LANGUAGE AND COGNITIVE PROCESSES, 2011, iFirst, 1–22

Psychology Press
Taylor & Francis Group

Time course analysis of the effects of distractor frequency and categorical relatedness in picture naming: An evaluation of the response exclusion account

Peter A. Starreveld[1], Wido La Heij[2], and Rinus Verdonschot[2]

[1]Brain and Cognition, University of Amsterdam, Amsterdam, The Netherlands
[2]Cognitive Psychology Unit, Leiden University, Leiden, The Netherlands

The response exclusion account (REA), advanced by Mahon and colleagues, localises the distractor frequency effect and the semantic interference effect in picture naming at the level of the response output buffer. We derive four predictions from the REA: (1) the size of the distractor frequency effect should be identical to the frequency effect obtained when distractor words are read aloud, (2) the distractor frequency effect should not change in size when stimulus-onset asynchrony (SOA) is manipulated, (3) the interference effect induced by a distractor word (as measured from a nonword control distractor) should increase in size with increasing SOA, and (4) the word frequency effect and the semantic interference effect should be additive. The results of the picture-naming task in Experiment 1 and the word-reading task in Experiment 2 refute all four predictions. We discuss a tentative account of the findings obtained within a traditional selection-by-competition model in which both context effects are localised at the level of lexical selection.

Keywords: Word production; Lexical selection by competition; Picture-word task; Distractor frequency effect; Semantic interference effect; Categorical relatedness.

All models of single-word production assume that during picture naming not only the name of the target picture is activated, but also the names of semantically related competitors. The present study focuses on whether the selection of the target word from this set of activated words is a competitive process (i.e., target selection is hampered by the presence of competitors) or a noncompetitive process (i.e., the highest activated word is selected).

In the literature, examples of both types of model can be found. A word production model that assumes lexical selection without competition is the spreading activation theory of word retrieval in sentence production proposed by Dell (1986). In this

---

Correspondence should be addressed to Peter A. Starreveld, Brain and Cognition, University of Amsterdam, Roetersstraat 15, 1018 WB, Amsterdam, The Netherlands. E-mail: P.A.Starreveld@uva.nl

Peter A. Starreveld, Brain and Cognition, University of Amsterdam; Wido La Heij, Cognitive Psychology Unit, Leiden University; Rinus Verdonschot, Cognitive Psychology Unit, Leiden University, and Leiden Institute for Brain and Cognition, Leiden University, and Centre for Linguistics, Leiden University.

We thank Wieger Manschot and Bart Rossmark for help in collecting the data.

http://www.psypress.com/lcp http://dx.doi.org/10.1080/01690965.2011.608026

theory, target representations are selected after a fixed number of time steps. At the time of selection, the highest activated representation is selected. Note that one of the major aims of the model was to account for error occurrence. As a consequence of the fact that the model selects a target after a fixed number of time steps, it is unable to simulate reaction times. Models that do simulate reaction times (e.g., Levelt, Roelofs, & Meyer, 1999) generally assume lexical selection by competition. In such models target selection occurs when the activation level of the target representation differs from the activation levels of competitors by some critical amount.

In a broader perspective, a mechanism of lexical selection by competition is not only assumed to be involved in word production, but also in many other types of language processing. Rahman and Melinger (2009) recently discussed its role in models of word recognition, sentence processing, lexical ambiguity resolution, and bilingual language processing. In addition, they argued that selection by competition can also be found in other cognitive domains. These observations are important because they show a possible unification of explanations within the myriad of models in cognitive psychology.

In a monolingual setting there are at least three lines of evidence that indicate that lexical selection is indeed by competition. The first line of evidence is based on the observation that pictures that are named in blocks that just contain other pictures from the same category (e.g., all vehicles) are named more slowly than pictures that are named in blocks that contain pictures from various other categories (e.g., Damian & Als, 2005; Damian, Vigliocco, & Levelt, 2001; Kroll & Stewart, 1994). This semantic blocking effect is explained by assuming that selection of the picture name in a semantically related block is hampered by the increased availability of semantically related distractor names. We will not pursue this approach here but instead focus on the other two lines of evidence.

The second line of evidence concerns findings obtained with the picture-word (PW) task. In the standard PW task, participants are asked to name a target picture and to ignore an accompanying distractor word. When the words are categorically related to the pictures (e.g., the picture of a cat accompanied by the word pig), they induce more interference than unrelated words (e.g., the picture of a cat accompanied by the word pin), the categorical interference effect (e.g., Glaser & Düngelhoff, 1984; Rosinski, 1977; Starreveld & La Heij, 1995, 1996; Underwood, 1976). This categorical interference effect is often explained as follows. As a result of picture processing, the lexical representation of categorically related distractor words receive additional activation from the processing of the target, a process of "reversed priming" (priming of the distractor due to target processing). Because reverse priming is absent in the case of unrelated distractors, semantically related distractors form stronger competitors for target selection than unrelated distractors.

The third line of evidence concerns the finding that many speech errors are semantically related to the intended target (e.g., Dell & Reich, 1981; Fromkin, 1973; Meringer & Mayer, 1895). This evidence strongly suggests that semantically related words are active during normal word production. It can be argued that the picture-word task opens the possibility to boost the activation level of such competitors by actually presenting them. Indeed, Starreveld and La Heij (1999) found that when response speed is stressed, errors in the picture-word task mainly consist of distractor errors, that is, participants produced the distractor word instead of the picture's name. An interesting feature of the distractor errors studied by Starreveld and La Heij (1999) was that they showed a categorical effect: distractor errors were much more frequent for the conditions in which the distractors were categorically related to the name of the pictures than when they were unrelated.

It is this pattern shown by these latter two lines of evidence, categorically related words cause more distractor errors in speeded naming and more interference in normal naming (as compared to unrelated words) that suggests that the same mechanism, lexical selection by competition, lies at the heart of both phenomena. Under speeded naming instructions, the mechanism causes semantic errors because at the time of target selection the competition is not yet resolved. However, under nonspeeded naming instructions, the same mechanism avoids the error of producing the distractor word instead of the picture's name. In general, a mechanism to avoid errors seems very worthwhile in normal word production. Note that the situation that is studied in the picture-word task (an externally provided word is available as competitor to an intended target) also occurs in many situations in normal life, for example when one is reading and talking at the same time, when one is interrupting—or is interrupted by—another speaker, or when one is having a conversation at a cocktail party (e.g., Harley, 1990; MacKay, 1987).

Recently, the lexical selection by competition account has been challenged. Miozzo and Caramazza (2003) reported several experiments in which pictures that were accompanied by unrelated distractors of high frequency (HF) were named faster than pictures that were accompanied by unrelated distractors of low frequency (LF), a distractor frequency effect. Miozzo and Caramazza (2003) argued that this distractor frequency effect was hard to account for by models of lexical selection by competition. Most of these models assume that the language frequency of a distractor word is reflected in its resting level of activation. The representations of words of HF would possess a higher resting level than the representations of words of LF (see, e.g., McClelland & Rumelhart, 1981, for a—competition based—word recognition model that incorporates this assumption). All other things being equal, such an account predicts that during processing in the picture-word task, representations of HF words reach a higher activation level than representations of LF words and, as a result, models based on lexical competition that adhere to this activation level hypothesis, predict that HF words should interfere more with picture naming than LF words. This prediction was refuted by the results of Miozzo and Caramazza (2003) and the corresponding activation level hypothesis therefore seems unlikely.

The findings of Miozzo and Caramazza (2003) (see also Catling, Dent, Johnston, & Balding, 2010) have been used as strong support for the model of Mahon, Costa, Peterson, Vargas, and Caramazza (2007). Mahon et al. (2007) assumed that lexical selection occurs without competition; it only involves the selection of the lexical representation with the highest activation level. Mahon et al. (2007) explained the interference induced by word distractors in the picture-word task by assuming that the picture naming process involves a single channel output buffer to which printed (and aurally presented) words have privileged access. When a distractor word is presented in a picture-word task, its production-ready representation automatically occupies the output buffer and has to be excluded from the buffer to allow the production-ready representation of the target name to take its place. We term this account the "response exclusion account" (REA). Within this framework, Mahon et al. (2007) accounted for the distractor frequency effect as follows: When all other word-properties are equal, HF and LF words take the same time to be excluded from the response buffer. The critical assumption to explain the distractor frequency effect is that production-ready representations of HF words enter the buffer *faster* and therefore can be excluded *earlier* than those of LF words. In other words, the size of the distractor frequency effect in picture naming is determined by the difference in time at which the HF and LF distractor words entered the output buffer. This process is illustrated in Figure 1a.

Because Mahon et al. (2007) assumed that lexical selection is not by competition, these authors were faced with the challenge to present an alternative account of

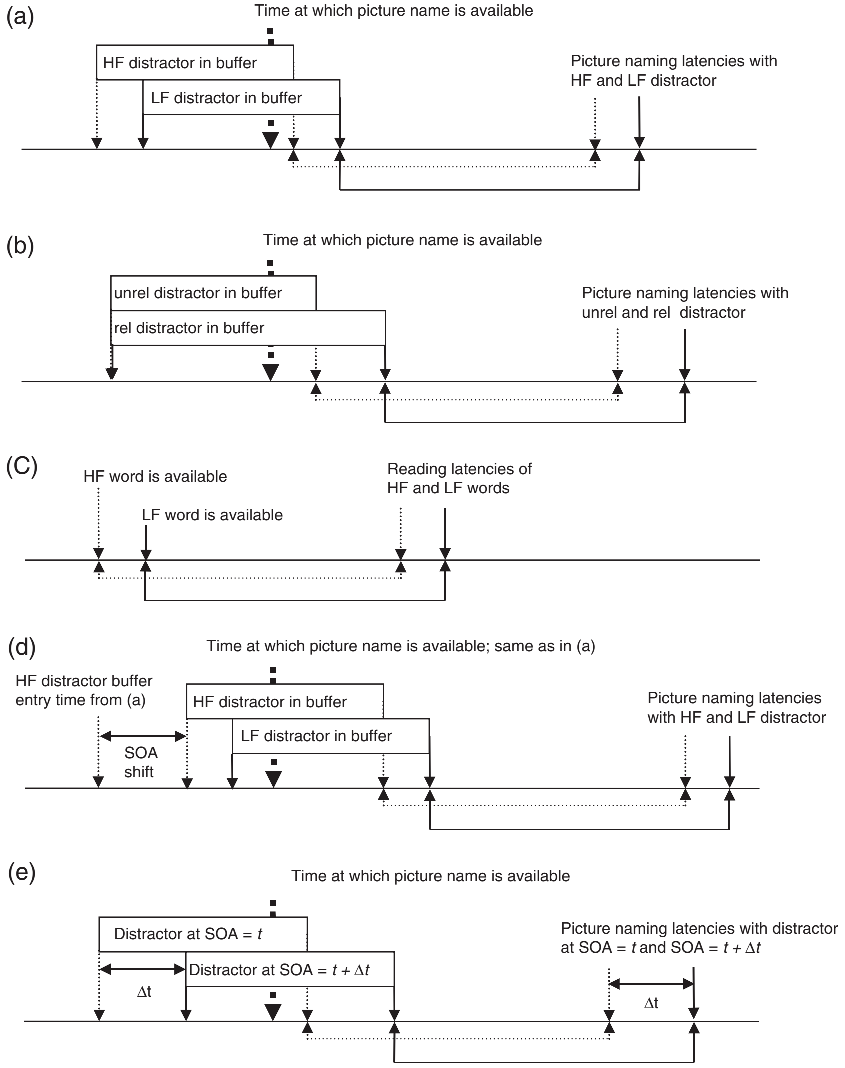

Figure 1. Hypothetical time course of picture and word processing according to the REA; time flows from left to right. The time at which the picture name is available is indicated by a bold dotted arrow. (a) The distractor frequency effect in picture naming; a production-ready representation of an HF distractor enters the buffer faster than that of an LF distractor. (b) The categorical relatedness effect in picture naming; a production-ready representation of a categorically related distractor remains in the buffer longer than that of an unrelated distractor. (c) The distractor frequency effect in reading aloud, the time at which production-ready representations of HF and LF distractors are available are identical to the buffer entering times in (a). (d) The distractor frequency effect in picture naming at a later SOA relative to (a); both HF and LF distractors enter the buffer later. (e) The effect of SOA manipulation on the interference induced by a specific distractor word; at SOA $= t + \Delta t$ buffer entry time for the word is shifted to the right by Δt ms. The length of the boxes represent the time needed to remove a distractor word from the response buffer. HF, high frequency; LF, low frequency; rel, categorically related; unrel, unrelated.

the categorical interference effect. Their account assumes that the effect is postlexical: Mahon et al. proposed that the exclusion mechanism that is able to remove the production-ready representation of the distractor word from the buffer takes more time when the distractor shares a response-relevant criterion with the target. An example of a response-relevant criterion is the broad semantic category of the target (e.g., "tool" or "fruit"). If a distractor shares this criterion with the correct target word, the response exclusion mechanism faces a more difficult challenge, thereby prolonging reaction times (RTs) for pictures accompanied by categorically related words as compared to RTs for pictures accompanied by unrelated words (see Figure 1b). Mahon et al. stress that response-relevant criteria are discrete, a distractor either satisfies a criterion or it does not. Also, it should be noted that, according to the authors, the number of possible response-relevant criteria is indefinite, giving the theory a huge power to account for various context effects.

To summarise, the REA as proposed by Mahon et al. (2007) contains two mechanisms to account for the amount of interference a distractor word produces.[1] The first mechanism is the speed with which a distractor enters the buffer; this speed is governed by properties of the distractor only, like its frequency. The second mechanism is the postlexical exclusion mechanism whose operation is governed by "general properties of the corresponding concepts (e.g., semantic category) as well as their source (picture or word)" (p. 524). In the present study, both mechanisms will be evaluated.

Although the REA explains the distractor frequency effect and the effect of categorical relatedness in the picture-word task, it needs theoretical concepts that lie outside the realm of word production models, like a response buffer and buffer-exclusion mechanisms. In contrast, selection-by-competition accounts attempted to explain observed effects in the picture-word task by the same mechanisms that are found in traditional word production models. Elsewhere (La Heij, Kuipers, & Starreveld, 2006) we argued on theoretical grounds that it might be too early to abandon these traditional accounts. In the present study, we put the REA to a number of empirical tests.

Due to its emphasis on the temporal aspects of distractor processing, the REA makes a number of strong predictions. According to the REA, production-ready representations of a distractor occupy a single channel output buffer and the picture's name has to wait until the distractor is excluded from the buffer. Although the REA differs from traditional "horse-race" models in that written words have a privileged access to the response buffer, in other respects it is very similar to horse-race models that were proposed to account for tasks in which multiple stimuli were presented (see, e.g., Dyer, 1973; Warren, 1972). In the following we derive four predictions from the REA.

First, as discussed above, the REA attributes the distractor frequency effect in the picture-word task to differences between the buffer-access times of HF and LF words (in colloquial terms, the HF horse starts earlier than the LF horse). More specific, the REA assumes that the size of the distractor frequency effect in the picture-word task is *identical* to the difference in buffer access times for HF words and LF words. In the present study, we estimated the latter time difference in the most direct way we could think of: we asked our participants to read the LF and HF distractor

---

[1] According to the REA, a distractor may also prime part of the target processing-pathway. When a distractor is, for example, phonologically related to the target, it primes phonological segments of the target. In the present study we did not investigate such priming effects.

words aloud and measured the response latencies, see Figure 1c. According to the REA, this time difference has a one-to-one relation to the distractor frequency effect in the picture-word task. Therefore, the first prediction we derive from the REA is that the distractor frequency effect in picture naming and in word reading should be of the same size.

Second, in the picture-word task, the relative presentation times of target and distractor can be manipulated by varying the stimulus-onset asynchrony (SOA) between the presentation of the target and the distractor. In this way, the time course of the effects involved can be studied (see, e.g., Glaser & Glaser, 1982, who dismissed a horse race account of the Stroop effect by means of a time course study). Distractors can be presented in advance of the presentation of the picture (negative SOA), simultaneously with (zero SOA), or after the presentation of the picture (positive SOA). According to the REA, the distractor frequency effect occurs exclusively as a result of differences in access times for HF and LF words. Due to this horse-race resembling aspect of the REA, the mechanism that causes the distractor frequency effect is independent from the timing of distractor presentation relative to the time at which the picture is presented. This point is illustrated in Figure 1(a, d), showing that a shift in relative presentation times between distractor word and target picture does not affect the *relative* buffer-entry times for HF and LF words. Therefore, the second prediction that we derive from the REA is that the distractor frequency effect should be independent of the SOA at which the distractor is presented. This prediction is—logically—restricted to the SOA range in which word distractors occupy the response buffer at the time the picture's name becomes available. It can be argued that at large negative SOAs the picture name becomes available after the buffer is already cleared and at large positive SOAs the production-ready representations of the distractor words simply arrive too late to affect target processing. Therefore, no distractor frequency effects are expected at such SOAs.[2] A way to evaluate whether word distractors actually occupied the response buffer at the time the picture's name became available is to test whether they caused more interference than a nonlexical control distractor which, by its very nature, cannot occupy a response buffer. If they do, the REA must assume that the word distractors occupied the response buffer when the picture's name became available.

Third, according to the REA, the amount of interference that a word induces is directly related to the time it enters the response buffer (this explains the distractor frequency effect). As a consequence of this horse-race like aspect, it holds that when the *same* distractors are presented y ms earlier, the interference they induce should be y ms less, and when they are presented y ms later, the interference they induce should be y ms more. In terms of SOA manipulation, the corresponding prediction is that when the same distractors are presented at $\text{SOA} = t$ ms and at $\text{SOA} = t + \Delta t$ ms, the interference effect at the latter SOA should be Δt ms larger than at the former SOA (see Figure 1e). Again, this prediction is restricted to the SOA range in which word distractors occupied the response buffer at the time the picture's name became available.

---

[2] Note that the REA also predicts that a small range of positive SOAs should exist in which a reversal of the polarity of the distractor frequency effect can be found. This should occur at SOAs in which HF distractors have already entered the buffer by the time the picture's name becomes available (thus causing huge interference because the buffer has to be cleared from their representations), but LF distractors arrive too late to influence target processing. If such a polarity shift would be obtained, it would present strong evidence for the REA.

Finally, because the categorical relatedness effect has been central in the development of selection-by-competition accounts, in the present study we investigated both the effects of categorical relatedness, distractor frequency, and their joint influences. According to the REA, the distractor frequency effect is assumed to influence the time at which a distractor is available for exclusion from the response buffer, whereas the categorical relatedness effect is assumed to reflect difficulties posed to the exclusion mechanism based on the distractor sharing response-relevant criteria with the target. Since both effects are linked to independent properties of the distractors and caused by independent mechanisms, it seems reasonable to assume that the effects of distractor frequency and categorical relatedness should be additive.

Miozzo and Caramazza (2003) performed an experiment (Experiment 5) using three SOAs (−100, 0 and 100 ms) in which they crossed the variables frequency of the distractor and categorical relatedness. Therefore, this experiment seems to provide the relevant information to evaluate our second, third, and fourth prediction. However, we believe that Miozzo and Caramazza's (2003) data are inconclusive for two reasons. First, although the distractors from the various conditions were carefully matched, they were not identical in the semantically related and unrelated conditions, leaving the possibility that the results were influenced by properties of the distractor words that were not controlled. Second, Miozzo and Caramazza used a between-participant design, leaving the possibility that the results obtained at the various SOAs were influenced by differences between the groups of participants, for example in general response speed. In order to avoid such potential problems, we used the same word distractor set in the semantically related and unrelated conditions and we used a within-participant design.

In Experiment 1 we put the four predictions we derived from the REA to the test. A PW task was used in which the variables categorical relatedness (of the distractor and the picture) and frequency (of the distractor) were factorially crossed. In addition, to evaluate our predictions that involved the time course of the effects, we varied the SOA in five small steps (−86, −43, 0, 43 and 86 ms). We used small SOA steps to avoid probing into times at which the response buffer might already have been cleared or distractor words had not yet entered it. To evaluate the second and third prediction, we added a control condition in which the distractor was an unpronounceable character string. Finally, after administering the PW task, we asked participants to read the HF and LF distractors words aloud to estimate the relative time at which their production-ready representations became available.

EXPERIMENT 1

Method

Participants

A total of 20 students at Amsterdam University participated for course credit. All had normal or corrected-to-normal vision.

Materials

Twenty-four pictures were selected from the corpus of black and white line drawings made available by Székely et al. (2003). Each picture was used in five distractor conditions: (1) control: the unpronounceable character string "#!/ + &"; (2) HFrel: HF words that were categorically related to the pictures; (3) HFunrel:

HF words that were unrelated to the pictures; (4) LFrel: LF words that were categorically related to the pictures; (5) LFunrel: LF words that were unrelated to the pictures. The HF words and the LF words from the categorically related conditions were paired with other pictures to form the HF unrelated condition and the LF unrelated condition, respectively. Due to a programming error, wrong distractors were presented for two pictures in the LFrel and LFunrel conditions. All data related to these two pictures were removed from all analyses reported below. A list of the remaining target pictures and the distractors appears in the Appendix 1. Distractors were presented in black on a white background.

Frequency counts (per million) for the distractors were taken from the corpus compiled by the Instituut voor Nederlandse Lexicologie and provided by CELEX (Burnage, 1990). The CELEX database contains data obtained from written Dutch documents. The means for the HF and LF conditions were 123.8 ($SD = 52.6$) and 15.7 ($SD = 38.8$), respectively. The difference between these means was significant, $t(38.7) = 7.76$, $p < .001$. Recently, Keuleers, Brysbaert, and New (2010) presented Dutch frequency counts (per million) based on an inventory of films and television series subtitles. The means for the HF and LF conditions calculated from these frequency counts were 90.1 ($SD = 92.3$) and 9.5 ($SD = 12.7$), respectively. The difference between these means was significant, $t(21.8) = 4.1$, $p = .001$. The large *SD* for the HF conditions was caused by the very high frequency of the HF word *car* (458). With this word removed, the mean for the HF conditions was 72.5 ($SD = 43.3$). The difference between the means for the HF and LF conditions remained highly significant, $t(23.3) = 6.4$, $p < .001$.

Distractor words from the HF and LF conditions did not differ in length in letters, $t(42) = -0.82$, $p > .4$; mean scores were 4.8 ($SD = 1.6$) and 5.1 ($SD = 1.3$), respectively. Also, the words in the HF and LF conditions did not differ in imageability (imageability ratings were taken from van Loon-Vervoorn, 1985; ratings were absent for four words used in the LF condition), $t(38) = 0.615$, $p > .5$; mean scores were 6.34 ($SD = 0.67$) and 6.23 ($SD = 0.39$), respectively. In addition, we collected semantic similarity ratings for all picture-distractor combinations used in the experiment. Thirty participants at Leiden University saw each picture-distractor combination on a computer screen, written out as target-distractor word-pairs (e.g., a picture name paired with a distractor from the HFrel condition would be presented as *book—newspaper*). The word pairs were presented individually and in a new random sequence for each participant. Participants were asked to judge the semantic relatedness of each word pair on a scale from 1 to 5 (1 = barely, 2 = a little bit, 3 = average, 4 = quite related, and 5 = strongly related). An independent *t*-test showed that distractors from the HFrel condition ($M = 3.9$, $SD = 0.6$) and the LFrel condition ($M = 3.7$, $SD = 0.7$) did not differ in their semantic similarity to the target, $t(42) = 0.628$, $p = .53$. Also, an independent *t*-test showed that distractors from the HFunrel condition ($M = 1.3$, $SD = 0.2$) and the LFunrel condition ($M = 1.2$, $SD = 0.2$) did not differ in their semantic similarity to the target, $t(42) = 0.799$, $p = .43$.

Note that we did not match the HF and LF distractors on a number of variables like bigram frequency, number of neighbours, initial phoneme, or age of acquisition (see, e.g., Brysbaert & Cortese, 2011, who discuss the importance of some of these variables in the determination of reading aloud latencies). Therefore, we cannot be sure whether an effect of our frequency manipulation was really caused by frequency only or whether, purely coincidently, other variables were involved. Importantly though, all these other variables are, in terms of the REA, associated with the speed in which distractors enter the response buffer and not with response-relevant criteria based on the corresponding concepts (see the introduction for a discussion of these two mechanisms of the REA). Therefore, with respect to our theorising about the REA, the meticulousness of our "frequency" manipulation is not relevant. For smoothness of presentation, we will use the terms frequency effect and frequency manipulation in the remainder of this article.

Apparatus

The experiment was performed using Presentation® software (Version 9.90, www.neurobs.com). Pictures were presented on a fast cathode ray tube monitor running on a refresh rate of 70 Hz. Responses were collected using a voice key and were measured to the nearest millisecond.

Procedure

Participants were seated in a dimly lit room, approximately 50 cm in front of a computer screen. The experiment consisted of three phases. In the training phase, participants were asked to name all of the pictures during four series in which the presentation sequence was randomised. In the first series all pictures were accompanied by their names. In the second series the names were removed, in the third series the names were replaced by the unpronounceable character string, and in the final practice series the names were replaced by distractors that were comparable to those of the experiment proper. During the four practice series, participants named all pictures, errors were corrected and pictures that were named incorrectly were repeated at the end of the series. In the first three series, participants named the pictures at ease. Before the fourth series, participants were asked to name the pictures as fast as possible without making errors for the rest of the experimental session. In addition, they were asked to ignore the accompanying distractors.

In the second phase, five experimental series were presented to the participants with time characteristics corresponding to the five SOAs used. The sequence of the series was counterbalanced according to a Williams (1949) design in which Latin squares are used that are controlled for carry over effects. As a result, the presentation at a particular SOA occurred equally often at each possible position in the sequence. In addition, the number of times that the presentation at a particular SOA was directly preceded by the presentation using each other SOA was equal. The presentation of the materials at each SOA started with five warm up trials, randomly selected from the materials from the last practice series. In order to avoid random noise due to repetition effects, the presentation sequence of the pictures was pseudo-random in such a way that no picture was repeated within the next two trials. After participants made an error, or after a trial in which the voice key malfunctioned, a filler trial was presented, which was randomly selected from the last series of practice trials.

A trial involved the following sequence. First, a fixation sign (+) appeared in the middle of the screen. After 500 ms an empty screen was shown for 200 ms. Subsequently, for negative SOAs, the distractor was shown at the place of fixation and, after a time delay corresponding to the absolute value of the SOA, the picture was added to the display. For positive SOAs, picture and distractor presentation order was reversed and for SOA 0 ms, the picture and the distractor appeared simultaneously. The picture-distractor combination remained on the screen until the participant made a response or 2,500 ms had elapsed. The experimenter then typed in a code into the computer indicating whether the response was correct or false or whether the voice key had malfunctioned (i.e., it triggered a response too early, too late, or not at all). After 500 ms, the next trial began.

In the third and final phase of the experiment, the participants read aloud the distractors that were presented without the pictures. In this phase the unpronounceable character string was replaced by the word "test", these trials were not analyzed. All distractors were presented twice, as they were for each SOA presentation during the previous phase. The sequence of events within a trial was identical to that of the previous phase.

Results

Picture naming

The preparation of the data for the analyses of picture naming times involved the following steps. First, RTs $>2,000$ ms and <300 ms were removed from the analysis; this accounted for 0.09% of the data. Second, RTs of incorrect responses and of trials in which the voice key malfunctioned were removed. This accounted for 1.1% and 2.3% of the data, respectively. Finally, RTs that deviated more than 3 *SD*s from their participant cell means were removed from the analyses; this accounted for 1.4% of the data. The remaining RTs were used in the calculation of the means. Table 1 shows the participant means for all SOAs and all conditions in the experiment; it also shows the corresponding number of errors.

In the introduction, we argued that the REA only accounts for effects in the SOA range in which word distractors occupied the response buffer at the time the picture's name became available. To evaluate whether this was the case in our Experiment 1, we analyzed whether the mean RTs for the word distractor conditions differed from the mean RT for the nonlexical control condition (it can safely be assumed that the unpronounceable string "#!/ + &" that we used in the control condition did not cause a code to enter the response buffer). For each SOA we analyzed simple contrasts

TABLE 1
Participant mean reaction times (in milliseconds) per SOA and per condition and number of errors (in parentheses) for Experiment 1

	SOA				
Condition	*−86*	*−43*	*0*	*43*	*86*
HFrel	653 (1)	669 (10)	704 (4)	663 (10)	602 (12)
HFunrel	652 (7)	659 (5)	688 (7)	655 (9)	597 (8)
LFrel	669 (5)	686 (10)	730 (6)	707 (4)	643 (6)
LFunrel	648 (5)	664 (7)	700 (4)	671 (6)	630 (4)
Control	614 (1)	628 (3)	638 (4)	585 (5)	548 (3)
Rel effect HF	1	10	16	8	5
Rel effect LF	21	22	30	36	13
Freq effect Rel	16	17	26	44	41
Freq effect Unrel	−4	5	12	16	33

Note: SOA, Stimulus Onset Asynchrony; HFrel, high frequency, categorically related; HFunrel, high frequency, categorically unrelated; LFrel, low frequency, categorically related; LFunrel, low frequency, categorically unrelated; Rel effect HF, the categorical interference effect for high frequency words computed as HFrel – HFunrel; Rel effect LF, the categorical interference effect for low frequency words computed as LFrel – LFunrel; Freq effect Rel, the distractor frequency effect for categorically related words, computed as LFrel – HFrel; Freq effect Unrel, the distractor frequency effect for categorically unrelated words, computed as LFunrel – HFunrel.

based on ANOVAs with condition (HFrel, HFunrel, LFrel, LFunrel, and Control) as within subjects variable. The corresponding analysis was also performed using the item means. Results showed that for each SOA the mean RTs in the control condition were faster than those in all other conditions (all $ps < .01$). These results clearly indicate that, in terms of the REA, all distractor words must have occupied the response buffer to cause the *additional* interference.

A $5 \times 2 \times 2$ analysis of variance (ANOVA) was performed on the participant means with SOA, frequency, and categorical relatedness as within participant variables (F_1); a Greenhouse–Geisser correction was used when Mauchly's test of sphericity had a p-value $< 10\%$. The corresponding analysis was also performed using the item means (F_2). The three main effects were significant. First, the effect of SOA was significant, Greenhouse–Geisser-corrected $F_1(2.9,\ 54.3) = 22.3$, $MSE = 5{,}052$, $p < .001$; Greenhouse–Geisser-corrected $F_2(2.9,\ 61.4) = 110.1$, $MSE = 1{,}133$, $p < .001$. Second, the effect of frequency was significant, $F_1(1,\ 19) = 42.8$, $MSE = 988$, $p < .001$; $F_2(1,\ 21) = 17.8$, $MSE = 2{,}306$, $p < .001$. Third, the effect of categorical relatedness was significant, $F_1(1,\ 19) = 22.9$, $MSE = 1{,}143$, $p < .001$; $F_2(1,\ 21) = 15.0$, $MSE = 1{,}690$, $p = .001$.

Interestingly, the interaction of SOA and frequency was significant, $F_1(4,\ 76) = 3.5$, $MSE = 984$, $p = .01$; $F_2(4,\ 84) = 4.8$, $MSE = 675$, $p = .002$. Inspection of Table 1 shows that the frequency effect got larger with increasing SOAs. Also, the interaction of frequency and categorical relatedness was significant, $F_1(1,\ 19) = 6.2$, $MSE = 1{,}090$, $p = .02$; $F_2(1,\ 21) = 5.0$, $MSE = 1{,}472$, $p = .037$. Inspection of Table 1 shows that the categorical relatedness effect was smaller for HF distractors than for LF distractors. The remaining two-way interaction of SOA and categorical relatedness was not significant. Finally, the three-way interaction of SOA, frequency, and semantic relatedness proved non significant.

In the introduction we showed that the REA predicts that when the same distractors are presented at $\text{SOA} = t$ ms and at $\text{SOA} = t + \Delta t$ ms, the interference effect at the latter SOA should be Δt ms larger than at the former SOA. To evaluate this prediction, we calculated for each condition the increases in RTs between SOAs. Because it can be argued that nonlexical properties of the distractors might have influenced the amount of interference through other means than the response buffer (by causing, e.g., warning effects or visual masking effects), for each SOA we first subtracted the RTs obtained in the control condition from the RTs for the word conditions. The resulting increases in RTs between SOAs are shown in Table 2. Because our SOA manipulation consisted of a regular increase of 43 ms, according to the REA each and every cell of Table 2 should resemble that value. Inspection of Table 2 clearly shows that this was not the case. More formally, the REA predicts that each entry should be much the same as 43 ms; whether the actual value is larger or smaller than 43 ms is just determined by chance. Therefore, according to the REA, the chance that *all* 16 entries in Table 2 would be numerically smaller than 43 ms, as was the case in the present experiment, is exactly 2^{-16}. Another way to test the REA's prediction is to evaluate the mean value in Table 2.[3] According to the REA,

---

[3] The non-significant differences between particular values in Table 2 and the test value 43 ms, as presented in Table 2, are not very informative, because the power to detect such differences was very low. For example, given the sample size and the observed sample standard deviation, the power to detect a difference between a sample value of 30 ms and the test value of 43 ms was only 17% for the by-participant analysis. This means that the chance of making a type II error (claiming that there is no difference when in fact there is one) was very high, in this example 83%.

TABLE 2
Observed increases in RT (in milliseconds) between SOAs per condition, relative to the control condition, for Experiment 1

	Observed increases between SOAs			
Condition	*−86 and −43*	*−43 and 0*	*0 and 43*	*43 and 86*
HFrel	$2^{*\#}$	24^{*}	$13^{*\#}$	$-24^{*\#}$
HFunrel	$-6^{*\#}$	$17^{*\#}$	$20^{\#}$	$-21^{*\#}$
LFrel	$4^{*\#}$	33	30	$-27^{*\#}$
LFunrel	$2^{*\#}$	$26^{\diamond}$	$24^{\#}$	$-4^{*\#}$

Note: SOA, Stimulus Onset Asynchrony; HFrel, high frequency, categorically related; HFunrel, high frequency, categorically unrelated; LFrel, low frequency, categorically related; LFunrel, low frequency, categorically unrelated. According to the REA, the expected increase in RTs between SOAs equals 43 ms for each cell, each cell entry was tested against this value with a one-sample *t*-test. Discrepancies between numbers presented in Table 1 and Table 2 are due to rounding.

$^{*}p < .05$ in the by-participant analysis. $^{\diamond}p < .1$ in the by-participant analysis. $^{\#}p < .05$ in the by item analysis.

the mean value should be 43 ms. The observed mean value was 7.1 ms and this value differed from the test value 43 ms, $t_1(19) = -15.7$, $p < .001$; $t_2(21) = -15.6$, $p < .001$. We conclude that this prediction failed completely.

Inspection of Table 1 shows that no speed-accuracy trade-offs were apparent in the data. Very few errors (1.1%) were made during the experiment. Therefore, no formal error analyses were performed.

Distractor reading

In the final phase of the experiment, participants read all distractor words twice. Data treatment was identical to that of the first phase, except that RTs < 200 ms were considered as extremes. The removal of extremes, incorrect responses, trials in which the voice key malfunctioned and RTs that deviated more than 3 *SD*s from the participant cell means, accounted for 0.0%, 0.2%, 2.0%, and 1.2% of the data, respectively. The remaining RTs were used in the calculation of the means.

The mean participant reading times were 479 ms ($SD = 54$) for the HF distractors and 477 ($SD = 49$) for the LF distractors. A paired *t*-tests (in the participant analysis) and an independent *t*-test (in the item analysis) showed no effect of distractor frequency, $t_1(19) = 0.86$, $p = .4$; $t_2(42) = 0.03$, $p = .97$. As expected, reading errors were virtually absent and error percentages were not analysed.

Interim discussion

The lack of a frequency effect in the distractor reading times was not expected but is, nevertheless, very problematic for the REA. A central assumption of the REA is that the distractor frequency effect in the picture-word task is caused by differences in response-buffer entering times between HF and LF distractors. If HF distractors enter the response buffer faster than LF distractors, a frequency effect in reading aloud should also be obtained, but we did not observe it for our materials. To establish this finding more strongly, in Experiment 2 we replicated the second, reading, phase of Experiment 1. In addition, Experiment 2 allowed us to examine an alternative interpretation of the lack of a frequency effect in the reading times obtained in Experiment 1. It is possible that we did not obtain a frequency effect in those reading times because the participants had already read the corresponding distractors silently several times during the first phase of the experiment. If the distractor frequency effect diminishes during repeated presentation of the distractors (but see Miozzo & Caramazza, 2003, who claim that repeated recognition does not cause diminishing interference), it could be the case that the frequency effect was present during the first phase of the experiment but absent in the second phase.

In Experiment 2, we also included a delayed-reading condition to test whether the lack of a frequency effect in reading was related to possible differences between HF and LF distractors in their ability to trigger the voice key. If the LF distractors were able to trigger the voice key faster than the HF distractors, a frequency effect in the reading aloud condition might have been cancelled out. However, in that case, a reversed frequency effect should arise in the delayed-reading condition.

EXPERIMENT 2

Method

Participants

A total of 20 Amsterdam University students participated in the normal-reading part. A total of 30 Leiden University students participated in the delayed-reading part. All had normal or corrected-to-normal vision and participated for course credit.

Materials and apparatus

The same materials and apparatus were used as in the final phase of Experiment 1. The delayed-reading part of the experiment was performed using E-prime software (version 2.0).

Procedure

For the normal-reading part of the experiment, the procedure was identical to the one used for the final phase of Experiment 1. For the delayed-reading part, the procedure was slightly adjusted, as follows. Upon presentation, the distractors remained on the screen for 750 ms. Next, an empty screen was shown, until, after a random interval ranging from 750 to 1,250 ms, a cue was presented. Participants were instructed to withhold their response until the cue appeared and were instructed to respond as fast as possible upon presentation of the cue, while avoiding errors.

Results

Normal reading

Data treatment was identical to that used in the analyses of the data from the final, reading phase of Experiment 1. The removal of extremes, incorrect responses, trials in which the voice key malfunctioned and RTs that deviated more than 3 *SD*s from the participant cell means, accounted for 0.0%, 0.0%, 0.4%, and 0.7% of the data, respectively. The remaining RTs were used in the calculation of the means.

The mean participant reading times were 473 ms ($SD = 65$) for the HF words and 478 ($SD = 68$) for the LF words. A paired *t*-tests (in the participant analysis) and an independent *t*-test (in the item analysis) showed no effect of distractor frequency, $t_1(19) = -0.87$, $p = .4$; $t_2(42) = -0.62$, $p = .54$. As expected, reading errors were virtually absent and error percentages were not analysed.

In the present experiment, participants read all distractor words twice, so the analysis above involved the mean of the RTs of two presentations of each distractor. To evaluate whether an effect of frequency was present during the first presentation only, we also performed a paired *t*-tests (in the participant analysis) and an independent *t*-test (in the item analysis) on the data of the first presentation. The mean participant reading times for the first presentation of the distractors were 479 ms ($SD = 71$) for the HF words and 486 ($SD = 72$) for the LF words. This numerically small frequency effect for the first presentation of the distractors was not significant, $t_1(19) = -1.2$, $p = .24$; $t_2(42) = -0.63$, $p = .53$.

Delayed reading

Data treatment was identical to that used in the analyses of the data from the normal-reading part. The removal of extremes, incorrect responses, trials in which the voice key malfunctioned and RTs that deviated more than 3 *SD*s from the participant cell means, accounted for 3.5%, 0.5%, 0.8%, and 1.2% of the data, respectively. The remaining RTs were used in the calculation of the means.

The mean participant reading times were 340 ms ($SD = 65$) for the HF words and 336 ($SD = 60$) for the LF words. A paired *t*-test (in the participant analysis) and an independent *t*-test (in the item analysis) showed no effect of distractor frequency, $t_1(29) = 1.12$, $p = .27$; $t_2(42) = 0.78$, $p = .44$. As expected, delayed-reading errors were virtually absent and error percentages were not analysed.

Interim discussion

In Experiment 2 we replicated the result of the second phase of Experiment 1. For a second time, we did not obtain an effect of distractor frequency in distractor reading times, now in a setting in which participants had not seen the distractors before. Therefore, this finding clearly shows that the lack of a frequency effect in distractor reading times obtained in the final phase of Experiment 1 is a robust finding which can not be attributed to the repeated presentation of the distractors during the first phase of that experiment. In addition, we did not obtain an effect of distractor frequency in distractor reading times in a delayed-reading task in which the participants could prepare their responses in advance of a cue. This result clearly shows that the lack of a frequency effect in reading times obtained in the final phase of Experiment 1 and in the normal-reading condition of Experiment 2 cannot be attributed to possible differences between HF and LF distractors in their ability to trigger the voice key. Next, we first discuss the implications of this finding.

GENERAL DISCUSSION

In the introduction we derived four predictions from the REA as proposed by Mahon et al. (2007). The first prediction was that the size of the distractor frequency effect obtained in the PW task should be similar to the size of the frequency effect obtained in the word-reading task. Although we did obtain clear-cut distractor frequency effects in the PW task (with a maximum of 44 ms with semantically related words and a maximum of 33 ms with unrelated words; see Table 1), we failed to find a frequency effect when the same participants read the distractor words aloud. This result is problematic for the response exclusion hypothesis because it indicates that the amount of interference induced by distractor words is unrelated to the ease with which a production-ready representation can be produced, which is a core assumption of the REA. In our Experiment 1, participants read the distractors after performing the experiment proper, so one might argue that the repeated exposure to the distractors might have caused the lack of a covert frequency effect. In order to test that hypothesis, in Experiment 2 we had a new group of participants who only read the HF and LF distractor words, but again no frequency effect was obtained. We also included a delayed-reading task to determine whether the HF or LF distractors differed in their ability to trigger the voice key but they did not. If one makes the reasonable assumption that the reading times present a direct measure of the time at which a production-ready representation is available in an output buffer, these results clearly undermine the heart of the theoretical rationale of the frequency account presented by both Miozzo and Caramazza (2003) and Mahon et al. (2007).

One could argue that a cause for our failure to obtain a frequency effect in reading times lies in the possibility that in our reading aloud task the distractor words were read sublexically, whereas in our picture-word task the distractors were read through a lexical route. However, there seems to be no a priori reason to assume that reading in the two tasks involved different processes. In addition, Mahon et al. (2007) assume that the privileged access to the articulators that printed words have occurs because “word reading benefits from the quasi rule-like relationships between orthography and phonology” (p. 524). Thus, Mahon et al. assume that a sublexical route is involved in the processing of distractor words in the picture-word task also, so the involvement of this route cannot be used to explain the lack of a frequency effect in reading aloud.

Our failure to find a frequency effect in reading times accords with other findings in the literature. Although a number of studies have reported a language frequency effect in word reading, detailed inspection of the relevant studies reveals that the effect is more elusive than is often assumed (see, for a similar conclusion, McCann & Besner, 1987). The effect was reported by Forster and Chambers (1973), who presented 15 HF (AA category: 100 or over per million) and 15 LF (1–3 per million) words in a series of trials in which nonwords were also presented and reported mean word-reading latencies of 508 and 579 ms, respectively. Frederiksen and Kroll (1976) presented words from four frequency classes (counts per million: ≥ 30, 6–29, 2–5, and ≤ 1) and obtained reading latencies (estimated from their Figure 1) of 525, 531, 545, and 561 ms, respectively. This finding indicated that a substantial proportion of the effect was due to the inclusion of words of (very) LF. Finally, McCann and Besner (1987) reported a significant correlation between language frequency (ranging between 1 and 1,000 occurrences per million) and word-reading latency.

However, in quite a number of studies the frequency effect was nonsignificant. Scarborough, Cortese, and Scarborough (1977; Experiment 3) failed to find a significant difference in reading latencies between LF (3 per million) and HF words (76 per million) and Richardson (1976) did not obtain a significant correlation between frequency and word-reading latencies. Moreover, in a number of experiments Waters and Seidenberg (1985) obtained a word frequency effect in reading irregular English words but not in reading regular words. Only in their Experiment 6, in which the frequency range was large (1,059 and 3.8 in the HF and LF conditions, respectively) a small but significant language frequency effect of 16 ms was observed.

Given these findings and the fact that (1) Dutch has a shallow orthography and (2) the range of frequencies used in the present study was rather limited (mean counts of 16 and 124 in the LF and HF conditions, respectively), the lack of a frequency effect in our word-reading tasks of Experiments 1 and 2 is not that surprising. Still, the fact that these words when used as distractors in the PW task did produce a distractor frequency effect is quite problematic for the REA.

The second prediction we derived from the REA was that when a frequency effect is obtained, it should be constant across SOAs. Our results showed otherwise, for example, in the SOA range of 0, +43, and +86 ms, the frequency effects (obtained with unrelated distracter words) were 12, 16, and 33 ms, respectively. This result seems hard to account for by the REA. In order to explain the obtained interaction, it would be necessary to assume that word processing time is dependent on SOA in such a way that when a word is presented earlier, the difference between the processing times of HF and LF words get smaller. It seems hard to defend such an assumption. Note that Miozzo and Caramazza (2003) showed that the ease with which a word can be recognised did not affect the size of the distractor frequency effect, indicating that word perception processes do not play a significant role in the causation of the distractor frequency effect. Instead, Miozzo and Caramazza (2003) argued that the ease with which a word can be produced affects the amount of interference obtained in a picture-word task.

The third prediction we derived from the REA was that an increase in SOA (later presentation of the distractor word) should result in an equivalent increase in the interference induced by that distractor word. For instance, an increase of SOA from 43 to 86 ms should—in comparison to the nonword control condition—result in an increase of interference induced by a distractor word of about 43 ms. This is not what we obtained, for example, interference induced by a distractor word even decreased when SOA increased from 43 to 86 ms. Similar results were obtained by Starreveld and La Heij (1996), who reported that unrelated distractor words induced 85 ms interference at SOA = 0 ms and 79 ms interference at SOA = + 100 ms.

The fourth prediction we derived from the REA was that the effects of categorical relatedness and distractor frequency should be additive. However, we obtained a clear interaction between the effects of categorical relatedness and distractor frequency. The experiment showed that the categorical relatedness effect was smaller for HF words than for LF words. This result contrasts with the results reported by Miozzo and Caramazza (2003) who obtained evidence for additivity of the two effects. As discussed in the introduction, this discrepancy might be due to a difference in design: whereas Miozzo and Caramazza (2003) used a between-participants design we used a within-participants design. This design difference has also been put forward as a possible cause of differences in the results of time course studies involving the effects of semantic and phonological relatedness (Starreveld, 2000). Another difference in the design is that Miozzo and Caramazza (2003) incorporated many filler trials whereas we only used filler trials after participants made an error or the voice key malfunctioned. Finally, although Miozzo and Caramazza carefully matched the distractors in the four conditions involved, they used different words in the categorically related conditions and in the unrelated conditions, which allows for the possibility that differences between these two conditions are due to uncontrolled variables. In contrast, in our experiment, the same words were used in the categorically related conditions and in the unrelated conditions, excluding that possibility. Given that we obtained a clear interaction between categorical relatedness and frequency, we conclude, following the same logic additive factors logic (Sternberg, 1969) that Miozzo and Caramazza did, that both effects are most probably localised at the same level of processing.

The finding of an interaction between categorical relatedness and frequency seems problematic for the response exclusion hypothesis (Mahon et al., 2007). According to this hypothesis, the categorical relatedness effect arises because "production-ready representations corresponding to unrelated distractor words can be excluded faster than representations corresponding to distractors that satisfy a response criterion demanded by the target pictures" (p. 523). Thus, this explanation is based on a specific property of the distractor (whether it shares a response criterion with the target or not). As discussed above, the distractor frequency effect was argued to arise because HF distractors are available for exclusion from the response buffer earlier than LF distractors. Thus, this explanation is based on a different property of the distractor than the one used to explain the categorical relatedness effect. Consequently, the two corresponding effects are expected to show additivity.

Elsewhere (La Heij et al., 2006), we argued that the activation-based account should not so easily be disposed with, because it presents a straightforward explanation of many experimental findings, among which are the time courses of categorical interference, phonological facilitation, and their interaction (Starreveld & La Heij, 1995, 1996; see also Damian & Martin, 1999, and Bonin & Fayol, 2000), the reversal from semantic interference into facilitation as a result of varying SOAs and distractor modality (Bloem & La Heij, 2003; Bloem, van den Boogaard, & La Heij, 2004), and the distractor-priming effect (e.g., Starreveld & La Heij, 1996). In addition, the activation account is backed up by computational models that are able to simulate the empirical effects based on the theoretical assumptions (Bloem et al., 2004; Starreveld & La Heij, 1996). How about the present results? Can activation-based models account for the four observations discussed above? In our view there are at least two options to pursue.

A first option to account for the distractor frequency effect in terms of activation-based models is to assume that an attentional mechanism is able to block out distractors using production rules (Roelofs, 2005). It is quite possible that by adding other production rules Roelofs' activation-based model is also able to account for the results presented here. Elsewhere (La Heij, Starreveld, & Kuipers, 2007), we argued against the use of production rules to solve *hard problems*, so we leave it to the developers of Weaver + + to account for the present data in terms of their model.

A second option to account for the distractor frequency effect in terms of activation-based models might be to build on a different account of frequency effects in word perception. To account for such frequency effects, activation-based accounts try to implement the assumption that less evidence is needed to recognise an HF word than an LF word. One way to accomplish this is to assume that frequency effects in word recognition are caused by differences in resting level activations, as discussed by Miozzo and Caramazza (2003). However, another way to accomplish this is to assume that frequency effects in word recognition are caused by differential thresholds for recognition. Although these accounts resemble each other very much and, in fact, Norris (2006) treats both accounts of word frequency as functionally equivalent, an account in terms of different thresholds might provide a possible explanation of distractor frequency effects in the picture-word task. If HF words have lower thresholds for recognition than LF words, representations of HF words will reach *lower* levels of activation than those of LF words. Therefore, representations of HF distractors will be weaker competitors to the target than those of LF distractors, causing the distractor frequency effect to emerge. The concept of differential thresholds has been around in the literature for many years, it was an integral part of the original logogen model of word perception (Morton, 1969), but it can also be found in models of word production (Jescheniak & Levelt, 1994).

An additional assumption, common to many activation-based models, is that upon recognition, the activation of the distractor word representation decays to its resting level. It is commonly assumed that the decay of a representation is proportional to its level of activation (exponential decay; e.g., Bloem, van den Boogaard, & La Heij, 2004; Levelt et al., 1999; McClelland & Rumelhart, 1981; Starreveld & La Heij, 1996). Based on these two assumptions the course of activation for the representations of HF and LF words can be derived (see Figure 2). Note that the resting levels for HF and LF words are identical in Figure 2. We call this account the differential-threshold account (DTA). The four main findings from the present study might be accommodated within this framework as follows.

First, in contrast to the REA, the DTA does not predict a one-to-one relation between the size of the frequency effect in word reading and the size of the corresponding distractor frequency effect in picture naming. As illustrated in Figure 2, different recognition thresholds for HF and LF words might have a very small effect on word recognition times (left, rising part of the activation curve in Figure 2), but may have a substantial effect on the amount of interference induced on target-name selection (right, decaying part of the activation curve in Figure 2).

Second, the DTA has little problem explaining our finding that the distractor frequency effect increases with an increase in SOA (in the SOA conditions −86, −43, 0, +43, and +86 ms, the average frequency effects observed were 6, 11, 19, 30, and 37 ms, respectively). This can again be illustrated with the help of Figure 2. The difference between HF and LF distractor words will be maximal when their peak activation levels coincide with the moment of target-name selection. This situation most probably occurs with a small postexposure of the distractor word. When the SOA decreases (earlier presentation of the distractor words), the difference in activation level between HF and LF words decreases, resulting in smaller distractor frequency effects; exactly what was obtained empirically.

Third, the DTA does not predict that when the same distractors are presented at SOA $= t$ ms and at SOA $= t + \Delta t$ ms, the interference effect at the latter SOA should be Δt ms larger than at the former SOA. Instead, the amount of predicted interference is a result of the competition between representations that are activated through word perception and word production processes. Relative to the situation at

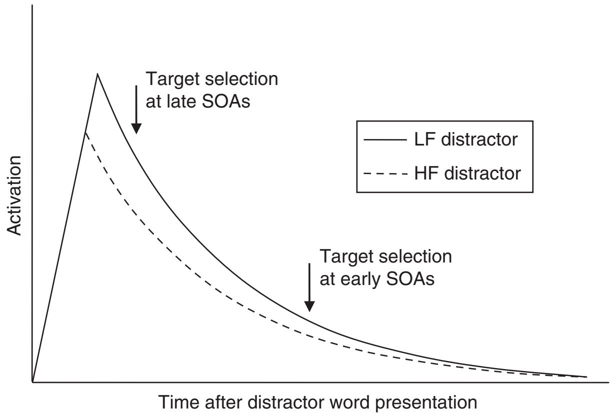

Figure 2. Hypothetical course of activation of the representations of HF distractors and LF distractors after visual encoding. Activation decays upon recognition. Arrows indicate the time of picture-name selection at late SOAs (at which the picture is presented first), and at early SOAs (at which the distractor is presented first).

SOA 0 ms, the interference induced by words presented at large negative SOAs is reduced as a result of decay of distractor representations. At large positive SOAs it is reduced because the target representation is already highly activated at the moment the distractors appear.

Finally, the DTA might explain the interaction of frequency and semantic relatedness along the following lines. As picture naming is faster in case of an HF distractor, less reverse priming (additional activation that reaches the node of a distractor as a result of picture processing) can reach the word representation of an HF word as compared to that of an LF word. As a result, an interaction of frequency and categorical relatedness might surface. This explanation is analogous to that of the interaction of the effects of categorical and phonological relatedness (see Starreveld & La Heij, 1996, for details and a computational model).

An activation-based explanation of distractor frequency effects in the picture-word task thus seems possible. However, we realise that it needs more empirical support. For example, the DTA predicts an interaction of the effects of distractor frequency and phonological relatedness between target and distractor, because both variables affect the same response-selection process. Indeed, with the concurrent presentation of target and distractor, Miozzo and Caramazza (2003) reported such an interaction, but a replication of this finding using a time course study would establish it more firmly. Also, the DTA would certainly gain strength by a computational implementation showing the validity of the arguments. It should be noted though, that the general assumptions made by activation-based models are backed up by computational implementations (e.g., Bloem et al., 2004; Levelt et al., 1999; Starreveld & La Heij, 1996) whereas computational implementations of the response exclusion hypothesis are, at present, lacking.

In conclusion, although activation-based models that encompass lexical selection by competition have their problems too (for discussion, see Finkbeiner & Caramazza, 2006; La Heij et al., 2006) the present results showed that the REA makes incorrect predictions regarding four important aspects of interference induced by HF and LF distractor words. The failure of the REA to account for our present findings can be attributed to its horse-race resembling assumptions about distractor processing and interference. Since these are core assumptions of the REA, we regard the account as seriously challenged.

REFERENCES

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.

Bonin, P., & Fayol, M. (2000). Writing words from pictures: What representations are activated, and when? *Memory & Cognition*, *28*, 677–689.

Brysbaert, M., & Cortese, M. J. (2011). Do the effects of subjective frequency and age of acquisition survive better word frequency norms? *The Quarterly Journal of Experimental Psychology*, *64*, 545–559.

Burnage, G. (1990). *CELEX – A guide for users.* Nijmegen, The Netherlands: CELEX.

Catling, J. C., Dent, K., Johnston, R. A., & Balding, R. (2010). Age of acquisition, word frequency, and picture–word interference. *The Quarterly Journal of Experimental Psychology*, *63*, 1304–1317.

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, and Cognition*, *31*, 1372–1384.

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, G., & 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.

Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, *20*, 611–629.

Dyer, F. N. (1973). The Stroop phenomenon and its use in the study of perceptual, cognitive, and response processes. *Memory and Cognition*, *1*, 106–120.

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

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, *12*, 627–635.

Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance*, *2*, 361–379.

Fromkin, V. A. (Ed.). (1973). *Speech errors as linguistic evidence*. The Hague: Mouton.

Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 875–894.

Glaser, W. R., & Düngelhoff, F. J. (1984). The time course of picture–word interference. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 640–654.

Harley, T. A. (1990). Environmental contaminations of normal speech. *Applied Psycholoinguistics*, *11*, 45–72.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word-frequency effects in speech production: Retrieval of syntactic information and phonological form. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *20*, 824–843.

Keuleers, E., Brysbaert, B., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*, 643–650.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–174.

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

La Heij, W., Starreveld, P. A., & Kuipers, J.-R. (2007). Structural complexity is not the (big) issue: A reply to Roelofs (2007). *Language and Cognitive Processes*, *22*, 1261–1280.

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

MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills.* New York: Springer.

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

McCann, R. S., & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 14–24.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception 1: An account of basic findings. *Psychological Review*, *88*, 375–407.

Meringer, R., & Mayer, K. (1895). *Versprechen und Verlesen: Eine psychologisch-linguistische Studie*. Stuttgart, Germany: Göschense Verlagsbuchhandlung. (Reissued, with introductory essay by A. Cutler & D. A. Fay, Amsterdam, The Netherlands: John Benjamins, 1978).

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

Morton, J. (1969). The interaction of information in word recognition. *Psychological Review*, *76*, 165–178.

Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review*, *113*, 327–357.

Rahman, R. A., & Melinger, A. (2009). Semantic context effects in language production: A swinging lexical network proposal and a review. *Language and Cognitive Processes*, *24*, 713–734.

Richardson, J. T. E. (1976). The effects of stimulus attributes upon latency of word recognition. *British Journal of Psychology*, *67*, 315–325.

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

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

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 1–17.

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

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

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

Starreveld, P. A., & La Heij, W. (1999). Word substitution errors in a speeded picture-word task. *American Journal of Psychology*, *112*, 521–553.

Sternberg, S. (1969). Discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, *30*, 276–315.

Székely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., et al. (2003). Timed picture naming: Extended norms and validation against previous studies. *Behavior Research: Methods, Instruments & Computers*, *35*, 621–633.

Underwood, G. (1976). Semantic interference from unattended printed words. *British Journal of Psychology*, *67*, 327–338.

Van Loon-Vervoorn, W. A. (1985). *Voorstelbaarheidswaarden van Nederlandse woorden*. Lisse, The Netherlands: Swets & Zeitlinger.

Warren, R. E. (1972). Stimulus encoding and memory. *Journal of Experimental Psychology*, *94*, 90–100.

Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: Time course and decision criteria. *Memory & Cognition*, *13*, 557–572.

Williams, E. J. (1949). Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Scientific Research*, *2*, 149–168.

APPENDIX 1
Stimulus material used in Experiments 1 and 2

Picture name	*HF categorically related*	*LF categorically related*
Boek (book)	Krant (newspaper)	Folder (flyer)
Ezel (donkey)	Hond (dog)	Cavia (guinea-pig)
Beer (bear)	Paard (horse)	Geit (goat)
Jas (jacket)	Broek (trousers)	Shirt (shirt)
Ober (waiter)	Leraar (teacher)	Chirurg (surgeon)
Vliegtuig (airplane)	Trein (train)	Metro (underground)
Borst (chest)	Rug (back)	Dij (thigh)
Duim (thumb)	Arm (arm)	Been (leg)
Skelet (skeleton)	Hart (heart)	Nier (kidney)
Hand (hand)	Schouder (shoulder)	Enkel (ankle)
Bus (bus)	Auto (car)	Truck (truck)
Cactus (cactus)	Boom (tree)	Riet (reed)
Agent (policeman)	Schrijver (writer)	Bakker (baker)
Zon (sun)	Ster (star)	Komeet (comet)
Dokter (doctor)	Rechter (judge)	Visser (fisherman)
Heks (witch)	Geest (ghost)	Spook (phantom)
Tent (tent)	Gebouw (building)	Woonboot (house-boat)
Zeilboot (sailboat)	Schip (ship)	Kano (canoe)
Koffer (suitcase)	Zak (Sack)	Tasje (purse)
Bever (beaver)	Kat (cat)	Mol (mole)
Boter (butter)	Melk (milk)	Yoghurt (yoghurt)
Woestijn (desert)	Bos (wood)	Jungle (jungle)

Note: Distractor words were presented in Dutch. English translations appear in parentheses. HF, high frequency; LF, low frequency.