Phonologically Related	Semantically Related
baby (baby)	beer (bear)	vader (father)
pauw (peacock)	paus (pope)	kalkoen (turkey)
regen (rain)	revolutie (revolution)	sneeuw (snow)
helikopter (helicopter)	hemel (heaven)	fiets (bike)
haak (hook)	haan (rooster)	spijker (nail)
leeuw (lion)	leeftijd (age)	tijger (tiger)
gordijn (curtains)	golf (golf)	behang (wallpaper)
been (leg)	beest (animal)	hoofd (head)
kurkentrekker (corkscrew)	kust (coast)	notenkraker (nutcracker)
ladder (ladder)	lawine (avalanche)	roltrap (escalator)
pruik (wig)	projector (projector)	baard (beard)
potlood (pencil)	pond (pound)	krijt (chalk)
stuur (steering wheel)	staal (steel)	wiel (wheel)
arm (arm)	ark (ark)	neus (nose)
kasteel (castle)	kalf (calf)	paleis (palace)
ridder (knight)	ring (ring)	prinses (princess)
man (man)	mast (mast)	vrouw (woman)
duim (thumb)	duik (dive)	vinger (finger)
schoorsteen (chimney)	schilder (painter)	haard (fireplace)
haar (hair)	hakenkruis (swastika)	oor (ear)
clown (clown)	k Luis (safe)	leeuwentemmer (lion tamer)
pomp (bicycle pump)	pols (wrist)	ketting (bicycle chain)
barbecue (barbecue)	bakker (baker)	fondue (fondue)
palmboom (palm tree)	passer (compasses)	cactus (cactus)
konijn (rabbit)	koper (buyer)	lam (lamb)
pincet (tweezers)	pil (pill)	scalpel (scalpel)
wortel (carrot)	wok (wok)	kool (cabbage)
eekhoorn (squirrel)	economie (economy)	mu is (mouse)
regenboog (rainbow)	rekening (bill)	zon (sun)
thermometer (thermometer)	tempel (shrine)	kom pas (compass)
hamer (hammer)	haven (harbour)	tang (plier)
kraan (tap)	krik (jack)	stop (plug)
vrachtwagen (truck)	vrucht (fruit)	tram (tram)
maan (moon)	maag (stomach)	ster (star)
vleugel (wing)	vlaming (Flemish man)	poot (leg)
tang (pliers)	tampon (tampon)	schroevendraaier (screwdriver)
bril (glasses)	broer (brother)	lens (lens)
haai (shark)	haak (hook)	dolfijn (dolphin)
koning (king)	cola (coke)	prins (prins)
bom (bomb)	bob (designated driver)	granaat (grenade)
stad (city)	stempel (stamp)	kolonie (colony)

Picture	Last word of Sentence	
	Phonologically Related	Semantically Related
kaas (cheese)	kano (canoe)	melk (milk)
mixer (mixer)	mist (mist)	deegrol (rolling pin)
boom (tree)	boon (bean)	struik (bush)
asbak (ashtray)	assistente (assistant)	lucifer (match)
pijl (arrow)	pijn (pain)	kogel (bullet)
kreeft (lobster)	kribbe (cradle)	schelp (shell)
schop (shovel)	schoorsteen (chimney)	emmer (bucket)
portefeuille (wallet)	postbode (postman)	handtas (purse)
cassette (tape)	kalender (calendar)	cd (cd)
wolk (cloud)	worst (sausage)	regenboog (rainbow)
kaars (candle)	kaas (cheese)	lamp (light)
peer (pear)	peter (godfather)	appel (apple)
telefoon (telephone)	tekenfilm (cartoon)	walkietalkie (walkie-talkie)
bijl (ax)	bijbel (bible)	schop (spade)
blik (can)	blad (magazine)	fles (bottle)
jas (coat)	jacht (hunt)	muts (cap)
kameel (camel)	kannibaal (cannibal)	ezel (donkey)
bank (bench)	balk (beam)	zetel (chair)
schouder (shoulder)	scheidsrechter (referee)	borst (breast)
cowboy (cowboy)	kous (stocking)	sheriff (sheriff)
postzegel (stamp)	porto (port wine)	brief (letter)
zon (sun)	zolder (attic)	melkweg (Milky way)
naaimachine (sewing machine)	nagel (nail)	strijkplank (ironing board)
eskimo (eskimo)	estafette (relay)	indiaan (Indian)
meisje (girl)	mijnenveld (mine field)	kind (child)
vlinder (butterfly)	vlek (stain)	lieveheersbeestje (ladybug)
kok (cook)	kop (head)	ober (waiter)
hert (deer)	hek (fence)	konijn (rabbit)
sleutel (key)	sloop (pillowcase)	deur (door)
rolschaats (roller-skate)	romp (torso)	ski (ski)
accordeon (accordion)	accu (battery)	orgel (organ)
piramide (pyramid)	piste (piste)	eifeltoren (Eiffel tower)
eikel (acorn)	eicel (ovule)	noot (nut)
brief (letter)	brug (bridge)	kaart (postcard)
strijkplank (ironing board)	steen (stone)	droogkast (tumble dryer)
sjaal (scarf)	chef (chef)	handschoen (mitten)
banaan (banana)	bal (ball)	kiwi (kiwi)
zeppelin (blimp)	zes (six)	luchtballon (hot-air balloon)
olifant (elephant)	olijf (olive)	neushoorn (rhinoceros)
cactus (cactus)	camera (camera)	palm (palm tree)
medaille (medal)	meter (meter)	beker (goblet)

Picture	Last word of Sentence	
	Phonologically Related	Semantically Related
krokodil (alligator)	krant (paper)	slang (snake)
hoef (hoof)	hoest (cough)	staart (tail)
tafel (table)	taart (cake)	bank (couch)
sigaar (cigar)	sirene (siren)	pijp (pipe)
tent (tent)	tekst (text)	caravan (caravan)
bus (bus)	burgemeester (mayor)	trein (train)
pelikaan (pelican)	pedofiel (paedophile)	flamingo (flamingo)
kroon (crown)	krul (curl)	helm (helmet)
verpleegster (nurse)	vulkaan (vulcano)	chirurg (surgeon)
pompoen (pumpkin)	post (post)	komkommer (cucumber)
schroevendraaier (screwdriver)	schaats (ice skate)	beitel (chisel)
trein (train)	tractor (tractor)	bus (bus)
vliegtuig (airplane)	vlees (meat)	ruimteschip (space craft)
vork (fork)	vos (fish)	lepel (spoon)
brood (bread)	bruidsmeisje (bridesmaid)	koekje (cookie)
broek (trousers)	bruidegom (groom)	kilt (kilt)

Appendix B

Stimuli used in Experiment 5

Picture	Last word of the Sentence	
	Phonologically Related	Homophone Related
moon	mood	star
envelope	invention	letter
money	monday	check
fly	flag	be
rabbit	rabbi	hair
glass	grass	bowl
toe	toast	heal
woman	wolf	mail
globe	glue	ball
leaf	leash	flour
pants	pans	genes
slide	slices	shoot
hoof	hook	pa
crayon	crater	pen
racquet	raccoon	bat
bridge	bricks	damn
clock	clot	watch

Picture	Last word of the Sentence	
	Phonologically Related	Homophone Related
swan	swallow	duck
necklace	never	ring
angel	angles	ferry
ear	eat	I
lighter	license	match
priest	peach	none
dollar	dolphin	bill

Acknowledgments

This research was supported by a grant from the Ghent University special research fund (BOF 011/098/04). This research was also supported by the Fund for Scientific Research (Flanders, Belgium), of which Elie Ratinckx is a post-doctoral research fellow and by NIH grant R01 HD051010. We thank Katie Doyle for assistance developing the materials and testing participants for the English experiment. We would also like to thank Sachiko Kinoshita and three anonymous reviewers for their helpful comments on an earlier version of this manuscript.

References

- Baars BJ, Motley JT, MacKay D. Output editing for lexical status from artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*. 1975; 14:382–391.
- Biedermann B, Blanken G, Nickels L. The representation of homophones: Evidence from remediation. *Aphasiology*. 2002; 16(10-11):1115–1136.
- Caramazza A. How many levels of processing are there in lexical access? *Cognitive Neuropsychology*. 1997; 14(1):177–208.
- Caramazza A, Costa A, Miozzo M, Bi YC. The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology-Learning Memory and Cognition*. 2001; 27(6):1430–1450.
- Cutting JC, Ferreira VS. Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1999; 25(2):318–344.
- Del Viso S, Igoa JM, Garcia-Albea JE. On the Autonomy of Phonological Encoding - Evidence from Slips of the Tongue in Spanish. *Journal of Psycholinguistic Research*. 1991; 20(3):161–185.
- Dell GS. A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review*. 1986; 93(3):283–321. [PubMed: 3749399]
- Dell GS, O'Seaghdha PG. Mediated and Convergent Lexical Priming in Language Production - a Comment. *Psychological Review*. 1991; 98(4):604–614. [PubMed: 1961775]
- Dell GS, Reich PA. Stages in Sentence Production - an Analysis of Speech Error Data. *Journal of Verbal Learning and Verbal Behavior*. 1981; 20(6):611–629.
- Dell GS, Schwartz MF, Martin N, Saffran EM, Gagnon DA. Lexical access in aphasic and nonaphasic speakers. *Psychological Review*. 1997; 104(4):801–838. [PubMed: 9337631]
- Ferreira VS, Griffin ZM. Phonological influences on lexical (mis)selection. *Psychological Science*. 2003; 14(1):86–90. [PubMed: 12564760]
- Goldrick M. Limited interaction in speech production: Chronometric, speech error, and neuropsychological evidence. *Language and Cognitive Processes*. 2006; 21(7-8):817–855.
- Goldrick M, Blumstein SE. Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes*. 2006; 21(6):649–683.
- Hamm S, Junglas K, Bredenkamp J. The central executive as a prearticulatory control device. *Zeitschrift für Psychologie*. 2004; 212(2):66–75.

- Hartsuiker RJ. The addition bias in Dutch and Spanish phonological speech errors: The role of structural context. *Language and Cognitive Processes*. 2002; 17(1):61–96.
- Hartsuiker RJ. Are speech error patterns affected by a monitoring bias? *Language and Cognitive Processes*. 2006; 21:856–891.
- Hartsuiker RJ, Anton-Mendez I, Roelstraete B, Costa A. Spoonish spanerisms: A lexical bias effect in Spanish. *Journal of Experimental Psychology-Learning Memory and Cognition*. 2006; 32(4):949–953.
- Hartsuiker RJ, Corley M, Martensen H. The lexical bias effect is modulated by context, but the standard monitoring account doesn't fly: Related reply to Baars et al. (1975). *Journal of Memory and Language*. 2005; 52(1):58–70.
- Hartsuiker RJ, Pickering MJ, De Jong NH. Semantic and phonological context effects in speech error repair. *Journal of Experimental Psychology-Learning Memory and Cognition*. 2005; 31(5):921–932.
- Kempen G, Huijbers P. The lexicalization process in sentence production and naming - indirect election of words. *Cognition*. 1983; 14(2):185–209.
- Levelt, WJM. *Speaking: From intention to articulation*. Cambridge: MA: MIT Press; 1989.
- Levelt WJM, Roelofs A, Meyer AS. A theory of lexical access in speech production. *Behavioral and Brain Sciences*. 1999; 22(1):1–75. [PubMed: 11301520]
- Martin N, Gagnon DA, Schwartz MF, Dell GS, Saffran EM. Phonological facilitation of semantic errors in normal and aphasic speakers. *Language and Cognitive Processes*. 1996; 11(3):257–282.
- Marx E. Gender processing in speech production: Evidence from German speech errors. *Journal of Psycholinguistic Research*. 1999; 28(6):601–621.
- Möller J, Jansma BM, Rodriguez-Fornells A, Münte TF. What the brain does before the tongue slips. *Cerebral Cortex*. 2006:1–6.
- Nooteboom, SG. The tongue slips into patterns. In: Sciarone, AG.; Van Essen, AJ.; Van Raad, AA., editors. *Nomen, Leyden studies in linguistics and phonetics*. The Hague: Mouton; 1969. p. 114-132.
- Nooteboom SG. Lexical bias revisited: Detecting, rejecting and repairing speech errors in inner speech. *Speech Communication*. 2005; 47(1-2):43–58.
- Oomen CCE, Postma A. Effects of time pressure on mechanisms of speech production and self-monitoring. *Journal of Psycholinguistic Research*. 2001; 30(2):163–184. [PubMed: 11385824]
- Rapp B, Benzing L, Caramazza A. The Autonomy of Lexical Orthography. *Cognitive Neuropsychology*. 1997; 14(1):71–104.
- Rapp B, Goldrick M. Discreteness and interactivity in spoken word production. *Psychological Review*. 2000; 107(3):460–499. [PubMed: 10941277]
- Roelofs A. Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: Comment on Rapp and Goldrick (2000). *Psychological Review*. 2004; 111(2):561–572. [PubMed: 15065924]
- Severens E, Van Lommel S, Ratinckx E, Hartsuiker RJ. Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*. 2005; 119(2):159–187. [PubMed: 15877979]
- Slevc LR, Ferreira VS. Halting in single word production: A test of the perceptual loop theory of speech monitoring. *Journal of Memory and Language*. 2006; 54(4):515–540. [PubMed: 17917694]
- Stemberger JP. Neighborhood effects on error rates in speech production. *Brain and Language*. 2004; 90:413–422. [PubMed: 15172557]
- Swinney DA. Lexical Access during sentence comprehension (re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*. 1979; 18:645–659.
- Tanenhaus MK, Leiman JM, Seidenberg MS. Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior*. 1979; 18:427–440.
- Vigliocco G, Hartsuiker RJ. The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin*. 2002; 128(3):442–472. [PubMed: 12002697]

- Vigliocco G, Vinson DP, Indefrey P, Levelt WJM, Hellwig F. Role of grammatical gender and semantics in German word production. *Journal of Experimental Psychology-Learning Memory and Cognition*. 2004; 30(2):483–497.
- Vitevitch MS. The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 2002; 28:735–747.
- Vitkovitch M, Humphreys GW. Perseverant Responding in Speeded Naming of Pictures - Its in the Links. *Journal of Experimental Psychology-Learning Memory and Cognition*. 1991; 17(4):664–680.
- Wilshire CE. Serial order in phonological encoding: an exploration of the ‘word onset effect’ using laboratory-induced errors. *Cognition*. 1998; 68(2):143–166. [PubMed: 9818510]
- Wilshire CE. The “tongue twister” paradigm as a technique for studying phonological encoding. *Language and Speech*. 1999; 42:57–82.
- Zwitserslood P. The Locus of the Effects of Sentential-Semantic Context in Spoken-Word Processing. *Cognition*. 1989; 32(1):25–64. [PubMed: 2752705]

Figure 1.

General model of language production based on Ferreira and Griffin (2003) and Cutting and Ferreira (1999). Representation of a homophone pair (*nun* and *none*), and for one meaning of the homophone a semantic competitor (*priest*) is shown. Furthermore, the representation of two phonologically related words (*present* and *priest*) is shown. Connections from concepts to semantically similar lemmas are represented by dashed lines.

Figure 2. Percentages of sentence completions for the semantic and phonological competitor sentences in both the related and unrelated conditions. Error bars indicate standard errors of the mean.

Figure 3. Percentages of sentence completions for the semantically related, phonologically related and unrelated competitor sentences. Error bars indicate standard errors of the mean.

Figure 4. Percentage of sentence completions for the semantically related, phonologically related, and unrelated competitor sentences. Error bars indicate standard errors of the mean.

Figure 5. Percentages of sentence-completions for the semantically related, phonologically related, and unrelated competitor sentences. Error bars indicate standard errors of the mean.

Experiment 5

Figure 6. Percentages of sentence completions for the homophone and phonological competitor sentences in both the related and unrelated condition. Error bars indicate standard errors of the mean.