The distractor frequency effect in picture-word interference: evidence for response exclusion

Elisah Dhooge and Robert J. Hartsuiker

Department of Experimental Psychology, Ghent University

Running head: distractor frequency effect

Address for correspondence:

Elisah Dhooge

Department of Experimental Psychology

Ghent University

Henri Dunantlaan 2

9000 Ghent

Belgium

e-mail: Elisah.DHooge@ugent.be

Abstract

In three experiments, subjects named pictures with low or high-frequency superimposed distractor words. In a first experiment, we replicated the finding that low-frequency words induce more interference in picture naming than high-frequency words (i.e., distractor frequency effect, Miozzo & Caramazza, 2003). According to the response exclusion hypothesis, this effect has its origin at a post-lexical stage and is related to a response buffer. The account predicts that the distractor frequency effect should only be present when a response to the word enters the response buffer. This was tested by masking the distractor (Experiment 2) and by presenting it at various time points before stimulus onset (Experiment 3). Results supported the hypothesis by showing that the effect was only present when distractors were visible, and if they were presented in close proximity to the target picture. These results have implications for the models of lexical access and for the tasks that can be used to study this process.

A key issue in word production concerns the mechanisms underlying the retrieval of lexical units from the mental lexicon: lexical access. It is generally accepted that during retrieval of an objects' name from the lexicon, a spreading activation mechanism operates at the conceptual level (e.g., Caramazza, 1997; Dell, 1986; Levelt, Roelofs, & Meyer, 1999). Due to this mechanism, the semantic representation of the target name (e.g., DOG) and that of closely related concepts (e.g., CAT) become active. The amount of activation depends on the semantic distance between the target and non-target nodes. For example, DOG will receive more activation from CAT than from FISH and vice versa, as CAT and DOG are semantically closer than FISH and DOG. It is also assumed that this activation spreads to the lexical level, hereby activating a set of lexical representations (for an alternative proposal see Bloem & LaHeij, 2003; Bloem, van den Bogaert, & LaHeij, 2004). Because of the multiple activation at the lexical level, there is a need for a decision mechanism to select the correct lexical unit for further processing. This paper contrasts two views of this decision mechanism: a view according to which competition between multiple lexical candidates delays lexical selection with a view according to which the activation of competitors is in principle irrelevant and according to which any delays result from post-lexical checking mechanisms.

A prominent view is that a competitive process underlies lexical selection (e.g., Belke, Meyer, & Damian, 2005; Bloem & La Heij, 2003; Bloem et al., 2004; Caramazza & Costa, 2000; Damian & Bowers, 2003; Damian & Martin, 1999; Damian, Vigliocco, & Levelt, 2001; Hantsch, Jescheniak, & Schriefers, 2005; Levelt et al., 1999; Roelofs, 1992, 1993, 2001, 2003; Starreveld & La Heij, 1995; Vigliocco, Lauer, Damian, & Levelt, 2002; Vigliocco, Vinson, Lewis, & Garrett, 2004). According to the 'lexical selection by competition' account, the time it takes to select the target node is affected not only by its own activation level, but also by the activation of other active nodes. These competition models predict an inverse relationship between the time to select the target node and the relative activation of the non-target nodes. The more the competitors are activated, the more time it takes to select the correct node. Computational formulations of this account are usually based on the Luce choice ratio (e.g., Roelofs, 1997). This ratio generates a probability that indicates the likelihood of selection at a given time step and is based on a division of the activation of a node by the activation of all active nodes in the lexical network. Thus, the more non-target nodes that are activated and the more highly activated these nodes, the more time needed for lexical selection.

Another view of lexical selection does not assume a competitive mechanism underlying lexical selection (e.g., Caramazza, 1997; Dell, 1986). In this non-competitive view, the ease with which the target lexical node is selected will depend solely on its own activation, regardless of the activation of other lexical nodes. These models assume either a threshold that has to be exceeded for the target node to be selected, or assume that after a certain number of time steps the node with the highest activation gets selected. Thus, even though non-target nodes are active, they do not influence target selection.

The key difference between these views is the influence of the activation levels of the non-target nodes. *Prima faecie* evidence for a competitive view comes from influences of semantically related distractor words on naming latencies in the picture-word interference (i.e., PWI) paradigm (for a review, see MacLeod, 1991). In this paradigm, a picture is presented with a superimposed word (i.e., the distractor). Subjects are asked to name the picture and ignore the distractor. The lexical selection by competition account predicts that semantic relatedness between the picture and the distractor should lead to semantic interference (longer naming latencies when the picture, e.g., of a dog, is accompanied by a semantically related word (CAT) than by an unrelated word (CUP)). This is because semantically related distractors will lead to more highly activated lexical nodes than do unrelated distractors, because of the spreading activation mechanism at the conceptual level (the conceptual representation of dog spreads activation to that of cat, which in turn spreads activation to the lexical representation CAT)[1]. With unrelated distractor-picture pairs this spreading activation will of course not be present. Therefore, if a distractor word is semantically related, its lexical activation will be higher than if it is unrelated, and so a related distractor leads to more competition and longer naming latencies. In contrast, non-competition models predict no interference because the distractor word's lexical representation would be irrelevant for the speed of lexical selection. Rather, they predict semantic facilitation, because of priming at the conceptual level (e.g., the distractor concept cat primes the target concept dog, resulting in more easy lexical selection of DOG).

Consistent with competition models, many studies have indeed found evidence for semantic interference in the picture-word paradigm (e.g., Caramazza & Costa, 2000; Damian & Martin, 1999; Lupker, 1979; Rosinski, 1977; Schriefers, Meyer, & Levelt, 1990; Starreveld & LaHeij, 1995). In these studies, distractor words were usually category coordinates of the target words. Recently, however, several studies have challenged the empirical generalization that the typical pattern is semantic interference (see Mahon, Costa, Peterson, Vargas, & Caramazza, 2007 for review and discussion). For example, Mahon et al. (2007) presented pictures with distractors that were semantically related. Importantly, semantically related items varied in the semantic distance to the target. For example, the picture of a horse could be paired with the semantically close distractor ZEBRA and the semantically farther distractor CAT. They found that naming latencies were longer for within-category far distractors than for within-category close distractors. Competition models would instead predict more interference from within-category close distractors, due to the larger priming at the conceptual level. Semantic facilitation has also been found when the grammatical class of picture names and distractor words was different. Mahon et al. found semantic facilitation when using verbs as distractors. That is, presenting the picture of a car with the semantically related verb DRIVE enabled subjects to name the picture faster than when the distractor was the verb SLEEP (but see Vigliocco, Vinson, & Siri, 2008).

Moreover, there is semantic facilitation with other kinds of non-category coordinate semantic relatedness, namely part-whole relationships, like car-engine (e.g., Costa, Alario, & Caramazza, 2005), and associative pairing of distractor and picture (e.g., cat – mouse; Abdel Rahman & Melinger, 2007; Alario, Segui, & Ferrand, 2000). Summarizing, results of the PWI task indicate that not all semantic relationships lead to interference, and that even category coordinates can lead to semantic facilitation.

Competition models are not only challenged by results from the classical PWI task. Versions of this task that varied the modality of the distractor or target have posed a difficulty as well. For example, Bloem and LaHeij (2003) showed that in a word translation task, context pictures induced semantic facilitation. Additionally, when both target and distractor are words, a semantic relationship between both does not induce longer naming latencies compared to an unrelated pair of words (Glaser & Glaser, 1989; see also, LaHeij, Happel, & Mulder, 1990). Results from visual-world eye-tracking have also been taken as evidence against the lexical selection by competition view. Huettig and Hartsuiker (2008) presented four objects in a visual display and subjects were asked to name a target picture on the basis of its visual shape or its category. Objects in the display were either semantically related or unrelated, in the condition with shape-instructions, and visually similar or dissimilar in the condition with category-instructions. There were more fixations on (semantically or visually) related objects than on unrelated objects, but this did not influence naming latencies. Thus, semantic activation that is strong enough to influence visual attention apparently does not influence naming latencies, and by extension, lexical selection.

Thus, semantic effects apparently do not allow for an easy distinction between the models. Another possibility of distinguishing the models would be to manipulate the activation of the non-target nodes irrespectively of any semantic relationship between target and non-target nodes. Competition models would predict an influence of any heightened activation of a non-target node, whereas non-competition models do not predict any effects. This can be done by examining the effects of the frequency of the distractor. If lexical nodes of high-frequency words have higher resting levels than low-frequency words (McClelland & Rumelhart, 1981), the difference in activation between the distractor node and target node will be smaller with a high than with a low-frequency distractor. In line with predictions, Klein (1964) and Fox, Shor, and Steinman (1971) indeed found more interference from high compared to low-frequency words. However, as argued by Miozzo and Caramazza (2003), these studies have serious methodological limitations. First, they argued that the low and high-frequency words were not matched for factors influencing word recognition and only a small set of items were used. Most importantly, the low-frequency words used by Klein were extremely rare, and could be seen as functionally equal to non-words. As non-words interfere less than words (Klein, 1964; Lupker 1979; Lupker & Katz, 1982; Rosinski, Golinkoff, & Kukish, 1975), this might explain the results. In contrast, after controlling for these factors, Miozzo and Caramazza (2003, see also Burt, 2002) found that a *low-frequency* word induced longer picture naming latencies. This “distractor frequency effect” thus provides strong evidence against competition models.

Although the findings we have reviewed pose a challenge to competition models, they cannot be explained easily by non-competition models either. Thus, although the empirical generalization may be that there is semantic facilitation in picture-word interference tasks, non-competition models still need to explain why there is usually semantic interference in the specific situation where category coordinate distractors are compared to unrelated distractors. Additionally, the distractor frequency effect is at odds with both competition models (which predict the opposite effect) and non-competition models (which predict no effect at all). Explaining these two effects becomes crucial for non-competition models to be viable, and this can only be done by making additional assumptions. These assumptions are made most explicit in the response exclusion account (Finkbeiner & Caramazza, 2006; Janssen, Schirm, Mahon, & Caramazza, 2009; Mahon et al., 2007; Miozzo & Caramazza, 2003). This account states that the semantic interference effect and the distractor frequency effect reflect post-lexical rather than lexical processes. The assumption is that words have a privileged relationship to the articulators. In order to produce the picture name, the articulators first need to be disengaged from the distractor word. Put differently, the language production system is characterized by an output buffer. This output buffer forms a bottleneck, as it can only hold one phonologically well formed response at a time. A control process operating over the response buffer first needs to eliminate the response to the distractor word before the response to the picture can enter the buffer and be pronounced. This also explains why in the context of a picture target only word distractors but not picture distractors lead to semantic interference effects. It also explains why no effects are found with word-word compounds as stimuli. The target word will enter the buffer first, and there is no need to remove it.

According to this account, two factors determine how fast a picture will be named in the context of a distractor. The first factor is how fast the response to the distractor can enter the buffer. All things equal, if a response can enter the buffer sooner, it can also be removed sooner and so picture naming can be initiated earlier as well. This factor explains the distractor frequency effect. Because of their lower resting levels (McClelland & Rumelhart, 1981), low-frequency words will enter the buffer later than high-frequency words. Consequentially, low-frequency words can only be purged from the buffer later compared to high-frequency words, resulting in longer picture naming latencies. The second factor is how soon the response can be removed from the output buffer. This is related to response relevance: the more likely that the response in the buffer is correct, the more difficult it will be to remove. Semantic interference effects in the PWI paradigm are explained by means of this factor. It is assumed that early on in picture processing, basic category information about the picture is available, for example, partial information about the objects' structural components. This information can be used to discriminate between the response to the picture and the response to the word. If picture and word are related, this information will not be useful, and it will take more time to reject the response to the distractor and remove it from the buffer. Put differently, semantic relatedness is confounded with a manipulation of response relevant criteria: semantically related and unrelated distractors differ in whether and to what extent they satisfy semantic constraints imposed by the picture and the task. If these constraints are kept constant, semantic facilitation is found, because of the spreading activation at the conceptual level. In this way, the response exclusion account can elegantly explain the results from the Mahon et al. (2007) experiments. When all distractors are semantically related, they will all satisfy the semantic constraints imposed by the picture. For example, if the target picture is a horse, the distractors ZEBRA and CAT will both satisfy the semantic constraints ‘has four legs’ and ‘has fur’. Thus, both types of distractors will be removed from the buffer equally fast. Differences are then due to differences in priming at the conceptual level. Semantically close distractors will lead to more priming and hence to more facilitation. The same line of reasoning can be used to account for the effect of verb distractors. When all pictures are objects, one of the semantic constraints imposed by the picture is ‘name an object’. Both related and unrelated distractors will not satisfy that constraint and can be removed from the output buffer fairly fast. Again, differences will only be due to conceptual priming. Summarizing, the response exclusion account views semantic interference and the distractor frequency effect as depending on a task-specific additional mechanism and not as reflecting the dynamics of lexical selection.

Support for the post-lexical origin of the semantic interference effect in the PWI task stems from a study by Janssen et al. (2008). In this experiment, pictures were presented with semantically related or unrelated distractors. Subjects named pictures in two conditions. One condition formed the standard PWI task where picture and distractor were presented simultaneously (immediate naming). In the second condition, pictures were presented 1000 ms before the distractor. Subjects were told to prepare their response but to name the picture only when the distractor appeared (delayed naming). They reasoned that in the delayed naming condition, lexical selection would already be finished when the distractor word appeared. Any effects of semantic relatedness should thus be due to post-lexical processes, that is, to the role of the response buffer. In the immediate naming condition, picture processing would not yet have finished, so effects can be due to lexical and post-lexical processes. According to the response exclusion account, semantic interference should be found in both conditions. Competition models, which predict a lexical origin of the semantic interference effect, predict semantic interference solely in the immediate naming condition. In support of the response exclusion account there was semantic interference in both conditions.

Further evidence for the response exclusion hypothesis stems from a study by Finkbeiner and Caramazza (2006). According to the response exclusion hypothesis, no effects of the distractor should be found when no response to the distractor is formed or when it cannot enter the buffer. To test this, the authors presented pictures with semantically related and unrelated distractors in two conditions. In one condition the distractor was masked, in the other it was visible. Further research has shown that masking a distractor will prevent the formulation of a phonologically well formed response (Finkbeiner & Caramazza, 2008). The response exclusion account predicts semantic facilitation under masked conditions as this condition reflects lexical and conceptual processing. Under visible conditions, post- lexical mechanisms will give rise to semantic interference. Competition models, however, predict that semantic interference will be found in both conditions, as both reflect lexical level processing. In line with the response exclusion account, semantic interference was only found in the visible condition. In the masked condition, where presumably the formulation of a response to the distractor was prevented, semantic facilitation occurred. These two experiments therefore provide evidence for a response related origin of semantic interference effects and show that only semantic facilitation reflects lexical processing.

Summarizing, the response exclusion account assumes a non-competitive lexical selection process. It explains the semantic interference effect and the distractor frequency effect by relating it to a response buffer. It further assumes that the effects found in the standard PWI paradigm are not necessarily indicative of lexical processes, but also reflect operations concerning a post-lexical buffer. The response exclusion account of the semantic interference effect is supported by two studies, but to the best of our knowledge there is no empirical support for this account's explanation of the distractor frequency effect.

The present paper therefore tested a response exclusion account of the distractor frequency effect. One finding that appears to argue against such an account was reported by Miozzo and Caramazza (2003). These authors manipulated the semantic relatedness between picture and distractor and the frequency of the distractor. Although both main effects were significant, the interaction was not. The authors therefore argued that the two effects have a different locus. However, the absence of an interaction is predicted by the response exclusion account. Even though according to the account the semantic interference effect and the distractor frequency effect are both related to the output buffer, the effects are attributed to two different stages: the speed by which a response is removed from the buffer versus the speed by which it enters the buffer. The difference in speed by which the response enters the buffer, will not affect the decision process that follows. It will still be more difficult to reject a semantically related distractor than it will be to reject a semantically unrelated distractor. Thus, the two effects have a related but not equal origin, explaining the absence of an interaction (cf., Finkbeiner & Caramazza, 2006; Finkbeiner & Caramazza, 2006b; Janssen et al., 2008; LaHeij, Kuipers & Starreveld, 2006; Mahon et al., 2007)

To test the response exclusion account of the distractor frequency effect, we manipulated two characteristics of the task. Experiment 2 employed masked presentation of the distractor word. The response exclusion account predicts that under these circumstances the distractor word will not enter the response buffer, and hence that the distractor frequency effect should disappear. Experiment 3 varied the time-course between presentation of the distractor and target picture. The response exclusion account assumes that the longer the time interval between the stimuli is, the more likely it is that the response to the distractor has been excluded, and so the smaller the distractor frequency effect is. But first we report Experiment 1, which replicates the distractor frequency effect in Dutch, the language used in all our experiments.

Experiment 1

The goal of the first experiment was replicate the distractor frequency effect in Dutch. We used the same basic design as Miozzo and Caramazza (2003), with a few exceptions. First, to examine the boundaries of the distractor frequency effect, Miozzo and Caramazza varied the frequency of the picture name. As this did not influence the effect, it was not varied in our experiments. Second, the authors also varied the proportion of low-frequency distractors. This did not influence the effect either, so in all our experiments half of the distractors were low-frequency words and the other half were high-frequency words. Third, Miozzo and Caramazza performed the task with native speakers of English, so that distractors were presented and pictures were named in English. We performed the experiment with native speakers of Dutch, with Dutch stimuli. This way, the validity of the effect can be established by investigating whether it generalizes to other languages. Finally, in the original experiments, the distractors were not controlled for age of acquisition, which is a strong predictor of picture naming latencies (e.g., Severens, Van Lommel, Ratinckx, & Hartsuiker, 2005). As this variable is highly correlated with frequency (with early learned words having a higher frequency; for a review, see Brysbaert & Ghyselinck, 2006), it was controlled in our experiments. Despite these changes, we predicted that pictures accompanied by a low-frequency distractor would be named more slowly compared to pictures presented paired with a high-frequency distractor.

Method

Subjects. Twenty undergraduate students of Ghent University (19 women and 1 man, age range 19 - 34) participated in the experiment for partial course credits. All subjects reported normal or corrected to normal vision, and were native speakers of Dutch. They all gave written informed consent and were naïve to the purposes of the experiment. None of them participated in Experiment 2 or Experiment 3.

Design. Naming latency was considered as the dependent variable. Frequency of the distractor was the within-subject and within-item independent variable. It included two levels: low-frequency and high-frequency.

Materials. Forty-two black and white line drawings were selected from the Severens et al. (2005) database. This database contains timed picture naming norms for 590 pictures together with a number of variables known to influence picture naming latencies, such as frequency of the picture name stemming from the Celex database. Only pictures with a picture name frequency within the medium frequency range (range = 15-30 counts per million) were retained. From this set, the pictures with the highest name agreement were selected. This yielded 12 pictures used only in the practice phase (mean frequency = 22.33 counts per million) and 30 pictures used only in the experimental phase.

Words to be presented with the pictures as distractors were selected from the same database. Consequentially, only nouns were used, hereby avoiding any effects of grammatical class. Distractors and pictures used in the practice phase were never used in the experimental phase. Both the distractors used in the practice phase and the distractors used in the experimental phase were semantically and phonologically unrelated to the picture with which they were paired. For the practice phase, 12 distractors from within the medium frequency range (mean = 17 counts per million) were selected. Two distractors were selected for each picture used in the experimental phase: one with a frequency from the high-frequency range and one with a frequency in the low-frequency range. In all experiments, we selected only low-frequency distractors with a count below 15 per million, and high-frequency distractors with a count above 80 per million. Selected low-frequency distractors had a frequency count below nine per million and a log frequency below 0.95. All selected high-frequency distractors had a frequency count between 81 and 3797 per million and a log frequency range between 1.91 and 3.58. Details on the selected stimuli can be found in Table 1. The stimuli themselves can be found in Appendix 1. The difference in mean frequency was significant, t(58) = 2.36, p < .05. The same was true for the log frequency, t(58) = 20.01, p < .001. Furthermore, there were no significant differences between low- and high-frequency distractors regarding the mean number of letters, t(58) = -.20, p = .84; the mean age of acquisition, t(58) = -.29, p = .77; the mean number of syllables, t(58) = -1.49, p = .14; the mean number of phonemes, t(58) = -.55, p = .58; the number of neighbors, t(57) = -.35, p = .73; or bigram frequency, t(57) = .19, p = .85.

Insert Table 1 about here

A plus sign (‘+’) served as fixation point. All stimuli were presented centrally on a 17in. monitor with a 60Hz refresh rate placed at a distance of 60 cm in front of the subject. Distractors were presented centrally in black capital letters in a Times New Roman font 26 points. Pictures were 300 x 300 pixels large and appeared centrally on screen. For a given picture, distractors always appeared centered in the middle of the picture. Stimulus delivery and millisecond accurate response registration was achieved by means of the Tscope package (Stevens, Lammertyn, Verbruggen, & Vandierendonck, 2006) run on a Pentium 4PC. Oral responses were collected through a NEVK voice key (Duyck, Anseel, Szmalec, Mestdagh, Tavernier, & Hartsuiker, 2008) connected to the parallel port.

Procedure. The experiment was run in a sound attenuated, dimly lit room and consisted of two phases. In a first phase, subjects were familiarized with the names of the pictures that were used in the second phase. This was done in two steps. In a first step, subjects were shown a booklet containing the 42 pictures with the correct names for the pictures printed below. They were instructed to study these attentively and to memorize the name assigned to each picture. Next, all pictures were presented on a computer screen. On-screen written instructions informed the subjects to name the picture as fast and accurately as possible using the names they were given before. Each trial started with the presentation of a fixation cross for 700 ms, which was followed by the picture. After a verbal response was given, the picture disappeared and the correct name was printed on screen for 1000 ms. The next trial was initiated after 1000 ms.

The second phase followed after a self-paced break. It consisted of a practice phase and an experimental phase. In both phases, each trial started with the presentation of a fixation cross for 700 ms. This was replaced by the picture with the distractor until a verbal response was given. After the response, the experimenter typed in whether the response was correct or not and/or if any voice key failure took place. The next trial was initiated 1000 ms afterwards. The practice phase consisted of one block of 12 trials in which a picture was presented with its distractor. The experimental phase consisted of three blocks. In each block, each picture was presented once with its low- and once with its high-frequency distractor. This resulted in 60 trials per block, yielding 180 trials in total. Trials were randomized per block with the constraints that (a) a picture could not be repeated before all other pictures were named once, (b) stimuli in the same frequency condition could not appear on more than three consecutive trials and (c) per set of 10 trials there had to be five pictures paired with a low-frequency word and five with a high-frequency word. Furthermore, whether a certain picture would appear with its low or high-frequency distractor first was counterbalanced across subjects. Between the practice and the experimental phase and between blocks, subjects could take a self-paced break. Written instructions appeared on screen before both phases. Subjects were instructed to name the picture using the names they had studied before and ignore the distractor. Speed as well as accuracy was emphasized. The subject was encouraged to ask questions if anything was unclear.

In total, the experiment lasted approximately 30 minutes.

Results and discussion

Responses scored as errors included (a) voice key malfunctioning and verbal disfluencies and (b) incorrect naming of the picture. In addition, all responses exceeding the subjects mean by three standard deviations were discarded from the analyses. These criteria were also used in all subsequent experiments. F_1 analyses (on subjects' means) and F_2 analyses (on items' means) were conducted on response latencies with distractor frequency (2, low vs. high-frequency) as a within-subject and within-item variable. Errors rates are typically low in the paradigm, and were not further analyzed (cf. Miozzo & Caramazza, 2003).

Errors and outliers accounted for 2.75% and 2.33% of the data. No naming errors were made. Naming latencies were 23 ms faster with a high-frequency distractor (718 ms, SD = 68 ms) compared to a low-frequency distractor (741 ms, SD = 63 ms). This difference was significant by subjects, $F_1(1,19) = 38.43$, MSE = 5388.02, $p < .001$, and by items, $F_2(1,29) = 19.13$, MSE = 8423.45, $p < .001$.

As we expected, we replicated the distractor frequency effect (Miozzo & Caramazza, 2003). It appears to be more difficult to name a picture when it is accompanied by a low-frequency word compared to when it is paired with a high-frequency word. The results show that this effect is not language-specific and that it does not depend on age of acquisition. After having replicated the effect, the focus of the next experiments will be on its origin. In the following experiments, the hypothesis that the distractor frequency effect is related to a response buffer was tested.

Experiment 2

The second experiment focused on the origin of the distractor frequency effect in terms of the response exclusion account. If this account is correct, the distractor frequency effect should vary depending on the presence versus absence of a response to the distractor in the buffer. That is, if no response can enter the response buffer, the distractor frequency effect should be absent. But if a response is formed, the distractor frequency effect should be present. In line with Finkbeiner and Caramazza (2006), this was manipulated by masking the distractor. They argued that masking should prevent the formulation of a verbal response, so that in the PWI paradigm, no response would enter the buffer.

We note that this assumption was validated by Finkbeiner and Caramazza (2008) using the masked congruence effect. This is the effect that when responding to a target, primes that point to a different response lead to interference compared with a neutral condition. For example, when the task is to judge whether a vegetable is red or green, responses to the target ‘cucumber’ will be slower when the prime is ‘red’ then when it is a row of X’s. This interference is thought to be due to a conflict between incompatible responses at the motor level (e.g., Koechlin, Naccache, Block, & Dehaene, 1999; Kunde, Kiesel, & Hoffmann, 2003; Naccache & Dehaene, 2001) The authors argued that, if masking prevents the formulation of a verbal response, the masked congruence effect should disappear when the prime is masked and responses have to be given verbally. This is because the absence of an incompatible response should lead to the absence of a response conflict. In line with this prediction, no effect was found when primes were masked and responses had to be given verbally, whereas the effect was present with manual responses.

We thus presented low- and high-frequency distractors in two conditions: masked and visible. In line with Finkbeiner and Caramazza (2006, 2008), we expected that masking should lead to the absence of a verbal response and consequentially, the absence of the distractor frequency effect. To ensure that distractors were processed under masked conditions, we used the same time parameters as Finkbeiner and Caramazza (2006), and also included control conditions in which picture and distractor were either semantically related or unrelated. If distractors are processed, semantic facilitation should be found in the masked condition and semantic interference in the visible condition.

Method

Subjects. Forty further subjects (30 women and 4 men, age range 18-25) took part in the experiment and received €5 for their participation. All gave written informed consent and reported normal or corrected to normal vision. All subjects were native speakers of Dutch and naïve to the purpose of the experiment.

Design. Naming latency was considered as dependent variable. There were two within-subject and within-item variables: stimulus type (2; frequency vs. semantic relatedness) and visibility condition (2; masked vs. visible).

Stimuli and material. Twenty black and white drawings were selected from the Severens et al. (2005) database for the experimental phase. All had a frequency from within the medium frequency range (range = 15-30 counts per million). The same 12 pictures as in Experiment 1 were used in the practice phase. None of the pictures or distractors from the practice phase were used in the experimental phase.

For each picture, a low- and high-frequency distractor, and a semantically related and unrelated distractor (of medium frequency) was selected. Care was taken to select related distractors that were semantically close to the target. As in Experiment 1, only nouns were used. All distractors were phonologically unrelated to the target picture name; distractors from the frequency condition were also semantically unrelated to the picture name. Selected low-frequency distractors had a count below 11 counts per million. High-frequency distractors had a count between 88 and 900 counts per million. Details on the selected stimuli can be found in Table 2, the stimuli themselves can be found in Appendix 2. Unlike in Experiment 1, distractors were now matched on number of letters per picture. For the frequency condition, there was a significant difference in frequency, $t(38) = -4.81$, $p < .001$; and log frequency, $t(38) = -16.77$, $p < .001$. There were no significant differences regarding age of acquisition, $t(38) = .65$, $p = .52$; number of syllables, $t(38) = .97$, $p = .34$; number of phonemes, $t(38) = .18$, $p = .86$; number of neighbors, $t(38) = .09$, $p = .93$; or bigram frequency, $t(38) = .44$, $p = .66$. Regarding the semantic condition, matching on age of acquisition was impossible due to lack of norm data. However, there were no significant differences in log frequency, $t(38) = .03$, $p = .98$; frequency, $t(38) = .04$, $p = .97$; number of syllables, $t(38) = -.21$, $p = .83$; number of phonemes, $t(38) = -.10$, $p = .91$; number of neighbors, $t(38) = .24$, $p = .81$ or bigram frequency, $t(38) = .40$, $p = .69$.

Insert Table 2 about here

Procedure. The experiment was run in a sound attenuated, dimly lit room. It consisted of a familiarization phase, an experimental phase, and a visibility test phase. All 32 pictures were presented once in the familiarization phase. A trial started with the presentation of a fixation cross (‘+’) for 700 ms, followed by the target picture. After 1000 ms, the correct picture name appeared under the picture and subjects named the picture. The experimenter typed in whether the response was correct or not and/or if any voice key failure took place. The next trial was initiated after 1000 ms.

In the experimental phase, subjects worked through four blocks of trial. Before the first and third block, a practice phase was administered. The trial sequence of these phases was identical to the trial sequence of the experimental phase that followed. All subjects started with the masked condition. One half of the subjects first named the pictures in the frequency condition, the other half started with the semantic condition. In both conditions, a trial started with the presentation of ten hash marks for 500 ms, followed by the presentation of the distractor for three refresh cycles (50 ms). This was immediately followed by the backward mask, which was a different randomly generated consonant string on each trial. The string appeared alone on screen for one refresh cycle (16.67 ms), after which the picture and backward mask were presented together and subjects named the picture. Picture and mask were removed after 500 ms or sooner when the response was initiated earlier. The next trial was initiated 1000 ms after the experimenter had typed in whether the response was correct or not and/or if any voice key failure took place. Subsequently, subjects were presented with 2 blocks of trials in which distractors were visible. The trial structure was identical to the masked condition, with the exception that no masks were used. Also, each trial started with a fixation cross for 700 ms.

Each picture was presented four times in each block: twice with its low-frequency distractor and twice with its high-frequency distractor in the frequency condition, and twice with its semantically related distractor and twice with its semantically unrelated distractor in the semantic condition. This resulted in 80 trials per block, yielding 320 trials in total. Trials were randomized with the same restrictions as in Experiment 1. In addition, whether a certain picture would appear with its low- or high-frequency distractor, or semantically related or unrelated distractor first was counterbalanced across subjects. Between blocks and between the practice and the experimental phase, subjects could take a self-paced break. Written instructions appeared on screen before the practice phases and the experimental phases and were the same as in Experiment 1.

After having completed the experimental phase, subjects were debriefed and asked whether they noticed any distractor in the masked condition and whether they had seen all distractors in the visible condition. Next, subjects were administered a visibility test. In this test, we presented new stimuli that were matched to the distractor words on log frequency, frequency, number of letters, number of syllables, number of phonemes, number of neighbors, and bigram frequency. It consisted of two blocks. In both blocks, all pictures were presented four times, once with each associated matched word. Whether the picture appeared first in the frequency or semantic condition depended on the order of presentation in the experimental phase, as was whether it was shown first with its low- or high-frequency and semantically related or unrelated matched distractor. This gave rise to 40 trials per block, resulting in a total of 80 trials. In the first block, the trial sequence was identical to the masked condition. However, now subjects were instructed to indicate whether or not they had seen a distractor, one of its letters, or had noticed something appearing without knowing what they saw. In the second block, the trial sequence was identical to the trial sequence in the visible condition. The instructions were the same. In this control experiment, no subject indicated having noticed distractors in the masked condition and all reported being able to see and identify the distractors in the visible condition. Results of the visibility test confirmed this. No distractors were detected in the first block, and all subjects correctly reported all distractors in the visible condition.

In total, the experiment lasted approximately 45 minutes.

Results and discussion. Analyses (see Figure 1 and Figure 2) were performed separately for the frequency and the semantic condition. F_1 analyses (on subjects' means) and F_2 analyses (on items' means) were conducted on response latencies with distractor frequency (2, low vs. high-frequency) or relatedness (2, semantically related vs. unrelated) and visibility condition (2, masked vs. visible) as within-subject and within-item variables. Errors and outliers accounted for 0.77% and 2.00% of the data respectively, and were not further analyzed. There were no naming errors or verbal disfluencies.

Insert Figure 1 and Figure 2 about here

Analyses of the frequency condition showed no effect of visibility condition by subjects, $F_1(1,39) = 2.63$, $MSE = 2820.94$, $p = .11$, but a significant effect by items, $F_2(1,19) = 8.68$, $MSE = 2406.18$, $p < .01$, with shorter naming latencies in the masked condition. The main effect of frequency was significant by subjects, $F_1(1,39) = 17.53$, $MSE = 5219.77$, $p < .001$, and by items, $F_2(1,19) = 7.49$, $MSE = 2692.83$, $p < .05$. Subjects responded faster when the picture was paired with a high- compared to a low-frequency word. Most importantly, the interaction between frequency and visibility condition was significant by subjects $F_1(1,39) = 42.24$, $MSE = 6528.12$, $p < .001$ and items, $F_2(1,19) = 14.85$, $MSE = 2429.45$, $p < .01$. Paired samples t-tests showed that this interaction was due to a significant distractor frequency effect in the visible condition, $t_1(39) = 6.23$, $p < .001$, $t_2(19) = 3.73$, $p < .01$, but not in the masked condition, $t_1(39) = -0.49$, $p = .63$, $t_2(19) = 0.15$, $p = .88$.

Analyses of the semantic conditions showed no significant main effect of relatedness by subjects, $F < 1$; or items, $F(1,19) = 1.29$, $MSE = 311.85$; $p = .27$. The main effect of visibility condition was significant by items, $F_2(1,19) = 51.21$, $MSE = 19293.03$, $p < .001$, and by subjects, $F_1(1,39) = 9.10$, $MSE = 18288.15$, $p < .01$. Again, responses were slower in the visible condition.

Like in the frequency condition, the interaction between the two variables reached significance by both subjects, $F_1(1,39) = 26.62$, MSE = 7494.45, $p < .001$ and items, $F_2(1,19) = 17.57$, MSE = 3754.76, $p < .001$. Paired samples t-tests showed that there was significant semantic interference in the visible condition by subjects, $t_1(39) = -3.81$, $p < .001$, and by items, $t_2(19) = -2.96$, $p < .01$. Most importantly, there was semantic facilitation in the masked condition that was significant by subjects, $t_1(39) = 5.05$, $p < .001$, and by items, $t_2(19) = 3.07$, $p < .01$.

These results confirmed the predictions made by the response exclusion hypothesis. Under masked conditions, where no response to the distractor should be formed, no distractor frequency effect was found. However, when the distractor was not masked, a clear distractor frequency effect was present. This experiment provides evidence for the assumption that the distractor frequency effect depends on the presence versus absence of a response to the distractor in the output buffer. One could argue of course that the absence of a distractor frequency effect under masked conditions is due to absence of processing the distractor, but this is unlikely. We introduced a control condition, in which distractors and picture were semantically related. If distractors are not processed under masked conditions, no effect should be found with semantically related distractors either. However, a semantic facilitation effect was present. As stimuli were presented under the same masking and timing parameters in both conditions, it seems unlikely that the same subjects could only process the distractors when they were semantically related to the target. Also, the same semantic facilitation effect was found by Finkbeiner & Caramazza (2006) who used the same presentation parameters as we did. These two controls make it unlikely that the absence of the distractor frequency effect is related to an absence of processing the distractor. Of course, the stimuli in the semantic condition were all of moderate frequency. Therefore, if it is assumed that processing capacity is correlated with frequency, it could still be argued that the high frequency distractors were processed, but not the low frequency distractors. This seems very unlikely. If a low frequency distractor should not have been processed, the situation in the frequency condition would be that the picture presented with a low frequency word corresponds to the condition where no distractor is present. This would be equal to a paradigm in which a picture is presented with a distractor (i.e., the case of the high frequency word) and without a distractor (i.e., the case of the low frequency word). A standard finding in the picture-word interference paradigm is that naming latencies are longer when the picture is presented with a word than when is it presented without a word (e.g., Ehri, 1976). Therefore, we should have found longer naming latencies with pictures accompanied with a high frequency distractor. This, however, was not the case. Thus, the experiment thus shows that a response buffer plays a role in the PWI task, and that effects depend on this factor. Furthermore, by replicating the semantic facilitation effect, this experiment also provides further evidence for a non-competitive model of lexical selection. Apparently, when post-lexical processes are eliminated and effects only reflect lexical processing, the standard semantic interference reverses into semantic facilitation.

The next experiment tested a further prediction of the response exclusion account, namely that the distractor frequency effect should disappear when there is time enough to exclude the response to the distractor from the response buffer.

Experiment 3

The second experiment showed that manipulating the presence versus absence of a response in the output buffer results in the absence versus presence of the distractor frequency effect. In the third experiment, the goal was to examine a gradual version of the response exclusion account of the distractor frequency effect. We hypothesized that the more time subjects had to remove the response to the distractor from the buffer, the smaller the distractor frequency effect would be. Therefore, distractors were presented before the pictures at four different points in time.

Method

Subjects. Twenty-four further subjects (22 women and 2 men, age range 18-32) took part in the experiment and received €5 for their participation. All gave written informed consent and reported normal or corrected to normal vision. All subjects were native speakers of Dutch and naïve to the purpose of the experiment.

Design. Naming latency was considered as dependent variable. There were two independent variables: frequency of the distractor and stimulus onset asynchrony (i.e., SOA). Frequency had two levels: high- and low-frequency. SOA had four levels: 0 ms, -100 ms, -200 ms and -300 ms.

Stimuli and material. The same stimuli and material was used as in Experiment 2 with the exception that no semantic condition was included.

Procedure. The experiment was run in a sound attenuated, dimly lit room and consisted of a familiarization phase and an experimental phase. The familiarization phase was identical to Experiment 2.

The experimental phase consisted of four blocks; each block corresponded to one SOA level. The order of blocks was counterbalanced across subjects. Before each block, subjects were administered a practice phase. This consisted of one block of 12 trials. Each trial started with the presentation of a fixation cross (‘+’) for 700 ms. Then, the distractor word appeared. The picture appeared either together with the distractor (0 ms SOA condition) or 100 ms, 200 ms or 300 ms after distractor onset, depending on SOA condition. Distractor and picture stayed on screen until the subject had named the picture or until 500 ms had elapsed. The next trial was initiated 1000 ms after the experimenter had typed in whether the response was correct or not and/or if any voice key failure took place. Each picture was presented four times in each block: twice with its low-frequency distractor and twice with its high-frequency distractor. This resulted in 80 trials per block and 320 trials in total. With which distractor a given picture appeared first was counterbalanced across subjects. Trials were presented in a different random order for each subject with the same restrictions as in the previous experiments. Between blocks and between the practice and the experimental phase, subjects could take a self-paced break. Written instructions appeared on screen before the practice phases and the experimental phases and were the same as in the previous experiments.

In total, the experiment lasted approximately 45 minutes.

Results and discussion. Analyses (see Figure 3) were performed with SOA (4; 0 ms, -100 ms, -200 ms and -300 ms) and frequency of the distractor (low vs. high-frequency) as within-subject and within-item variables. Errors and outliers accounted for 1.29% and 2.15% of the data and were not further analyzed. No naming errors or verbal disfluencies were recorded.

Insert Figure 3 about here

There was a significant distractor frequency effect, both by subjects, $F_1(1,23) = 33.36$, MSE = 6856.38, $p < .001$, and by items, $F_2(1,19) = 11.67$, MSE = 5338.96, $p < .01$. The main effect of SOA was significant as well, both by subjects, $F_1(3,21) = 39.44$, MSE = 43542.95, $p < .001$, and items, $F_2(3,17) = 140.90$, MSE = 36389.46, $p < .001$. Naming latencies decreased as the SOA increased. Paired samples t-tests showed that there was a significant difference between the SOA 0 ms and -100 ms, $t_1(23) = 7.79$, $p < .001$, $t_2(19) = 17.17$, $p < .001$, between 0 ms and -200 ms, $t_1(23) = 7.55$, $p < .001$, $t_2(19) = 18.31$, $p < .001$, and between 0 ms and -300 ms, $t_1(23) =$

8.97, $p < .001$, $t_2(19) = 19.30$, $p < .001$. No other comparison reached significance by subjects (smallest p-value = .11). By items, there were additional differences between -100 ms and -300 ms, $t_2(19) = 3.48$, $p < .01$, and between -200 ms and -300 ms, $t_2(19) = 5.85$, $p < .001$, but not between -100 ms and -200 ms, $t_2(19) = -1.58$, $p = .13$. In line with our predictions, the interaction between the two variables was significant by subjects, $F_1(3,21) = 3.90$, MSE = 1018.55, $p < .05$, and marginally significant by items, $F_2(3,17) = 3.17$, MSE =814.01, $p = .05$. Paired samples t-tests showed that there was a significant distractor frequency effect at 0 ms, $t_1(23) = 4.90$, $p < .001$, $t_2(19) = 3.84$, $p < .01$, at -100 ms, $t_1(23) = 3.84$, $p < .01$, $t_2(19) = 2.31$, $p < .05$, a marginally significant effect at -200 ms, $t_1(23) = 1.89$, $p = .07$, $t_2(19) = 1.44$, $p = .17$, but no effect at -300 ms, $t_1(23) = .74$, $p = .47$, $t_2(19) = .81$, $p = .43$.

We also performed a linear regression with SOA as predictor and the difference in naming latencies between low and high-frequency distractors as dependent variable. Results showed that the effect decreased linearly with increasing SOA, $F(1,95) = 13.30$, $p < .001$, standardized ß = -.35.

These results support the response exclusion account. The distractor frequency effect decreased linearly with increasing SOA. That is, the more time subjects had to remove the response from the buffer, the smaller the effect became. The distractor frequency effect was clearly present in the 0 ms and 100 ms SOA condition, but decreased strongly in the SOA 200 ms and was negligible in the 300 ms condition. These results allow us to draw conclusions about the time course of the distractor frequency effect and the response buffer. Apparently, the control process operating over the response buffer takes no more than 200 ms to completely remove the response from the buffer. Results are also in line with Miozzo and Caramazza’s (2003, Experiment 5) data. They presented a low- or high-frequency distractor either together with the picture, or 100 ms before or 100 ms after the presentation of the picture. They found a reliable distractor frequency effect at all SOAs. We extended this experiment by showing that the distractor frequency effect is indeed present at an SOA of -100 ms, but decreases at longer SOAs.

General Discussion

In this paper, we investigated the origin of the distractor frequency effect in the PWI paradigm (Miozzo & Caramazza, 2003). We started from the response exclusion hypothesis that states that this effect is related to an output buffer that forms a bottleneck (Finkbeiner & Caramazza, 2006; Janssen et al., 2009; Mahon et al., 2007; Miozzo & Caramazza, 2003). It is assumed that words have a privileged relationship to the buffer. In the PWI paradigm, the response to the distractor word will be formed first and has to be purged from the buffer before the picture can be named. As low-frequency words have a lower resting level (McClelland & Rumelhart, 1981), they enter and are removed from the buffer later compared to high-frequency words. This gives rise to the distractor frequency effect. In a first experiment the distractor frequency effect was replicated even after having changed some of the parameters used by Miozzo and Caramazza (2003). This experiment showed that the distractor frequency effect is a robust finding. The next experiments tested the response exclusion hypothesis directly. In Experiment 2, distractors were presented under masked and visible conditions. Masking should prevent the formulation of a response to the distractor (Finkbeiner & Caramazza, 2008), leading to the absence of a frequency effect. In line with predictions, an effect of distractor frequency was only found in the visible condition. A control condition in which semantic relatedness between picture and distractor was manipulated suggests that distractors were processed under masked conditions, as semantic facilitation was found. To examine the response selection hypothesis in a different way, we also tested a more gradual formulation of the hypothesis. In Experiment 3, distractors were presented at various time points before the picture. The more time there was between the presentation of the distractor and the picture the smaller the distractor frequency effect became. This suggests that the more time there was to remove the response from the buffer, the less the response interfered with picture naming, and the interference effect even disappeared at 200 ms.

These experiments are the first to report evidence for the linkage between the response exclusion hypothesis and the distractor frequency effect. They have provided support for the assumption that the distractor frequency effect arises at a post-lexical stage at the level of an output buffer. The hypothesis that the effect arises at a post-lexical origin is consistent with findings reported by Miozzo and Caramazza (2003). In one experiment, they presented subjects with a PWI task in which distractors were or were not repeatedly read aloud by the subject before the start of the experiment. The authors hypothesized that the effect of both high-frequency words and of previously read distractors should be attributable to the fact that they were repeatedly selected for production. If repeated production plays a critical role in the distractor frequency effect, previously read aloud words should interfere less compared to words that have not been read aloud. In line with the output locus, pictures presented with a distractor that had been read aloud previously were named faster, similar to the effect of a high-frequency distractor. Further evidence for an output locus comes from the observation that pictures were named more slowly when accompanied by a low-frequency distractor, but only when picture and distractor were not phonologically related (Miozzo & Caramazza, 2003; Experiment 7). The interaction was interpreted as evidence for a common locus of both effects at output level. Experiment 2 and Experiment 3 of this paper provide further evidence for an output locus of the distractor frequency effect, and at the same time refine the output account: the effect has to be situated at the level of an output buffer that forms a bottleneck within word production. If no response to the distractor can be formed or if the response is already purged from the buffer when the picture name becomes available, the distractor will not influence picture naming. However, when the response to the distractor still occupies the buffer when the response to the picture becomes available, picture naming has to be postponed until the initial response is purged from the buffer.

The post-lexical origin of the distractor frequency effect allows drawing a parallel with the semantic interference effect (e.g., Lupker, 1979). Recent research leads to the assumption that this effect arises at a post-lexical origin as well (e.g., Finkbeiner & Caramazza, 2006; Janssen et al., 2008). According to this assumption, the semantic interference effect and the distractor frequency effect have a similar origin as they are both related to the response buffer but differ in the point in time at which the effects arise. Additional support for a linkage between both effects comes from an experiment by Bloem et al. (2004). They show that the semantic interference effect has a similar time course as the distractor frequency effect. They presented semantically related and unrelated distractors with to be named pictures at different SOAs. They found semantic interference when the distractor was presented in close proximity to the picture, but facilitation when the distractor was presented 400 ms before the target. They explained this result by assuming that activation at the lexical level decays more rapidly than activation at the conceptual level. However, the data of Experiment 3 suggest that responses to distractors can be removed as early as 200 ms after presentation. Thus, one can assume that a semantically related distractor is already removed from the buffer after 400 ms. But when picture and distractor are presented in close temporal proximity, the distractor will not yet have left the buffer, leading to a semantic interference effect (see also Finkbeiner & Caramazza, 2006).

The results of the experiments have, albeit indirect, implications for the current models of word production and the mechanisms of lexical selection they propose. As stated earlier, competition models cannot explain the distractor frequency effect (e.g., Levelt et al., 1999; Roelofs, 1992, 1993, 2001, 2003), but neither can non-competition models (e.g., Caramazza, 1997; Dell, 1986). As the activation of the surrounding nodes should not influence lexical selection, there should not be an effect of distractor frequency. Additional assumptions about the origin of the effect, made by the response exclusion hypothesis, are thus needed. Our experiments provided clear evidence for these assumptions, thereby considerably strengthening the basis for non-competition models. Of course, experiments designed to perform a direct comparison between both types of models are needed.

However, the response exclusion account is presently rather underspecified. Specifically, the nature of the control mechanisms operating on the response buffer remains unclear. The data of these and other experiments support non-competition models, but also indicate that non-competition models need to be extended with additional assumptions to explain key effects like the semantic interference effect. That is, they have to assume a response buffer over which a control mechanism operates and checks the response for correctness; but the nature of this control mechanism has so far not been specified. Postulating a new mechanism, that would be only operative in situations like the PWI task, is of course not a very parsimonious road to take. Instead, our tentative suggestion is that the function of checking the output buffer, and purging incorrect responses from it, is subserved by a mechanism that is always in place in speech production, namely the verbal self-monitoring system (e.g., Hartsuiker & Kolk, 2001; Levelt, 1989; Levelt et al., 1999; but see Mahon et al., 2007). The normal function of the self-monitoring is to inspect internal and external speech for problems, such as speech errors, inappropriate lexical choice, and so on. It is not a far-fetched step to assume that during picture-word interference, the self-monitor also checks whether a speech plan is consistent with the speaking task at hand, and initiates a (time-consuming) correction when this is not the case.

Indeed, some findings appear to be consistent with this proposal. First, several authors have argued that the monitoring process employs the criterion of lexicality (i.e., is it a word?). Because of this property, it is easier for the monitor to detect and correct a non-word error compared to a word error, thus leading to the production of more word errors than non-word errors (the lexical bias effect, e.g., Baars, Motley, & MacKay, 1975; Hartsuiker, Corley, & Martensen, 2005; Levelt et al.,1999; Roelofs, 2004). This lexicality criterion is consistent with the finding that in a PWI task, a non-word interferes less than a word (e.g., Klein, 1964; Lupker 1979; Lupker & Katz, 1982; Rosinski et al., 1975): a non-word is more easily detected as inconsistent with the task at hand, and so is excluded as a response earlier. Second, Starreveld and LaHeij (1999) showed that in a speeded PWI task, subjects often named the distractor instead of the picture. Whereas normally the self-monitor has enough time to check the response for correctness and delete incorrect responses (i.e., responses to the distractor), the time pressure may sometimes prevent this process. As the response to the distractor is formed first, subjects will on some trials respond to the distractor and not to the picture. Even more, the errors also showed a semantic effect: more errors were made when picture and distractor were semantically related compared to when they were unrelated. Thus, when the distractor is related, it appears to be even more difficult to be detected by the self-monitor, and consequentially will be rejected less compared to an unrelated distractor. Finally, we note that monitoring accounts have most often been proposed in the context of speech-error research (see Hartsuiker, 2006; for review and discussion). However, several studies suggest that monitoring mechanisms can also have repercussions for the *time-course* of language production, which is of course most relevant for picture-word interference (Boland, Hartsuiker, Pickering, & Postma, 2005; Hartsuiker, Pickering, & De Jong, 2005; Van Wijk & Kempen, 1987). Boland et al., for example, showed that when subjects had prepared a noun phrase to describe a stimulus (e.g., *blue square* or *light blue square*), but then had to revise their utterance (because a change in context rendered the prepared utterance under- or over-informative), there was a larger reaction time cost when they needed to add a word than delete a word. No such additional cost was observed when they said the original utterance aloud, instead of preparing it, excluding the possibility that the effect resulted from lexical priming. Instead, they argued that the reaction time cost resulted from covert editing to a prepared response, which was more time-consuming when an additional word needed to be retrieved from the lexicon. Thus, equating the control process with the self-monitoring system would not only be more parsimonious, it would also allow for an integration between different domains in language production, that is, between error monitoring and models of lexical selection. Of course, further research is needed to investigate the nature of the control mechanism operating over the response buffer. A fruitful direction in our opinion would be to investigate the parallels with the self monitoring system.

Finally, these and related experiments (e.g., Janssen et al., 2008, Finkbeiner & Caramazza, 2006) also have implications for the tasks that could be used when studying lexical access. The results support the assumption that the standard PWI task might not always be suitable to study language production. The response exclusion account assumes that with the standard PWI paradigm, the response to a word will always be formed first. Effects of the distractor could thus be due to operations related to the response buffer. Therefore, results might not be informative of lexical selection *per se* but could also reflect post-lexical operations related to the buffer. This implies that if no attention is paid towards the role of the response buffer, the results of PWI-studies must be interpreted with great caution. We suggest a number of recommendations for the use of the PWI paradigm that would result in minimizing the potential role of a buffer. Regarding the task itself, some adjustments can be made to eliminate the potential interference caused by the response buffer. For example, our experiments show that masking the distractor (see also Finkbeiner & Caramazza, 2008) or presenting it early enough before picture presentation could be a useful technique. Regarding the stimuli, two main factors have to be taken into account. First, as shown by the distractor frequency effect, stimuli should not differ in how fast they can enter the response buffer. This implies that care has to be taken to match stimuli on variables that might influence this factor such as frequency, stimulus quality etc. Second, as shown by the semantic interference effect, stimuli should not differ in to what extent they satisfy constraints imposed by the picture, or more broadly, in how easily they are removed from the buffer. Another possibility, of course, would be to consider alternative methods, such as the semantic blocking paradigm (e.g., Damian & Als, 2005; Damian et al., 2001; Howard, Nickels, Coltheart, & Cole-Virtue, 2006), translation tasks (Bloem & La Heij, 2003), or visual world eye-tracking tasks (Huettig & Hartsuiker, 2008), instead of or in addition to the PWI task. Converging results from different paradigms would surely add to the reliability of the evidence and conclusions that are drawn.

In conclusion, the experiments reported in this paper add to the existing evidence in favor of the response exclusion hypothesis. We have shown that this account elegantly explains the distractor frequency effect, which has implications for models of lexical selection and for the tasks that can be used to study lexical access. Future research should focus on the difference between the two classes of models of lexical selection, and on the nature of the control mechanism operating over the response buffer.

References

Abdel Rahman, R., & Melinger, A. (2007). When bees hamper the production of honey: Lexical interference from associates in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 604-614.

Alario, F. X., Segui, J., & Ferrand, L. (2000). Semantic and associative priming in picture naming. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 53A*, 741–764.

Baars, B. J., Motley, J. T., & MacKay, D. (1975). Output editing for lexical status from artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior, 14*, 382-391.

Belke, E., Meyer, A., & Damian, M. F. (2005). Refractory effect in picture naming as assessed in a semantic blocking paradigm. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *58A*, 667-692.

Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. *Journal of Memory and Language*, *48*, 468-488.

Bloem, I., van den Boogaard, S., & La Heij, W. (2004). Semantic facilitation and semantic interference in language production: Further evidence for the conceptual selection model of lexical access. *Journal of Memory and Language*, *51*, 307-323.

Boland, H. T., Hartsuiker, R. J., Pickering, M. J. & Postma, A. (2005). Repairing inappropriate utterances: Revision or restart? *Psychonomic Bulletin & Review, 112,* 472-477.

Brysbaert, M., & Ghyselinck, M. (2006). The effect of age of acquisition. Partly frequency related, partly independent. *Visual cognition*, *12,* 992-1011.

Burt, J. S. (2002). Why do non-color words interfere with color naming? *Journal of Experimental Psychology: Human Perception and Performance, 28*, 1019–1038.

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

Caramazza, A., & Costa, A. (2000). The semantic interference effect in the picture–word interference paradigm: Does the response set matter*? Cognition*, *75*, B51-B64.

Costa, A., Alario, F. X., & Caramazza, A. (2005). On the categorical nature of the semantic interference effect in the picture–word interference paradigm. *Psychonomic Bulletin & Review, 12,* 125–131.

Damian, M. F., & Als, L. C. (2005). Long-lasting semantic context effects in the spoken production of object names. *Journal of Experimental Psychology: Learning, Memory & Cognition, 31*, 1372-1384.

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

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

Damian, M. F., Vigliocco, V., & Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, *81,* B77–B86.

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

Duyck, W., Anseel, F., Szmalec, A., Mestdagg P., Tavernier, A., & Hartsuiker, R. J. (2008). Improving accuracy in detecting acoustic onsets. *Journal of Experimental Psychology:*

Human Perception and Performance, 43, 317-326.

Ehri, L; C. (1976). Do words really interfere in naming pictures? *Child Development, 42,* 502-505.

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

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

Finkbeiner, M., & Caramazza, A. (2008). Modulating the masked congruence priming effect with the hands and the mouth. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 894-918.

Fox, L. A., Shor, R. E., & Steinman, R. J. (1971). Semantic gradients and interference in naming color, spatial direction, and numerosity. *Journal of Experimental Psychology, 91,* 59–65.

Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General, 118,* 13-42.

Hantsch, A., Jescheniak, J. D., & Schriefers, H. (2005). Semantic competition between hierarchically related words during speech planning. *Memory & Cognition, 33,* 984–1000.

Hartsuiker, R. J. (2006). Are speech error patterns affected by a monitoring bias? *Language and Cognitive Processes, 21,* 856-891.

Hartsuiker, R. J., Corley, M., & Martensen, H. (2005). 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, 52,* 58-70.

Hartsuiker, R. J., & Kolk, H. H. J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology, 42,* 113-157.

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

Huettig, F., & Hartsuiker, R. J. (2008). When you name the pizza you look at the coin and the bread: Eye movements reveal semantic activation during word production. *Memory & Cognition, 36*, 341-360.

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: experimental and computational studies. *Cognition, 100*, 464-482.

Janssen, N., Schirm, W., Mahon, B.Z., & Caramazza, A. (2008). Semantic interference in a delayed naming task: Evidence for the response exclusion hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 249-256.

Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *American Journal of Psychology, 77,* 576–588.

Koechlin, E., Naccache, L., Block, E., & Dehaene, S. (1999). Primed numbers: Exploring the modularity of numerical representations with masked and unmasked semantic priming. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1882–1905.

Kunde, W., Kiesel, A., & Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition, 88,* 223–242.

LaHeij, W., Happel, B., & Mulder, M. (1990). Components of Stroop-like interference in word reading. *Acta Psychologica, 73,* 115–129.

LaHeij, W., Kuipers, J.-R., & Starreveld, P. A. (2006). In defence of the lexical selection by competition account of picture-word interference: a comment on Finkbeiner and Caramazza (2006). *Cortex, 42,* 1028-1031.

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

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

Lupker, S. J. (1979). The semantic nature of response competition in the Picture–Word Interference Task. *Memory & Cognition, 7,* 485–495.

Lupker, S. J., & Katz, A. N. (1982). Input, decision, and response factors in picture–word interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 7,* 269–282.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203.

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 503-533.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception. *Psychological Review, 88,* 375–407.

Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distractor frequency in picture–word interference paradigm. *Journal of Experimental Psychology: General, 132,* 228–252.

Naccache, L., & Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition, 80,* 215–229.

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

Roelofs, A. (1993). Testing a non-decompositional theory of lemma retrieval in speaking: Retrieval of verbs. *Cognition, 47,* 59–87.

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

Roelofs, A. (2001). Set size and repetition matter: Comment on Caramazza and Costa (2000). *Cognition, 80,* 283–290.

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

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

Rosinski, R. R. (1977). Picture–word interference is semantically based. *Child Development, 48,* 643–647.

Rosinski, R. R., Golinkoff, R. M., & Kukish, K. S. (1975). Automatic semantic processing in a picture–word interference task. *Child Development, 26,* 247–253.

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

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

Starreveld, P. A., & La Heij, W. (1999). Word substitution errors in a speeded picture-word task.

American Journal of Psychology, 112, 521-553.

Severens, E., Van Lommel, S., Ratinckx, E., & Hartsuiker, R.J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica, 119*, 159-187.

Stevens, M., Lammertyn, J., Verbruggen, F., Vandierendonck, A. (2006). Tscope: A C library for programming cognitive experiments on the MS Windows platform. *Behavior Research Methods, 38*, 280-286.

van Wijk, C., & Kempen, G. (1987). A dual system for producing self-repairs in spontaneous speech: Evidence from experimentally elicited corrections. *Cognitive Psychology, 19*, 403-440.

Vigliocco, G., Lauer, M., Damian, M. F., & Levelt, W. J. M. (2002). Semantic and syntactic forces in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 46–58.

Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology, 48*, 422-488

Vigliocco, G., Vinson, D. P., & Siri, S. (2005). Semantic similarity and grammatical class in naming actions. *Cognition, 94*, B91-B100.

Appendix 1. Materials from Experiment 1 (English translations between parentheses).

Picture Name	Low-frequency	High-frequency
Aap *(monkey)*	Ballon *(balloon)*	Trein *(train)*
Appel *(apple)*	Broek *(pants)*	Bloem *(flower)*
Baard *(beard)*	Kalkoen *(turkey)*	Soldaat *(soldier)*
Bom *(bomb)*	Oorbel *(earring)*	Bureau *(desk)*
Eend *(duck)*	Slee *(sled)*	Glas *(glass)*
Emmer *(bucket)*	Banaan *(banana)*	Vinger *(finger)*
Haak *(hook)*	Clown *(clown)*	Stoel *(chair)*
Heks *(witch)*	Gieter *(watering can)*	Koning *(king)*
Kaars *(candle)*	Schommel *(swing)*	Telefoon *(telephone)*
Ketting *(chain)*	Mug *(mosquito)*	Bij *(bee)*
Kooi (cage)	Zwaan *(swan)*	Brief *(letter)*
Konijn *(rabbit)*	Fakkel *(torch)*	Vuur *(fire)*
Kroon *(crown)*	Hengel *(fishing*	Vogel *(bird)*

	pole)	
Kussen *(pillow)*	Hooi *(hay)*	Wijn *(wine)*
Leeuw *(lion)*	Lat *(ruler)*	Raam *(window)*
Lepel *(spoon)*	Kers *(cherry)*	Kerk *(church)*
Lucifer *(match)*	Egel *(hedgehog)*	Hart *(hart)*
Mand *(basket)*	Papegaai *(parrot)*	Schouder *(shoulder)*
Masker *(mask)*	Slak *(snail)*	Steen *(stone)*
Pet *(hat)*	Kikker *(frog)*	Dokter *(doctor)*
Pijl *(arrow)*	Mais *(corn)*	Tafel *(table)*
Pijp *(pipe)*	Bever *(beaver)*	Geest *(ghost)*
Sigaar *(cigar)*	Stok *(cane)*	Tand *(tooth)*
Slang *(snake)*	Kano *(canoe)*	Blik *(can)*
Tent *(tent)*	Rits *(zipper)*	Stad *(city)*
Trommel *(drum)*	Mier *(ant)*	Muur *(wall)*
Varken *(pig)*	Robot *(robot)*	Beeld *(statue)*
Vlag *(flag)*	Spin *(spider)*	Blad *(leaf)*
Wiel *(tire)*	Draak *(dragon)*	Vrouw *(woman)*
Zwembad *(swimming pool)*	IJsje *(ice cream)*	Paard *(horse)*

Appendix 2. Materials used in experiment 2 and 3 (English translations between parentheses).

Picture name	Low-frequency	High-frequency	Related	Unrelated
Aap *(monkey)*	Stok *(cane)*	Film *(film)*	Beer *(bear)*	Riem *(belt)*
Appel *(apple)*	Robot *(robot)*	Beeld *(statue)*	Peer *(pear)*	Geit *(goat)*
Boerderij *(farm)*	Vork *(fork)*	Neus *(nose)*	Schuur *(shed)*	Cadeau *(present)*
Bom *(bomb)*	Gebit *(teeth)*	Tafel *(table)*	Granaat *(grenade)*	Tribune *(stand)*
Boot *(boat)*	Sjaal *(scarf)*	Vogel *(bird)*	Kano *(canoe)*	Slak *(snail)*
Duim *(thumb)*	Kalkoen *(turkey)*	Soldaat *(soldier)*	Pink *(little finger)*	Pauw *(peacock)*
Hemd *(shirt)*	Gieter *(watering can)*	Koning *(king)*	Vest *(vest)*	Wieg *(cradle)*
Lamp *(lamp)*	Hooi *(hay)*	Kerk *(church)*	Kaars *(candle)*	Boter *(butter)*
Leeuw *(lion)*	Oorbel *(earring)*	Bureau *(desk)*	Tijger *(tiger)*	Wekker *(alarm clock)*
Lepel *(spoon)*	Zwaan *(swan)*	Bloem *(flower)*	Mes *(knife)*	Ton *(barrel)*
Pot *(pot)*	Draak *(dragon)*	Vrouw *(woman)*	Deksel *(lid)*	Masker *(mask)*
Revolver	Mier *(ant)*	Blad *(leaf)*	Geweer *(rifle)*	Terras *(deck)*

(revolver)				
Ring *(ring)*	Schommel *(swing)*	Schouder *(shoulder)*	Ketting *(necklace)*	Lucifer *(match)*
Rok *(skirt)*	Hert *(deer)*	Wijn *(wine)*	Broek *(pants)*	Vloer *(floor)*
Sigaar *(cigar)*	Bever *(bever)*	Geest *(goast)*	Pijp *(pipe)*	Nest *(nest)*
Televisie *(television)*	Egel *(hedgehog)*	Stad *(city)*	Radio *(radio)*	Motor *(motorbike)*
Trommel *(drum)*	Slee *(sledge)*	Glas *(glass)*	Cimbaal *(cymbal)*	Wasbeer *(raccoon)*
Vliegtuig *(airplane)*	IJsje *(ice cream)*	Paard *(horse)*	Helikopter *(helicopter)*	Handschoen *(glove)*
Wortel *(carrot)*	Fluit *(flute)*	Steen *(stone)*	Erwt *(pea)*	Pomp *(pump)*
Zetel *(couch)*	Kers *(cherry)*	Been *(leg)*	Stoel *(chair)*	Aarde *(earth)*

Footnotes

[1] Of course, there is spreading of activation in the reverse direction too. But most models assume that the representation of the target concept is more active than that of the distractor concept, because the target concept is activated directly (by the picture) rather than indirectly (from the word) and because the target concept would receive additional ‘task activation’. Therefore, there is more activation spreading from the concept dog to concept cat than vice versa.

Table 1. Properties of low and high-frequency items used in Experiment 1 (standard deviations between parentheses).

	Low Frequency	High Frequency
Log frequency	0.60 (0.30)	2.27 (0.33)
Length	5.00 (1.29)	5.04 (1.20)
AOA	5.29 (1.10)	5.52 (0.98)
Syllables	1.53 (0.57)	1.33 (0.54)
Phonemes	4.37 (0.96)	4.29 (1.04)
Neighbors	6.40 (5.89)	5.96 (4.78)
Bigram frequency	36614.13 (22470.24)	37397.58 (23936.87)

Table 2. Properties of items used in experiments 2 and 3 (standard deviations between parentheses).

	Low-frequency	High-frequency	Related	Unrelated
Frequency	5.30 (3.47)	195.95 (177.28)	27.20 (33.73)	26.75 (34.07)
Log frequency	0.61 (0.35)	2.21 (0.24)	1.20 (0.49)	1.19 (0.48)
Length	4.95 (1.10)	4.95 (1.10)	5.30 (1.63)	5.30 (1.63)
AOA	5.75 (1.04)	5.53 (1.13)	N.A.	N.A.
Syllables	1.45 (0.51)	1.30 (0.47)	1.60 (0.82)	1.65 (0.67)
Phonemes	4.25 (0.91)	4.20 (0.83)	4.60 (1.60)	4.65 (1.49)
Neighbors	5.70 (5.58)	5.55 (5.01)	5.85 (6.52)	5.40 (5.33)
Bigram frequency	38839.75 (21017.92)	35805.45 (22560.58)	44779.10 (24231.98)	41628.25 (26043.93)

Note. N.A., not available.

Figure captions

Figure 1. Mean naming latencies with confidence interval bars for low- and high frequency words in the frequency condition, analyses by subjects, Experiment 2.

Figure 2. Mean naming latencies with confidence interval bars for low- and high-frequency words in the semantic condition, analyses by subjects, Experiment 2.

Figure 3. Mean naming latencies with confidence interval bars of low- and high-frequency words in the four SOA conditions, analyses by subjects, Experiment 3.

Figure 1.

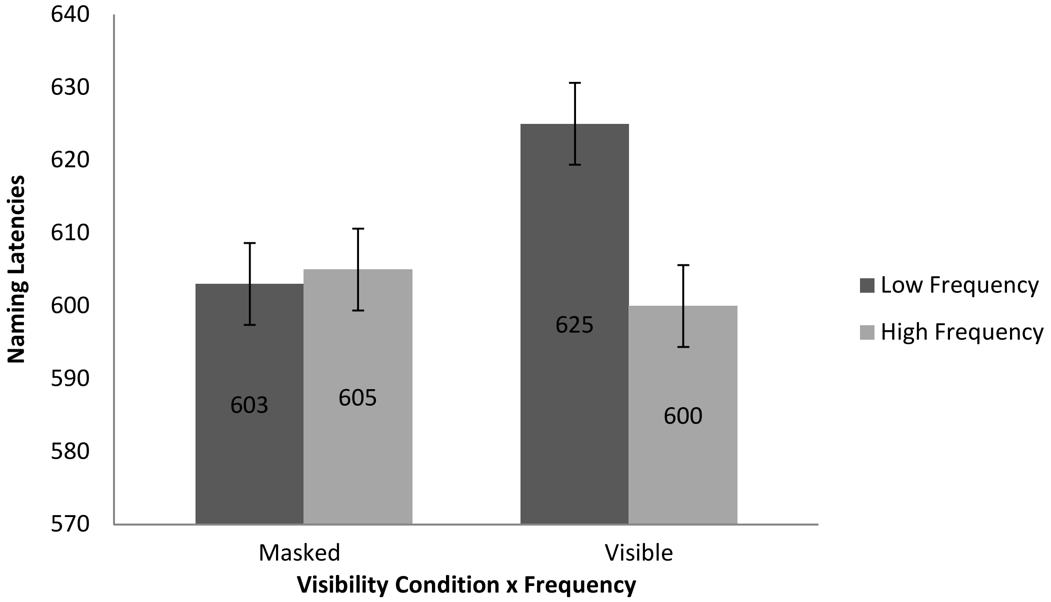

640
630
620
610
600
590
580
570
Naming Latencies
603
605
625
600
Masked
Visible
Visibility Condition x Frequency
Low Frequency
High Frequency

Figure 2.

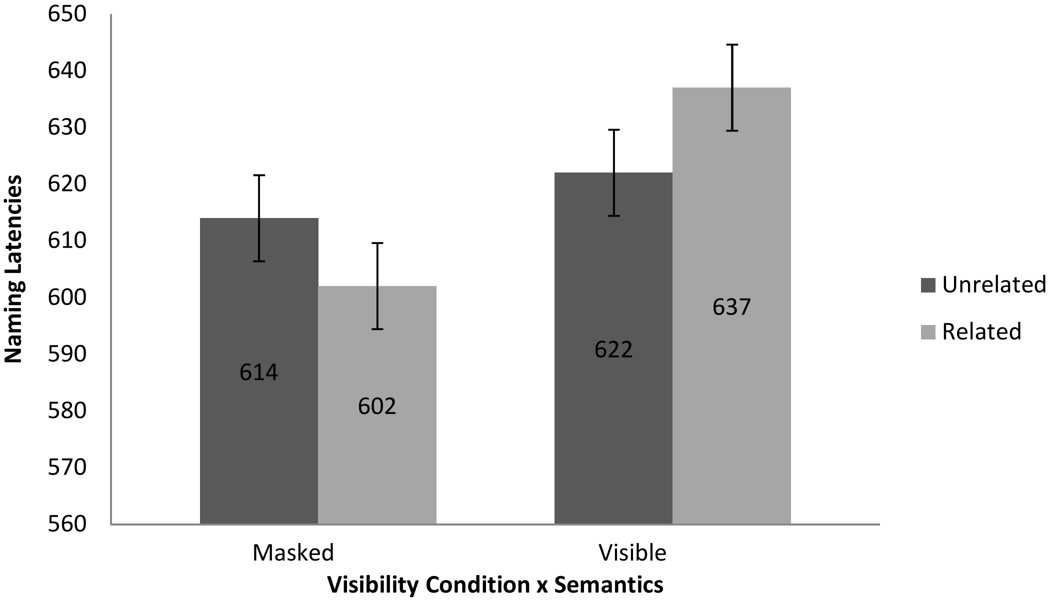
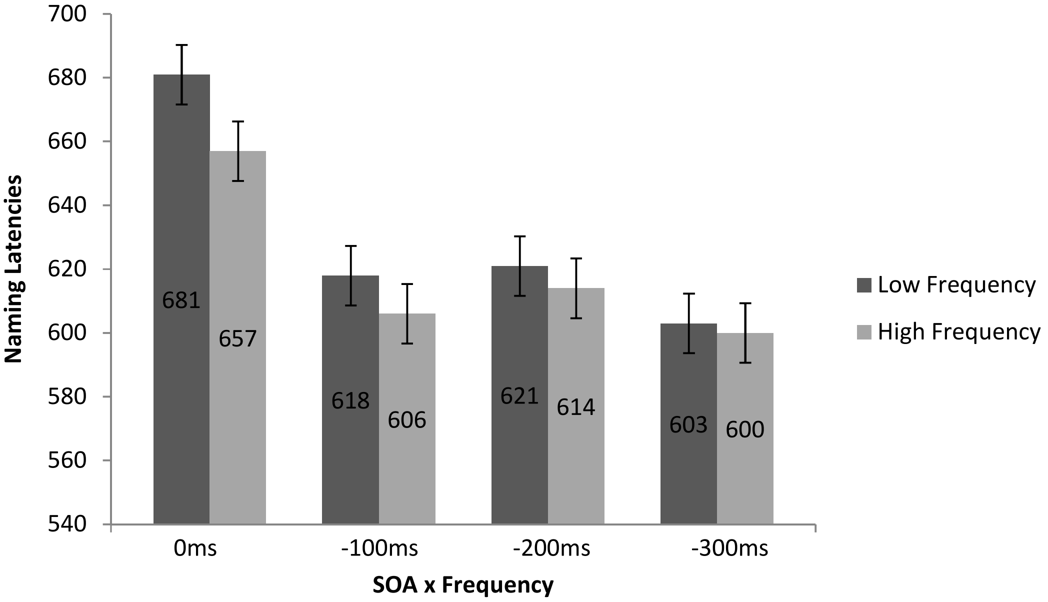

650
640
630
620
610
600
590
580
570
560
Naming Latencies
614
602
622
637
Masked
Visible
Visibility Condition x Semantics
Unrelated
Related

Figure 3.


700
680
660
640
620
600
580
560
540
Naming Latencies
681
657
618
606
621
614
603
600
0ms
-100ms
-200ms
-300ms
SOA x Frequency
Low Frequency
High Frequency