

# Utterance Format Affects Phonological Priming in the Picture–Word Task: Implications for Models of Phonological Encoding in Speech Production

Jörg D. Jescheniak

Max Planck Institute of Cognitive Neuroscience and  
University of Potsdam

Herbert Schriefers

Nijmegen University

Ansgar Hantsch

Max Planck Institute of Cognitive Neuroscience

Picture–word experiments investigating the production of multiword utterances with distractors that are phonologically related to words in noninitial position have yielded inconsistent results, ranging from facilitation to inhibition. A comparison of these studies is complicated by differences in detail. In parallel to the empirical inconsistencies, different theoretical accounts of phonological encoding in speech production have been provided. In the present article, the authors propose a unitary account, which can in principle account for facilitation, null effects, and inhibition. It assumes a graded activation pattern of the elements within the scope of phonological advance planning. The account is tested in an experiment varying utterance format while keeping all other aspects constant. The results are consistent with the proposed unitary account.

Recent theories of speech production distinguish four main levels of processing: conceptual preparation, grammatical encoding, phonological encoding, and articulation. In parallel to the distinction of these processing levels, accessing words in the mental lexicon is assumed to comprise two steps. On the one hand, the conceptual input guides access to a lexical lemma representation, that is, a nonphonological representation specifying the word's syntactic (Levelt, Roelofs, & Meyer, 1999) or semantic and syntactic properties (Vigliocco, Vinson, Martin, & Garrett, 1999). On the other hand, there is access to a word's phonological form, which serves as input in the generation of articulatory motor programs (see Garrett, 1988, and Levelt, 1999, for reviews).

Over the past 10 years, experimental evidence for these steps of lexical retrieval in language production has been accumulating from chronometric tasks investigating the production of single words, most notably nouns (e.g., Damian & Martin, 1999; Jescheniak & Levelt, 1994; Jescheniak & Schriefers, 1998; Jescheniak, Schriefers, & Hantsch, 2001; Levelt et al., 1991; Peterson & Savoy, 1998; Roelofs, 1992; Schriefers, Meyer, & Levelt, 1990). More recently, chronometric approaches to language production

have been going beyond the production of single words, focusing on the production of multiword utterances, most notably noun phrases with a definite determiner or an adjective or both (e.g., Costa & Caramazza, 2002; Miozzo & Caramazza, 1999; Schriefers, 1992, 1993; Schriefers & Teruel, 1999), and only occasionally sentences (Jescheniak et al., 2001; Meyer, 1996). Many of these studies have used the picture–word task and variants thereof. It requires participants to name (or describe) pictures of objects (or configurations of objects) while ignoring distractor words. This task was also used in the experiment reported in the present article.

The production of multiword utterances introduces a number of new topics that do not play a role in the production of simple nouns. A first issue concerns the question of whether initiation of articulation of the first word is dependent on any aspect of later words in the eventual utterance. A second and related issue concerns the question of how the retrieval procedures for the different words of a multiword utterance are aligned in time (e.g., Levelt & Meyer, 2000). The present article primarily addresses the former issue.

In the following, we illustrate this issue with reference to the types of utterances studied in the present article, namely, German noun phrases consisting of a bare noun (e.g., “*Hund*”—dog), a gender-marked article and a noun (e.g., “*der Hund*”—the<sub>[masc]</sub> dog), or a gender-marked article, a size adjective, a color adjective, and a noun (e.g., “*der große rote Hund*”—the<sub>[masc]</sub> big red dog).<sup>1</sup> Hereafter, we refer to these three utterance formats as *bare noun*, *simple noun phrase*, and *complex noun phrase*, respectively. In our experiment, these utterances were produced in the context of the picture–word task. Participants named pictures of colored objects presented in different sizes by one of the three utterance formats.

---

Jörg D. Jescheniak, Max Planck Institute of Cognitive Neuroscience, Leipzig, Germany, and Department of Psychology, University of Potsdam, Potsdam, Germany; Herbert Schriefers, Nijmegen Institute for Cognition and Information, Nijmegen University, Nijmegen, the Netherlands; Ansgar Hantsch, Max Planck Institute of Cognitive Neuroscience.

Our research was supported by Grant DFG JE229/2-3 and a Heisenberg Fellowship from the German Research Council awarded to Jörg D. Jescheniak. We thank three anonymous reviewers for helpful comments and suggestions. Many thanks go also to Stefanie Hoffman, Susanne Lamm, and Jutta Müller for their help in data collection.

Correspondence concerning this article should be addressed to Jörg D. Jescheniak, Max Planck Institute of Cognitive Neuroscience, P.O. Box 500355, D-04303 Leipzig, Germany. E-mail: jeschen@cns.mpg.de

---

<sup>1</sup> We use subscripts for indicating the grammatical gender of the German noun determining the form of the article; masc = masculine.

In addition, participants heard auditory distractor words they were instructed to ignore. These distractor words were either semantically related to the noun of the target utterance or phonologically related or unrelated.

Access to the noun lemma is required for all three utterance formats in order to be able to initiate the utterance. For bare nouns this is obvious. For the other two types of utterances, the lemma of the noun provides information about its grammatical gender, which, in turn, determines the initial element of the eventual utterance, that is, the definite article.<sup>2</sup> Put differently, in these utterances retrieval of the initial word is necessarily dependent on retrieval of the lemma of a word occurring only later in the utterance. Lemma retrieval can be tapped by the inhibition effect from semantically related distractors (e.g., Damian & Bowers, in press; Schriefers et al., 1990): Semantically related distractors slow down lemma selection, relative to unrelated distractors. In light of the previous considerations, we would thus expect inhibition from distractors that are semantically related to the noun in our experiments. This effect should obtain for all three utterance formats and should be of the same size.

For phonological encoding, the situation is less clear. In an extrapolation of the WEAVER (Word Encoding by Activation and VERification) model of lexical access (Roelofs, 1997) to multiword utterances, Levelt and Meyer (2000) proposed that the phonological forms of the morphemes or words making up a multiword utterance become available in successive packages, that is, word by word (or morpheme by morpheme) in a left to right manner. A similar view is taken by Dell (1986; see also Dell, Burger, & Svec, 1997). In this model, the first syllable of a multisyllabic word first receives stronger activation while the second syllable receives less activation. When the phonological segments of the first syllable are selected, the main flow of activation is switched to the second syllable. Generalizing this view to multiword utterances, the phonological forms of successive words become activated over time from left to right.

However, this proposal does not specify the point in time at which articulation is actually initiated. Here, there are three possibilities. First, according to a *minimum position*, articulation might be initiated as soon as the first word's phonological form has become available and before any of the other words' phonological forms has received activation. Second, according to a *maximum position*, initiation of articulation is delayed until the complete utterance has become phonologically encoded. Third, according to an *intermediate position*, articulation is initiated as soon as the first word's phonological form has become available, but at that point in time the other words' phonological forms have also received some activation. When we combine this intermediate position with the proposal by Levelt and Meyer (2000) and Dell (1986; Dell et al., 1997), the activation of these other word forms decreases as a monotonic function of their position in the eventual utterance. That is, taking a snapshot of the activation pattern at the moment articulation is about to be initiated, the phonological form of the first word is highest, and the activation of the phonological forms of later words monotonically decreases with their position.<sup>3</sup>

What would be the predictions concerning the effects from distractors phonologically related to the noun, relative to unrelated distractors in the three scenarios? When only the first word of the eventual utterance has become phonologically encoded while the other word forms have not yet received any activation at utterance

onset (minimum position), we should obtain facilitation for bare noun utterances but no effect whatsoever for simple and complex noun phrase utterances. This should be the case because in the latter two cases, the noun's phonological form has not received any activation before utterance onset. Therefore, priming this form by presenting a phonologically related distractor should not have any specific effect on utterance onset latencies.

What about the assumption that phonological encoding is completed for the whole utterance before articulation is initiated (maximum position)? Most likely, a distractor phonologically related to the noun should speed up the retrieval of its phonological form and thus lead to a faster completion of the phonological encoding of the whole utterance. This should in turn become visible as facilitation from phonologically related distractors relative to unrelated distractors, for utterances of any complexity.

Finally, there is the possibility of a graded activation of the words' phonological forms as a function of their position in the utterance (intermediate position). In this case, the phonological form of the noun receives some activation before utterance onset, but it has not yet become selected. A distractor that is phonologically related to the noun adds to this activation. Furthermore, utterance onset depends on the selection of the first word's phonological form. As the phonological form that should *not* be selected at that point in time (i.e., the noun) has accumulated additional activation from the distractor, competition between the noun's word form and the first word's phonological form increases, and thus lead to reduced facilitation or even inhibition relative to a situation in which an unrelated distractor was presented.

Let us now turn to the available experimental evidence and its theoretical interpretation in the literature. Meyer (1996) reported one of the first picture-word studies on the production of multiword utterances. Her participants named two simultaneously presented objects either by a noun phrase coordination (e.g., *the bag and the arrow*) or by simple sentences (e.g., *the bag is next to the arrow*). Semantically related distractors slowed down naming speed, and this was the case for distractors related to the first noun as well as for distractors related to the second noun. Distractors phonologically related to the first noun yielded significant facilitation effects. Most important for the present purpose, distractors phonologically related to the second noun showed a small but significant inhibition effect of about 20 ms at stimulus onset asynchronies (SOAs) of 300 ms and 0 ms in Experiments 3 and 4, respectively. Following Dell (1986), Meyer interpreted this result as follows. Words are encoded phonologically according to their order in the utterance, which is specified during grammatical encoding. The phonological form of the first noun is initially activated more strongly than the phonological form of the second noun. When the distractor is phonologically related to the second

<sup>2</sup> Recent models of the production lexicon consider grammatical gender as a syntactic property stored with the lemma of the noun (see Schriefers & Jescheniak, 1999, for a review). Hence, accessing a noun's gender requires accessing its lemma.

<sup>3</sup> A similar proposal has been put forward by Dell and O'Seaghdha (1992), who assumed that the forms of words in different sentence positions may be prepared to different degrees or in different ways before utterance onset.

noun, the activation of the phonological form of the second noun is increased. Thus, the phonological form of the second noun becomes a strong competitor to the phonological form of the first noun, that is, the phonological form that has to be actually selected to initiate articulation. This competition becomes visible as an increase in naming latency. Thus, these results and their interpretation are in line with the notion of a graded activation of the phonological forms of successive words as a function of their linear position.

Schriefers and Teruel (1999) studied the production of German noun phrases consisting of a prenominal adjective and a noun in a picture–word experiment. They observed facilitation from distractors phonologically related with the beginning of the first word (i.e., the prenominal adjective), but no effect from distractors phonologically related to the second word (i.e., the noun). The authors interpreted this result as evidence that speakers can initiate articulation after encoding the form of the first word or perhaps even only the first syllable (see Bachoud-Lévi, Dupoux, Cohen, & Mehler, 1998). These results can thus be interpreted as supporting the view that only the first word's phonological form has become activated and selected before utterance onset.

The studies sketched thus far showed either inhibition or null effects from distractors phonologically related to a lexical element occurring in noninitial position. However, other studies have been showing facilitation effects, which were reduced when compared to single word utterances. In a recent study, Jescheniak et al. (2001) compared the production of bare nouns and the production of noun phrases consisting of a definite determiner and a noun. When participants named pictures of simple objects by producing a bare noun, facilitation effects of 71 ms from auditory distractors phonologically related to the noun were obtained at an SOA of 0 ms. By contrast, this facilitation effect was reduced to 25 ms when participants named the same objects by producing a definite noun phrase consisting of a determiner and a noun. This reduction of the priming effect across utterance formats was highly reliable.

In another recent study, Costa and Caramazza (2002) studied the production of three different utterance formats in English, bare nouns, noun phrases consisting of a determiner and a noun, and noun phrases consisting of a determiner, a prenominal color adjective, and a noun. The only SOA tested was one of 0 ms. Distractors that were semantically related to the noun yielded inhibition effects for all three utterance formats of about the same size (inhibition of 27 ms, 25 ms, and 21 ms for the three utterance formats). For distractors that were phonologically related to the noun, significant facilitation effects were obtained for all three utterance formats (39 ms, 28 ms, and 24 ms, respectively). According to the authors, facilitating the phonological encoding of the noun allows more processing resources to be allocated to the phonological encoding of preceding elements, which then leads to a faster utterance onset. Note that this interpretation is actually a version of the view that the phonological encoding of a whole noun phrase is completed before articulation is initiated. At closer inspection, however, the data pattern presented by Costa and Caramazza (2002) shows a clear trend toward a decrease of phonological facilitation as a function of the primed element's linear position: from 39 ms in bare nouns over 28 ms in determiner–noun utterances to 24 ms in determiner–adjective–noun. This reduction was not reliable, but possibly only because of a lack in statistical power.<sup>4</sup>

Whereas all these data seem to suggest that phonological priming for noninitial elements is reduced or absent, or even turns into inhibition, a study by Miozzo and Caramazza (1999, Experiment 4) yielded different results. Miozzo and Caramazza tested the effect from distractors phonologically related to the noun on the production of bare nouns or noun phrases consisting of a determiner and a noun in Italian. Priming effects of 29 ms were obtained for both utterance formats at an SOA of 0.

Taken together, the resulting picture is confusing. Distractors that are phonologically related to later words in an utterance have been found to yield inhibition (e.g., Meyer, 1996), null effects (e.g., Schriefers & Teruel, 1999), reduced facilitation (e.g., Costa & Caramazza, 2002; Jescheniak et al., 2001), or full-blown facilitation (Miozzo & Caramazza, 1999). Just as the results appear to be contradictory, so are the theoretical interpretations. They range from the assumption that only the first word's phonological form is activated at utterance onset, over the assumption that activation for utterance noninitial elements is reduced, to the assumption that the complete utterance has become phonologically encoded before utterance onset. Furthermore, a direct comparison of the different studies is complicated by the fact that they differ on a number of relevant dimensions. Just to name a few, distractors were sometimes presented visually and sometimes auditorily, some studies tested a range of SOAs whereas others did not, and the studies tested different languages.

Obviously, two things are needed given this state of affairs. First, one needs a unitary theoretical account that can accommodate the fact that phonologically related distractors can yield effects ranging from facilitation over null effects to inhibition. Second, the conditions that lead to the different effects must be systematically tested in a way that avoids the above-mentioned problems concerning the comparability of the available studies. In the remainder of this article we first sketch such a unitary account, and we then put this account to an empirical test.

Our account extends the one proposed by Meyer (1996) and shares main features with Dell's (e.g., Dell, 1986) model of speech production. It starts from the assumption that the phonological forms of the successive words receive a graded pattern of activation before articulation is initiated. That is, the subsequent words differ with respect to their activation level, decreasing from left to right, such that activation strength translates into the position an element occupies in the utterance. Elements outside the phonological advance planning scope have an activation of zero. Further, we assume that any distortion of this graded activation pattern—and thus a disturbance of the linear position coding—leads to interference during phonological encoding. According to such a model, primes that enhance the activation of the utterance-initial element are most beneficial because they preserve the pattern and thus should be able to speed the encoding process without much costs. However, primes that enhance the activation of noninitial elements might disturb the graded activation pattern such that the primed element moves to a wrong (i.e., too early) position. Though some

<sup>4</sup> Note that the participant sample in the study by Jescheniak et al. (2001), in which a significant reduction of the phonological facilitation effect from bare nouns to simple determiner + noun phrases was observed, was more than twice as large as the participant sample in the study by Costa and Caramazza (2002).

benefit from activating part of the target utterance should arise in this case as well, it might (in part or fully) be consumed by costs associated with the need for reestablishing the original order among the elements. These costs are not necessarily independent of the position in which the primed element is intended to occur. For example, depending on the parameters of the phonological encoding system, priming an element at a later position of the actual utterance might lead to a more severe displacement of this element from its original position. Hence the costs associated with reordering the elements might be higher. In consequence, then, the facilitation effect from phonologically related distractors might either be reduced, or be absent, or even turn into inhibition when priming an utterance noninitial element, depending on the relative amount of benefit and costs. We return to a more detailed description of this account in the general discussion.

In the experiment we report here, we systematically varied the linear position of the noun, by comparing three different utterance formats: bare nouns, simple noun phrases (determiner + noun), and complex noun phrases (determiner + size adjective + color adjective + noun). Except for the variation of utterance format, the three corresponding subexperiments were completely identical. There were four distractor conditions: one with distractors semantically related to the noun, a second one with distractors phonologically related to the noun, and two corresponding unrelated control conditions. As pointed out above, for the semantic condition, we expected inhibition of approximately the same size for all three utterance formats. As for the phonological condition, the graded activation account predicts a systematic decrease of the phonological facilitation effect with increasing complexity of the utterance format, with perhaps even inhibition for the most complex utterance format. To rule out that the priming effect from phonologically related distractors just shifts to a later time window as the primed noun moves to a later position in the utterance rather than changing its size or direction, the experiment tested three different SOAs: 0 ms, 150 ms, and 300 ms.

## Method

### Participants

A total of 108 native speakers of German, most of them students from the University of Leipzig, took part in the experiment. They were randomly assigned to one of three groups consisting of 36 participants each. The first group produced bare noun utterances, the second group produced simple noun phrase utterances, and the third group produced complex noun phrase utterances. Participants were paid DM 13 (approximately U.S.\$5.50). They had no known hearing deficit, and they had normal or corrected-to-normal vision. Participants with overall error rates exceeding 15% were replaced.

### Materials

The materials were taken from a previous study (Jescheniak et al., 2001). There were line drawings of 24 different objects that all had one unambiguous name. An equal number of objects had names with masculine, feminine, or neuter gender. For each object, distractors were selected in the following way: Semantically related distractors came from the same semantic category as the picture name but had no phonological overlap with it. Phonologically related distractors minimally shared the initial consonant–vowel segments with the object name but had no obvious semantic relation with it. The semantically related and the phonologically related distractor word for a given target object had the same grammatical

gender, which was different from the gender of the picture name. Across the whole set of items, distractor words from the three gender classes occurred equally often, and all possible combinations of target object gender and distractor gender were represented equally often (see the Appendix for a complete list of the experimental materials). Semantically and phonologically unrelated control conditions were created by reassigning the related distractors to the experimental pictures with the constraint that the related and the unrelated distractor for each object had the same gender.

Pictures of each object were prepared in two different sizes: small (approximately 70 × 70 mm) and large (approximately 123 × 123 mm). During the experiment, the pictures were presented either in red or in blue. The acoustic distractor words were spoken by a female native speaker of German. Semantically related distractors varied in duration from 425 ms to 879 ms with an average of 548 ms ( $SD = 115$  ms), and phonologically related distractors varied in duration from 382 ms to 721 ms with an average of 557 ms ( $SD = 86$  ms). All auditory materials were digitized at a sampling rate of 22 KHz for presentation during the experiment.

An additional set of eight objects with appropriate distractors was selected for the construction of practice and warm-up trials. Thirty-two trials were constructed, 17 of these to be used in a practice block and 5 as warm-ups at the beginning of each of three experimental blocks.

### Design

Utterance format was tested between participants. For each utterance format, there were two pairs of critical comparisons of distractor conditions (semantically related vs. semantically unrelated and phonologically related vs. phonologically unrelated). SOA was varied in three steps (0 ms, 150 ms, 300 ms) as a within-subject variable. Thus, there were three crossed variables: the three-level variables utterance format and SOA and the two-level variable relatedness.

Each participant received each target picture exactly once in each of the 12 conditions formed by fully crossing the three levels of the SOA variable and the four levels of the distractor variable, resulting in a total of 288 experimental trials. Regardless of the actual utterance format, participants received colored pictures of varying size with the size of the target picture (small vs. large) and its color (red vs. blue) being systematically counterbalanced. Across groups of participants, the items were reassigned to these four conditions in a way such that, overall, each item occurred in each condition equally often. As a consequence, across participants and within each SOA, the four versions of a target picture (resulting from the complete crossing of size and color) occurred equally often in each of the four distractor conditions.<sup>5</sup> SOAs were blocked with the sequence of SOA blocks being counterbalanced across participants using a Latin square procedure. Twelve different experimental lists were created according to the following general criteria: (a) Identical distractors did not follow in adjacent trials, (b) repetitions of a target picture were separated by at least two intervening trials, (c) no more than three trials with the same distractor condition or the same target word gender followed each other, and (d) the target pictures were not of the same size or of the same color in more than four subsequent trials. The 12 experimental lists were used equally often.

### Procedure

Each participant was tested individually in a session lasting about 1 hr. The participant was comfortably seated in a dimly lit room, separated from

<sup>5</sup> This is of particular importance for the complex noun phrase format. For this format, a distractor starting with the initial consonant of one of the two size adjectives or one of the two color adjectives could have an effect on the phonological encoding of the corresponding adjective. However, because of our counterbalancing procedure, such an effect would contribute to all distractor conditions and all SOAs to the same extent.

the experimenter by a partition wall. The visual stimuli were presented on a 37.4-cm (17-in.) CRT as colored line drawings on a light gray background. Viewing distance was about 60 cm. The presentation of the visual and auditory stimuli and the on-line collection of the data were controlled by a computer with a Pentium processor (Intel Corporation, Santa Clara, CA). Auditory distractors were presented with headphones at a comfortable listening volume. Speech-onset latencies were measured to the closest millisecond with a voice-key connected to the computer.

A trial was structured as follows. Target pictures were presented for 1,700 ms in the center of the CRT. Auditory distractors were presented either simultaneously with the onset of the target picture (SOA = 0 ms), 150 ms later (SOA = 150 ms), or 300 ms later (SOA = 300 ms). Participants described the target picture as quickly as possible in the format specified by the instructions. Speech onset latencies were measured from the onset of the target picture. The total length of one trial was about 4,500 ms.

The actual experiment consisted of three parts: a study phase, a practice phase, and the main session. During the study phase participants studied a written instruction booklet that emphasized both the speed and accuracy of their responses. Participants also received a booklet showing all experimental pictures. Next to each picture, the depicted object's name was printed. Participants were instructed to use these names only. Then the 17 practice trials were administered. No auditory distractors were presented during this practice phase. The experimenter monitored whether participants used the appropriate utterance format and corrected them if necessary. Then the main experiment started with the first of three SOA blocks. There were short breaks between these blocks. At the beginning of each experimental block, five warm-up trials containing practice pictures preceded the experimental trials.

## Results

Observations were coded as erroneous and discarded from the reaction time analyses whenever any of the following conditions held: (a) a picture had been named with other than the expected name; (b) a nonspeech sound preceded the target utterance, triggering the voice key; (c) a dysfluency occurred or an utterance was repaired; or (d) a speech onset latency exceeded 1,500 ms. Observations deviating from a participant's and an item's mean by more than two standard deviations were considered as outliers and also discarded from the reaction time analyses. According to these criteria, 1,532 observations (4.9%) were marked as erroneous and 486 observations (1.6%) as outliers.

Averaged reaction times were submitted to analyses of variance (ANOVAs), separately for the semantic distractor conditions and the phonological distractor conditions. Statistical analyses involved the three fixed variables utterance format (bare noun, simple noun phrase, complex noun phrase), relatedness (semantically related vs. semantically unrelated or phonologically related vs. phonologically unrelated), and SOA (SOAs = 0 ms, 150 ms, 300 ms). Two complementary analyses were computed, one treating participants and one treating items as random variables (Clark, 1973).

Table 1 displays mean reaction times and error rates per SOA and distractor type for the three utterance formats. Figure 1 displays the reaction time differences between related and unrelated conditions for the semantic and the phonological conditions separately for each utterance format and each SOA.

### *Effects From Semantically Related Distractors*

Naming responses were slower and less accurate for complex noun phrases than for the other two utterance formats, yielding a

reliable effect of utterance format, for naming latencies,  $F_1(2, 105) = 59.61, p < .01, MSE = 66,384.75$ ;  $F_2(2, 46) = 1,610.60, p < .01, MSE = 1,665.59$ ; and for error rates,  $F_1(2, 105) = 36.85, p < .01, MSE = 3.11$ ;  $F_2(2, 46) = 37.57, p < .01, MSE = 4.57$ . Naming latencies and error rates decreased from SOA = 0 ms to SOA = 300 ms, as reflected in a significant SOA effect for naming latencies,  $F_1(2, 210) = 51.25, p < .01, MSE = 5,413.78$ ;  $F_2(2, 46) = 183.55, p < .01, MSE = 1,052.90$ ; and for error rates,  $F_1(2, 210) = 9.87, p < .01, MSE = 1.54$ ;  $F_2(2, 46) = 4.65, p < .05, MSE = 4.91$ . Relatedness was also significant, reflecting an inhibition of naming responses by a semantically related distractor when compared to an unrelated distractor (14 ms/2.0% inhibition); for naming latencies,  $F_1(1, 105) = 33.68, p < .01, MSE = 936.58$ ;  $F_2(1, 23) = 7.91, p < .05, MSE = 3,657.70$ ; and for error rates,  $F_1(1, 105) = 35.81, p < .01, MSE = 1.06$ ;  $F_2(1, 23) = 10.55, p < .01, MSE = 5.41$ . The interaction of this variable with utterance format was reliable in the analyses of response latencies and error rates; for naming latencies,  $F_1(2, 105) = 3.49, p < .05, MSE = 936.58$ ;  $F_2(2, 46) = 3.27, p < .05, MSE = 1,079.23$ ; and for error rates,  $F_1(2, 105) = 9.69, p < .01, MSE = 1.06$ ;  $F_2(2, 46) = 9.22, p < .01, MSE = 1.68$ . At first sight, this seems to suggest that the overall inhibition effect from semantically related distractors effect was not equal for the different utterance formats. However, a closer inspection of the data pattern reveals that the effect did not differ in size but only in its extension over SOAs, being obtained at more SOAs for complex and simple noun phrases than for bare nouns; we return to this issue in the Discussion. When the analyses were restricted to the one SOA at which semantic inhibition was observable for all utterance formats (SOA = 0 ms), semantic inhibition was found to be independent of utterance format, in line with our prediction (22 ms/1.2%, 25 ms/3.0%, and 25 ms/2.8% inhibition); main effect of relatedness: for naming latencies,  $F_1(1, 105) = 43.85, p < .01, MSE = 706.56$ ;  $F_2(1, 23) = 8.19, p < .05, MSE = 3,216.46$ ; for error rates,  $F_1(1, 105) = 16.68, p < .01, MSE = 1.36$ ;  $F_2(1, 23) = 4.68, p < .05, MSE = 7.27$ ; interaction of relatedness and utterance format:  $F_1$  and  $F_2 < 1$ , for naming latencies; and for error rates,  $F_1(2, 105) = 1.29, p = .28, MSE = 1.36$ ;  $F_2(2, 46) = 1.91, p = .16, MSE = 1.38$ .

The interaction of utterance format and SOA was significant in the item analyses of reaction times and error rates, reflecting a differential decrease of naming latencies and error rates over SOAs for the three utterance formats; for naming latencies,  $F_1 < 1$ ;  $F_2(4, 92) = 3.14, p < .05, MSE = 705.07$ ; and for error rates,  $F_1(4, 210) = 1.32, p = .26, MSE = 1.54$ ;  $F_2(4, 92) = 2.44, p = .05, MSE = 1.25$ . The interaction of relatedness and SOA, finally, was significant in the analyses of reaction times and error rates (though only marginally so in the item analysis of error rates); for naming latencies,  $F_1(2, 210) = 7.14, p < .01, MSE = 593.53$ ;  $F_2(2, 46) = 3.24, p < .05, MSE = 1,018.46$ ; and for error rates,  $F_1(2, 210) = 4.14, p < .05, MSE = 1.05$ ;  $F_2(2, 46) = 2.51, p = .09, MSE = 2.59$ . It reflects the fact that inhibition from semantically related distractors was strongest at SOA = 0 ms and smallest at SOA = 300.

### *Effects From Phonologically Related Distractors*

As in the analysis of the semantic conditions, naming responses for complex noun phrases were slower and less accurate than for the other two utterance formats, yielding a reliable effect of utter-

Table 1  
*Mean Reaction Times (in Milliseconds) and Error Rates (in Percentages) by Stimulus Onset Asynchrony, Utterance Format, and Distractor Type*

Distractor	Stimulus onset asynchrony					
	0 ms		150 ms		300 ms	
	<i>M</i>	%	<i>M</i>	%	<i>M</i>	%
Bare noun						
SEM-REL	642	3.9	601	3.5	554	2.7
SEM-UNR	620	2.7	601	3.1	551	2.7
Difference	+22 <sup>***/*</sup>	+1.2	0	+0.4	+3	0.0
PHO-REL	592	3.6	554	1.5	541	2.4
PHO-UNR	642	3.4	607	3.5	555	2.4
Difference	-50 <sup>****</sup>	+0.2	-53 <sup>****</sup>	-2.0 <sup>**/*</sup>	-14 <sup>**/****</sup>	0.0
Simple noun phrase						
SEM-REL	613	6.6	565	5.4	520	1.9
SEM-UNR	588	3.6	556	3.2	519	2.7
Difference	+25 <sup>****</sup>	+3.0 <sup>*/*</sup>	+9 <sup>†/ns</sup>	+2.2 <sup>*/*</sup>	+1	-0.8
PHO-REL	583	4.3	534	2.7	522	2.0
PHO-UNR	603	4.7	555	3.8	517	3.0
Difference	-20 <sup>**/*</sup>	-0.4	-21 <sup>****</sup>	-1.1 <sup>†/ns</sup>	+5	-1.0
Complex noun phrase						
SEM-REL	848	12.1	826	11.4	790	8.6
SEM-UNR	823	8.3	802	5.9	771	5.8
Difference	+25 <sup>**/†</sup>	+3.8 <sup>**/*</sup>	+24 <sup>**/*</sup>	+5.5 <sup>****/*</sup>	+19 <sup>†/†</sup>	+2.8 <sup>*/*</sup>
PHO-REL	855	10.4	824	7.6	781	7.0
PHO-UNR	827	7.6	820	6.0	773	7.2
Difference	+28 <sup>**/*</sup>	+2.8 <sup>†/†</sup>	+4	+1.7	+8	-0.2

Note. SEM-REL = semantically related to the target object name; SEM-UNR = semantically unrelated to the target object name; PHO-REL = phonologically related to the target object name; PHO-UNR = phonologically unrelated to the target object name. Positive difference scores reflect interference, and negative difference scores reflect facilitation. Results from the *t* tests by participant precede the results from the *t* tests by item. †*p* < .10. \**p* < .05. \*\**p* < .01. \*\*\**p* < .001.

ance format; for naming latencies,  $F_1(2, 105) = 64.08, p < .01, MSE = 68,916.11; F_2(2, 46) = 1977.87, p < .01, MSE = 1,498.21$ ; and for error rates,  $F_1(2, 105) = 36.29, p < .01, MSE = 2.40; F_2(2, 46) = 43.84, p < .01, MSE = 2.98$ . Naming responses were faster and more accurate at longer SOAs, as reflected in a reliable SOA effect; for naming latencies,  $F_1(2, 210) = 45.29, p < .01, MSE = 5,610.78; F_2(2, 46) = 206.41, p < .01, MSE = 859.03$ ; and for error rates,  $F_1(2, 210) = 6.85, p < .01, MSE = 1.51; F_2(2, 46) = 7.69, p < .01, MSE = 2.02$ . Overall, phonologically related distractors speeded up naming latencies when compared to unrelated distractors, yielding a significant effect of relatedness (-13 ms priming);  $F_1(1, 105) = 23.94, p < .01, MSE = 1,060.42; F_2(1, 23) = 7.72, p < .01, MSE = 2,013.36$ . Relatedness interacted with SOA in the analysis of naming latencies, reflecting the fact that overall priming was strongest at SOA = 150 ms (-23 ms), somewhat smaller at SOA = 0 ms (-14 ms), and not present at SOA = 300 ms (0 ms),  $F_1(2, 210) = 10.64, p < .01, MSE = 672.92; F_2(2, 46) = 6.72, p < .01, MSE = 704.30$ . Most important, relatedness interacted with utterance format; for naming latencies,  $F_1(2, 105) = 35.15, p < .01, MSE =$

$1,060.42; F_2(2, 46) = 24.80, p < .01, MSE = 1,004.56$ ; for error rates,  $F_1(2, 105) = 5.18, p < .01, MSE = 0.95; F_2(2, 46) = 4.32, p < .05, MSE = 1.71$ . Although priming was strongest with bare nouns (-39 ms/-0.6%) and reduced with simple noun phrases (-12 ms/-0.8%), a trend toward inhibition was obtained with complex noun phrases (13 ms/1.4%). Further analyses explored this interaction in detail. They revealed that the effect from phonologically related distractors differed pairwise for all three utterance formats as indicated by reliable interactions of relatedness and format. These interactions were significant in the comparison of bare nouns and simple noun phrases; for reaction times,  $F_1(1, 70) = 29.91, p < .01, MSE = 654.99; F_2(1, 23) = 29.05, p < .01, MSE = 433.27$ ; for error rates,  $F_1$  and  $F_2 < 1$ ; in the comparison of bare nouns and complex noun phrases, for reaction times,  $F_1(1, 70) = 54.99, p < .01, MSE = 1,355.28; F_2(1, 23) = 30.95, p < .01, MSE = 1,609.95$ ; and for error rates,  $F_1(1, 70) = 5.74, p < .05, MSE = 1.09; F_2(1, 23) = 4.16, p = .05, MSE = 2.26$ ; and also in the comparison of simple noun phrases and complex noun phrases; for reaction times,  $F_1(1, 70) = 15.11, p < .01, MSE = 1,170.99; F_2(1, 23) = 12.70, p < .01, MSE = 970.48$ ; and for error

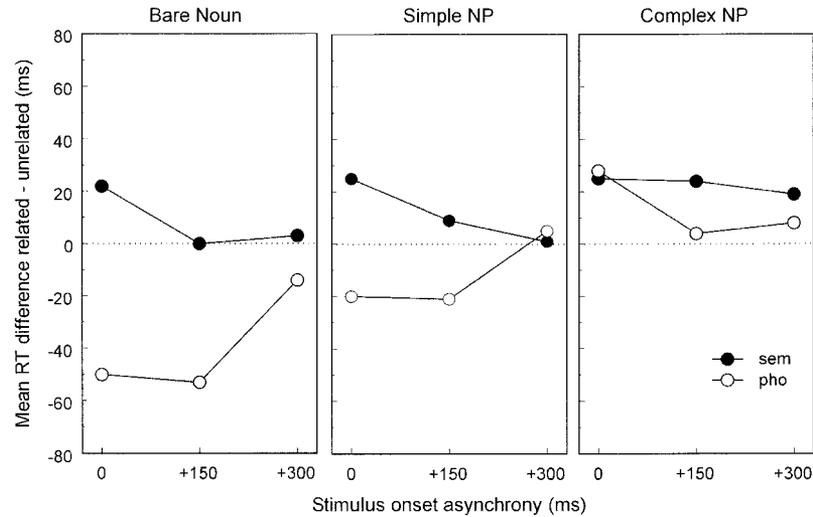


Figure 1. Mean reaction time (RT) differences (related – unrelated) for the semantic (sem) and phonological (pho) conditions broken down by stimulus onset asynchrony and utterance format. NP = noun phrase.

rates,  $F_1(1, 70) = 7.01, p < .01, MSE = 1.19$ ;  $F_2(1, 23) = 8.21, p < .01, MSE = 1.52$ . The same pattern was obtained when constraining the analyses to  $SOA = 0$  ms, as we had done for the semantic conditions. The interaction of utterance format and SOA was significant in the item analysis of reaction times, reflecting a differential decrease of naming latencies over SOAs for the three utterance formats ( $F_1 < 1$ ),  $F_2(4, 92) = 4.33, p < .01, MSE = 595.70$ .

Finally, there was a reliable interaction of utterance format, SOA, and relatedness in the analysis of naming latencies, reflecting that for complex noun phrases inhibition from phonologically related primes rather than facilitation was obtained, but only at  $SOA = 0$  ms,  $F_1(4, 210) = 6.21, p < .01, MSE = 672.92$ ;  $F_2(2, 46) = 8.92, p < .01, MSE = 347.82$ .

Before turning to the Discussion section, we should address two additional issues. First, although the overall pattern of results is very consistent, there seems to be one exception. Whereas for the bare noun format and the simple noun phrase format the means of the phonologically unrelated condition decreased from  $SOA = 0$  ms to  $SOA = 150$  by 35 ms and 48 ms, respectively, the decrease for the complex noun phrase format was only 7 ms. Therefore, one could object that the difference between the phonologically related and the phonologically unrelated condition observed at  $SOA = 0$  ms for the complex noun phrase format is not due to longer naming latencies in the related condition (reflecting interference) but rather to unexpectedly fast naming latencies in the unrelated condition.

We explored this possibility as follows. For each utterance format, there were 12 different experimental lists, with three participants assigned to each list. Across these 12 lists, all important variables (e.g., sequence of SOAs, possible combinations of size adjective, color adjective, and object name) were fully controlled. In a reanalysis of the data from the complex noun phrase format, we assigned the two participants from each list who showed the strongest decrease in the phonologically unrelated condition from  $SOA = 0$  ms to  $SOA = 150$  ms to one group

(Group A) and the remaining participant to a second group (Group B). This procedure ensured that for both groups the important variables were controlled in the same way as they had been for the complete sample. For Group A, there was a decrease of reaction times in the phonologically unrelated condition from  $SOA = 0$  ms to  $SOA = 150$  ms of 43 ms, whereas for Group B there was an increase of 64 ms. Thus, the decrease for Group A was similar to the decrease observed for the participant samples tested with bare nouns and simple noun phrases (35 ms and 49 ms, respectively). If the objection sketched above is correct, inhibition from phonologically related distractors at  $SOA = 0$  ms should be observed for Group B but not for Group A. However, an analysis of the reaction times at  $SOA = 0$  ms with the variables group (A vs. B) and relatedness (phonologically related vs. phonologically unrelated) showed that this was not the case. For Group A, there was a 23-ms inhibition, and for Group B, there was a 40-ms inhibition. The main effect of relatedness was significant,  $F_1(1, 34) = 10.71, p < .01, MSE = 1,356.97$ ;  $F_2(1, 23) = 5.51, p < .05, MSE = 4,884.04$ , whereas the interaction of relatedness with group was not ( $F_1 < 1$ );  $F_2(1, 23) = 1.40, p = .25, MSE = 2,114.65$ .<sup>6</sup> Put differently, the inhibitory effect from phonologically related distractors is observed if there is a decrease in reaction times from  $SOA = 0$  ms to  $SOA = 150$  ms in the unrelated condition similar to the one obtained for the other utterance formats as well as with the reverse pattern alike.

Second, as pointed out above, inhibitory effects from phonologically related distractors have until now been found only in one study (Meyer, 1996), and they were rather small. Thus, one might wonder which factors determine whether an inhibitory effect is

<sup>6</sup> Subsequently computed pairwise comparisons showed that the 40-ms inhibition effect was significant,  $t_1(11) = 3.20, p < .01$ ;  $t_2(23) = 2.22, p < .05$ , whereas the 23-ms inhibition effect just failed to reach significance,  $t_1(23) = 1.99, p = .06$ ;  $t_2(23) = 1.70, p = .11$ .

obtained. One possible factor is the size of what one might call the *base priming effect*, that is, a given phonologically related distractor's potential of facilitating the retrieval of the corresponding target word's phonological form. The amount of facilitation in bare noun production provides an index of this base priming effect. It is possible that an inhibitory effect for an element in noninitial position is only obtained for items with a relatively small base priming effect, whereas no such effect (or maybe even a slight facilitation effect) is obtained for items with a relatively large base priming effect. We investigated this possibility as follows. First, we divided the set of 24 pictures in two subsets, the criterion being the amount of priming in bare noun production at SOA = 0 ms. The first subset contained the 12 pictures with a small base priming effect, with a mean facilitation effect of 27 ms (Set 1). The second subset comprised the remaining 12 pictures with a relatively large base priming effect, with a mean facilitation effect of 71 ms (Set 2). Next, we repeated the analysis of the results for the simple noun phrases and the complex noun phrases at SOA = 0 ms, by including item subset as an additional factor. The results are presented in Figure 2.

For the subset of items with a large base priming effect, the significant 71-ms facilitation effect obtained for bare noun utterances,  $t_1(1, 35) = 7.20, p < .01$ ;  $t_2(11) = 9.59, p < .001$ , reduces to a 49-ms facilitation effect for simple noun phrase utterances,  $t_1(1, 35) = 5.39, p < .01$ ;  $t_2(11) = 3.93, p < .01$ , and turns into a nonsignificant 16-ms inhibition effect for complex noun phrase utterances,  $t_1(1, 35) = .96, p = .34$ ;  $t_2(11) = .63, p = .54$ . For the subset with a small base priming effect, the significant 27-ms facilitation effect for bare noun utterances,  $t_1(1, 35) = 3.53, p = .01$ ;  $t_2(11) = 3.69, p < .01$ , reduces to a nonsignificant 10-ms inhibition effect for simple noun phrase utterances,  $t_1(1, 35) = 1.31, p = .20$ ;  $t_2(11) = .38, p = .71$ , and turns into a significant 41-ms inhibition effect for complex noun phrase utterances,  $t_1(1, 35) = 4.21, p < .01$ ;  $t_2(11) = 3.63, p < .01$ .

This pattern indeed suggests that whether one observes reliable inhibition for complex noun phrase utterances appears to depend,

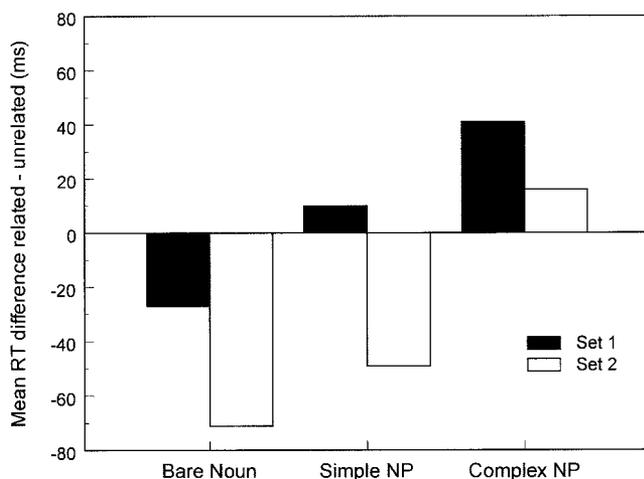


Figure 2. Mean reaction time (RT) differences (related – unrelated) for the phonological conditions at a stimulus onset asynchrony of 0 ms, broken down by utterance format and item subset. Set 1: items with a small priming effect in bare noun production. Set 2: items with a large priming effect in bare noun production. NP = noun phrase.

at least in part, on the strength of the base priming effect. One could thus hypothesize that the occurrence of an inhibitory effect from phonologically related distractors depends on at least two factors: the position of the primed element in the target utterance and the base priming effect for this element. The first factor leads to a gradient of priming effects, with systematic reduction of facilitation effects with later positions in the eventual utterance. The second factor appears to shift this gradient in a largely additive way. This shift can but need not lead to inhibitory effects from phonologically related primes. Of course, this hypothesis is based on a post hoc analysis and should be put to a direct experimental test in future research.

## Discussion

Extant picture–word experiments have yielded very different results concerning the effects of distractors that are phonologically related to a word in noninitial position of a multiword target utterance. The results reported range from full facilitation (relative to one-word utterances) to reduced facilitation, no effect, and (at the other extreme) inhibition. Furthermore, different theoretical accounts with respect to the phonological encoding process have been proposed for these different patterns of results, and these accounts partly contradict each other. Finally, a direct comparison of the different studies is complicated by the fact that they differ in a number of aspects, such as the utterance formats being used, the number and range of SOAs being tested, and the language under investigation.

In view of the seemingly conflicting findings, we proposed a unitary account, the graded activation account. This account holds that the phonological codes of successive words of a multiword utterance receive some activation before utterance onset, with the strength of activation an element receives decreasing as a monotonic function of its position in the utterance. The graded activation account predicted that the typical phonological facilitation effect obtained with single word utterances should gradually decrease as the primed word moves to later positions in a multiword utterance. These predictions were put to test by comparing the effect of distractors that were phonologically related to the noun in bare noun naming, simple noun phrase production, and complex noun phrase production. In addition, the experiment included a condition with distractors that were semantically related to the noun of the target utterance. Because for all three utterance formats access to the noun lemma is necessarily required before utterance onset, we expected a semantic interference effect for all three utterance formats. Except for the variation of utterance format, the corresponding parts of the experiment were identical, thus allowing for a nonconfounded comparison of the different utterance formats.

As expected, semantic interference was obtained for all three utterance formats, and it was comparable in size. Phonological facilitation was largest in bare noun production and reduced in simple noun phrase production. No priming effect was observed in complex noun phrase production at SOAs of 150 ms and 300 ms, and at an SOA of 0 ms phonologically related distractors even led to inhibition. Hence, the whole range of effects from phonological distractors as reported in the literature could be observed in the present experiment, by systematically varying the utterance format. At a global level, the overall pattern of results is in line with

our predictions. However, the time course of the obtained effects requires further discussion.

We start with the semantic interference effect, which was observed for all three utterance formats. When moving from bare nouns over simple noun phrases to complex noun phrases, the semantic interference effect extended to longer positive SOAs. This could be due to the fact that, unlike bare noun utterances, the latter two utterance formats required the retrieval of the noun's grammatical gender, which, in turn, might lead to a prolongation of the lemma retrieval process (cf. La Heij, Mak, Sander, & Willeboordse, 1998). For complex noun phrase utterances, the semantic interference effect extended even to the longest positive SOA tested, at least descriptively. At the same time, overall naming latencies for complex noun phrases were some 200 ms longer when compared to either bare noun utterances or simple noun phrase utterances. These longer naming latencies can be explained by an extended conceptual preparation stage, needed for extracting and coding additional stimulus attributes (i.e., color and size). According to the model by Levelt et al. (1999), a lemma representation is activated when the corresponding concept has become activated. Hence, any extension of the conceptual preparation phase should also enlarge the time window in which semantically related distractors can interfere with the lemma processing stage. One possibility of dealing with the difference in overall naming latencies across utterance formats is to consider SOAs relative to mean speech-onset time rather than relative to picture onset. If one does so, SOAs of 150 ms and 300 ms in complex noun phrase production can be viewed as roughly analogous to SOAs of 0 ms and 150 ms in bare noun and simple noun phrase production. In fact, the pattern for the semantic conditions was comparable in simple noun phrase production and in complex noun phrase production when considering the adjusted SOAs.<sup>7</sup> The same SOA adjustment is applied in the discussion of the phonological conditions.

The effects from phonologically related distractors were systematically modulated by utterance format. For bare nouns, a substantial facilitation effect was obtained at SOAs of 0 ms and 150 ms. For simple noun phrases, the facilitation effects at these SOAs were approximately reduced by half, and for complex noun phrases, no effects were present at SOAs of 150 ms and 300 ms (i.e., the SOAs which after adjustment should be comparable to SOAs of 0 ms and 150 ms in bare noun production and in simple noun phrase production). Thus, at these two SOAs, shifting the linear position of the phonologically primed target word systematically changed the size of the effect from full-blown facilitation through reduced facilitation to a null effect.<sup>8</sup>

Related to these SOA adjustments is another issue that needs some consideration. Naming latencies for complex noun phrases were not only about 200 ms longer than for the other two utterance formats, but there must also have been clear differences in utterance duration between the utterance formats. Articulation of the two adjectives in complex noun phrases must have taken approximately 600 to 700 ms. Thus, relative to distractor onset, the noun in complex noun phrase utterances was articulated some 800 to 900 ms later than in the other two utterance types. In view of this difference, one could speculate whether the absence of a facilitation effect in complex noun phrases was simply due to this rather long temporal distance between distractor presentation and actual realization of the noun. This explanation, however, appears un-

likely as it would predict a null effect but not an inhibition effect as empirically observed.

Before elaborating on how the graded activation account can explain the present data, we comment on the possible locus of inhibition effects from phonologically related distractors, as observed by Meyer (1996) and in the present experiment (complex noun phrases, SOA = 0 ms). In Meyer's experiment, distractors that were phonologically related to the second noun in coordinated noun phrase utterances (e.g., *arrow* in the utterance "*the bag and the arrow*") inhibited the naming response. As discussed above, the author's account for this inhibition effect is very similar to the graded activation account proposed here. However, one could conceive an alternative interpretation that locates the source of this effect at a nonphonological processing level.<sup>9</sup> According to that account, a distractor phonologically related to the second noun might have increased the activation of the lemma of this noun, via cohortlike processing of the auditory distractor (e.g., Roelofs, Meyer, & Levelt, 1996). As the syntactic frame of the utterance contains two slots that are labeled as noun slots (i.e., these slots can only be filled by a noun lemma), this increase of activation for the second noun lemma might have resulted in an enhanced competition of the two noun lemmas for the first noun slot. Note that this explanation is identical to the one that is used to account for overt word exchange errors, and more specifically for the fact that overt word exchange errors almost exclusively concern words of the same syntactic category. For the issue at hand, this account makes an important prediction: Distractors that are phonologically related to a word in a later position of the upcoming utterance should involve costs (i.e., inhibition) only if there is an earlier slot in the syntactic frame that is labeled with the primed word's syntactic category. The utterance format used in Meyer's experiments does not allow one ultimately to rule out this alternative account, as there were two noun slots. However, it cannot be applied to the present study, as the syntactic frame for the utterances tested contained only one noun slot. Thus, the phonological interference for complex noun phrases observed in the present experiment has to be located at a processing level that is not governed by syntactic constraints, namely, at the level of phonological encoding. This is exactly what the graded activation account does. In the following, we apply in more detail this account to the pattern of phonological facilitation and inhibition effects observed in the present experiments.

<sup>7</sup> Statistical analyses revealed the following pattern: Format  $\times$  Relatedness:  $F_1$  and  $F_2 < 1$  for reaction times;  $F_1(1, 70) = 2.28, p = .14, MSE = 1.11$ ;  $F_2(1, 23) = 1.73, p = .20, MSE = 2.20$  for error rates; Format  $\times$  SOA  $\times$  Relatedness:  $F_1$  and  $F_2 < 1$  for reaction times;  $F_1 < 1$ ;  $F_2(1, 23) = 1.05, p = .32, MSE = 1.43$ .

<sup>8</sup> Subsequent analyses showed that for each utterance format, the priming effect was comparable at the respective two SOAs, the only exception from this general pattern being observed in the error rate analysis for bare noun utterances: Relatedness  $\times$  SOA in the analysis of noun utterances:  $F_1$  and  $F_2 < 1$  for reaction times;  $F_1(1, 35) = 4.81, p < .05, MSE = .52$ ;  $F_2(1, 23) = 5.97, p < .05, MSE = .63$ , for error rates; in the analysis of simple noun phrase utterances:  $F_1$  and  $F_2 < 1$  for reaction times;  $F_1$  and  $F_2 < 1$  for error rates; in the analysis of complex noun phrases:  $F_1$  and  $F_2 < 1$  for reaction times;  $F_1$  and  $F_2 < 1$  for error rates.

<sup>9</sup> We thank an anonymous reviewer for drawing our attention to this alternative account.

According to the graded activation account, two processes determine the net effect of phonological distractors: on the one hand, the effect of priming an element within the current phonological planning unit, and on the other hand the effect of disturbing a graded activation pattern coding the relative position of the individual elements within that unit. To illustrate, we consider the following simple model. It starts from the following assumptions:

1. At the lemma level different syntactic frames are generated for the different utterance formats ( $N$  for bare noun,  $Det-N$  for simple noun phrases, and  $Det-Adj-Adj-N$  for complex noun phrases), and the slots of these frames are filled by the corresponding lemmas.
2. The lemmas of this frame initially activate their corresponding phonological forms to some extent. The sequence of the lemmas in the syntactic frame is indexed as a graded activation pattern at the phonological level such that activation of the corresponding phonological word forms decreases monotonically with serial position. In our example we assume that the initial element receives 3.0 activation units and each subsequent element half of the activation of the element immediately preceding it.
3. Compared to an unrelated distractor, a phonologically related distractor facilitates the phonological encoding process of some part of the utterance. This gain is modeled by adding 2.0 activation units to the primed element.
4. Any disturbance of this graded pattern (i.e., disturbance of serial position coding) is associated with some costs. The assumption is that the later the linear position of the primed element, the higher the costs resulting from priming. This is the case because priming the element makes it momentarily a candidate for an earlier position, that is, priming displaces this element from its intended position. The larger the span over which this displacement occurs, the larger the costs for reestablishing the original serial position coding. Here we assume that these costs amount to 1.0 activation units per position displacement.

Perhaps these assumptions concerning displacement of phonological elements and costs for reestablishing the original serial position coding can best be viewed in analogy to monitoring and (covert) self-repairs (see Levelt, 1989, for a comprehensive treatment). Let us take the case of a complex noun phrase. Displacing the noun by one position results in a locally correct utterance up to the noun (determiner–adjective–noun). However, monitoring processes signal that there still is a second adjective that has to be integrated into the eventual utterance. A covert repair process, then, relocates this element to its intended position. Displacing the noun by two positions also results in a locally correct utterance up to the noun (determiner–noun). However, in this case monitoring processes signal that two elements still need to be integrated. However, as more elements need to be integrated, the covert repair process should now be more demanding (i.e., involve larger costs) when compared to the first case. Note that exactly such repair processes become visible in overt repairs of spontaneous speech

(e.g., *the big house, eh, the big red house*), and participants in our experiments did indeed occasionally produce such overt repairs.

Under assumptions 1–4, the facilitation and cost components result in a net effect as empirically observed at SOAs of 0 ms and 150 ms in bare noun and simple noun phrase production and at SOAs of 150 ms and 300 ms in complex noun phrase production. For a primed element occurring in first position, there is 2.0 activation units gain and no cost, resulting in a net effect of 2.0 activation units, compared to the situation with an unrelated distractor. For primed elements occurring in second or third position, there is a gain of 2.0 activation units and a cost of 1.0 activation units (displacement by one position), resulting in a net effect of 1.0 activation units. Finally, for a primed element occurring in fourth position, there is a gain of 2.0 activation units and a cost of 2.0 activation units (displacement by two positions), resulting in a net effect of 0.0 activation units. Hence, the pattern of net activation (2.0, 1.0, 1.0, and 0.0 activation units, respectively) translates into the empirically observed reduction of phonological priming as the primed element moves to later positions in the utterance.

We now turn to the inhibition effect observed at  $SOA = 0$  ms for complex noun phrases. When considering SOAs relative to mean speech-onset time,  $SOA = 0$  ms represents a particularly early point in time during the phonological encoding of complex noun phrases. Possibly, at this point the different elements' positions are not yet coded in a robust graded activation pattern, such that it can more easily be disturbed than during later processing. One way of thinking about this state of affairs is that the activation difference between adjacent elements is less pronounced during the initial processing phase (i.e., the activation gradient flattened). In our example, this is modeled by assuming that the utterance initial element has only gathered 1.0 activation units at this early point in time (as opposed to 3.0 activation units during the later point in time). When applying the same parameters as before, smaller activation differences between adjacent elements result. What about the net effect of priming for the different positions in this situation? Compared to the situation sketched earlier, the net effects of priming elements occurring in first and second position remain unchanged. For a primed element occurring in third position, however, there is a gain of 2.0 activation units and a cost of 2.0 activation units (displacement by two positions), resulting in a net effect of 0.0 activation units. For a primed element occurring in fourth position, there is a gain of 2.0 activation units and a cost of 3.0 activation units (displacement by three positions), resulting in a net effect of  $-1.0$  activation units. Thus, the priming effect now turns into inhibition, if the element occurs at a very late position in the utterance, as observed in the empirical data.

Above, we had shown that the occurrence of an inhibitory effect for complex noun phrases appears to depend on the base priming effect, as indexed by the size of the facilitation effect obtained in the bare noun condition. The inhibition effect was carried to a large extent by items for which base priming was small. This finding is in line with our simple model. In the model, the base priming effect of 50 ms for bare noun production at  $SOA = 0$  ms was captured as a gain of 2.0 activation units. Taking the division into a subset with a large base priming effect (71 ms) and a subset with a small base priming effect (27 ms), the size of these priming effects translates into a gain of approximately 3.0 units for the former subset and a gain of 1.0 unit for the latter subset. When applying the flattened activation gradient (assumed to reflect the

state of activation at SOA = 0 ms during complex noun phrases production), the results are as follows. For the subset with small base priming, there is a gain of 1.0 unit from priming and a displacement cost of 3.0 units, resulting in a net inhibition effect of -2.0 activation units, which is larger than the one predicted for the complete item set. By contrast, for the subset with large base priming, this situation results in a 3.0 unit gain from priming that, however, is cancelled out by a 3.0 displacement cost, resulting in a net effect of 0 activation units, which is smaller than the one predicted for the complete item set. This pattern is in complete agreement with the empirical data. Put differently, the modulation of the inhibitory effect for complex noun phrase production at SOA = 0 ms arises, because in our model any activation boost is enough to move an item to the initial position and therefore to make it suffer maximum interference. Larger base priming, then, does not make an item suffer more interference (as it is already at maximum), but rather it only has the effect of helping the phonological spell-out of that item, thus reducing the overall interference effect.

One should note, however, that our simple model does not yet make the correct predictions for simple noun phrases. Here we need to apply the original (steeper) activation gradient. When doing so, we arrive at the following predictions. For the subset with small base priming, priming leads to a 1.0 unit gain and no displacement cost, and thus a net facilitation effect of 1.0. For the subset with large base priming, priming leads to a 3.0 unit gain and a 1.0 unit displacement cost, and thus a net facilitation effect of 2.0 activation units. Although the relative ordering of net effects for the two subsets is in line with the empirical data, the model predicts some facilitation for the subset with small base priming in simple noun phrase production, which, however, was not observed in the empirical data.

All in all, it seems that our simple model accounts for the overall pattern of results and provides at least a reasonable approximation to the more fine grained pattern that is obtained when one distinguishes between subsets of items with differently sized base priming.

The model also provides a reasonable qualitative account for the results from some of the other studies mentioned in the introduction, as shown next. To start with, Costa and Caramazza (2002) obtained a 39-ms facilitation effect for bare noun utterances and facilitation effects of 28 ms and 24 ms for determiner-noun utterances and for determiner-adjective-noun utterances, respectively. In terms of the parameters used above, the 39-ms facilitation effect for bare noun utterances should translate into approximately 1.5 units gain and no displacement cost. For determiner-noun utterances and determiner-adjective-noun utterances, this 1.5 unit gain should be reduced by 1.0 unit displacement cost, resulting in a net facilitation effect of 0.5 activation units. This pattern is in line with the observed reduction in facilitation for determiner-noun and determiner-adjective noun utterances. In another study, Miozzo and Caramazza (1999, Experiment 4) obtained a 29-ms facilitation effect for bare noun utterances as well as for determiner-noun utterances. The 29-ms facilitation effect for bare noun utterances corresponds to a 1.0 unit gain and no displacement cost, resulting in a net facilitation effect of 1 activation unit. As the 1.0 unit gain is too small to lead to a displacement of the noun in determiner-noun utterances, a net facilitation effect of 1 activation unit is predicted for these utterances as well. Thus,

priming effects of equal size are predicted for the two utterance types, just as observed in the empirical data.

Jescheniak et al. (2001) compared the same two utterance formats. This study revealed a much larger facilitation effect for bare nouns (71 ms) and a reduced facilitation effect for determiner-noun utterances (25 ms). The large priming effect for bare nouns translates into a 3.0 unit gain without cost. For determiner-noun utterances, however, the large gain leads to a displacement of the primed noun and thus a displacement cost of 1.0 unit, resulting in a reduced net priming effect of 2.0 activation units. Again, this difference in net activation for the two utterance formats is in line with the empirical pattern. Finally, Schriefers and Teruel (1999) obtained a 69-ms facilitation effect for bare noun utterances but only a nonsignificant 4-ms facilitation effect for adjective-noun utterances. As the base priming effect is similar to the one observed in the study by Jescheniak et al., the model once more predicts a reduced but sizeable phonological priming effect if the noun appears in second position in the utterance. Contrary to this prediction, however, no such effect was obtained, and thus the model cannot provide an accurate account of this data set yet.

When evaluating all these data against the model predictions, one should keep in mind that our model is a very simple, deterministic one, with more or less arbitrarily set parameters. For this reason, we consider it only as a means for illustrating our assumptions concerning the mechanisms that might be involved in the generation of multiword utterances, rather than a fully developed computational model of phonological encoding. The point we wish to make is that a few basic assumptions enable us to account for the full range of effects from phonologically related distractors in the production of single-word and multiword utterances as observed in the present experiment, in a unitary way. Furthermore, this simple model also gives a qualitative account for at least the majority of the results reported in the literature. It shows that seemingly contradictory results, like for example those reported by Miozzo and Caramazza (1999, Experiment 4) and Jescheniak et al. (2001), can be accounted for within the same framework when the size of the basic facilitation effect for bare nouns is taken into account. The most central assumption of the model is that, when generating multiple-word utterances, all elements within an advanced planning unit are phonologically activated to some degree before articulation starts, with the amount of activation decreasing from the initial to the final position.

The term *planning unit* requires some additional remarks. Within the framework of the graded activation account, any element that receives some phonological activation prior to utterance onset is part of the advance planning unit. If one obtains facilitation or inhibition for an element in a given linear position, one can be sure that the corresponding element belongs to the current planning unit. The situation is less obvious if a null effect is obtained. It could either mean that the primed element did not receive any activation prior to utterance onset (and thus was outside the scope of phonological advance planning), or it could be the net effect of facilitation during retrieval of the corresponding phonological form and of inhibition caused by the disturbance of the graded activation pattern coding the element's linear position. It appears, then, that any clear inference concerning the scope of advance planning for a position exhibiting a null effect requires that one test the next linear position (and a range of SOAs) as well. If there is an inhibitory effect at the next position, the critical

element (i.e., the element showing the null effect) was presumably within the planning scope. If the next position also shows a null effect, the critical element was most likely not part of the current planning unit.

A second remark with respect to the scope of phonological advance planning concerns the question of whether it is confined to a single phonological word. A phonological word is defined as a content word plus other function words that attach to it. Costa and Caramazza (2002) have argued that the advance planning scope exceeds the phonological word (see also Miozzo & Caramazza, 1999). The present results converge with this conclusion. Although the bare noun utterances and simple noun phrase utterances tested in our experiments consisted of just one phonological word, the complex noun phrase utterances consisted of three phonological words (e.g., “*der große // grüne // Tisch*”—the big // green // table; phonological word boundaries indicated by //). Nevertheless, our results indicate that for all three utterance formats, the phonological form of the utterance final noun has received some activation prior to utterance onset.

A final remark concerns the possibility that the phonological planning unit is stochastically variable, such that it spans only one element on some trials whereas it spans several elements on other trials (cf. Schriefers & Teruel, 1999). Under such a view, one would expect that less priming is obtained for utterance noninitial elements than for utterance-initial elements, as was in fact observed in the present experiment. The prediction holds because in the former case only a fraction of all trials would contribute to the overall priming effect, whereas in the latter case all trials would contribute to it. However, this account also predicts that priming effects for elements at very late positions should be absent at most. It cannot explain inhibition from phonologically related distractors when priming elements at late positions in an utterance. Such inhibition, however, was observed in the present experiment and confirms Meyer's (1996) earlier finding.

In summary, the present data show that the position of the phonologically primed element in an utterance appears to be a major predictor determining whether full facilitation, or reduced facilitation, or no effect or inhibition is obtained. In addition, the size of the base priming effect (as indexed by the size of the facilitation effect in bare noun production) appears to have some impact, as our post hoc analyses revealed. This full pattern can be accounted for in one theoretical framework, the graded activation account, which can be viewed as a natural extension of current models of phonological encoding (e.g., Dell, 1986; Levelt & Meyer, 2000). In the introduction, we had discussed that the different types of effects obtained in the literature have led to different and partly even contradictory theoretical accounts. In view of the data presented in this article, it appears to be likely that one account can ultimately handle the whole range of results. As was demonstrated above, even a rather simple and deterministic version of a corresponding model can handle some of the seeming inconsistencies in the literature. Nevertheless, a more decisive statement should be deferred until a more sophisticated computational model along the lines proposed above is available, replacing our “pocket calculator” model. Finally, the present experimental data show that drawing theoretical conclusions concerning the phonological encoding of multiword utterances by testing for effects of phonologically related distractors at a single linear position and a single SOA can be misleading.

## References

- Bachoud-Lévi, A.-C., Dupoux, E., Cohen, L., & Mehler, J. (1998). Where is the length effect? A cross-linguistic study of speech production. *Journal of Memory and Language*, *39*, 331–346.
- Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335–359.
- Costa, A., & Caramazza, A. (2002). The production of noun phrases in English and Spanish: Implications for the scope of phonological encoding in speech production. *Journal of Memory and Language*, *46*, 178–198.
- Damian, M. F., & Bowers, J. S. (in press). Locus of semantic interference in picture-word interference tasks. *Psychonomic Bulletin and Review*.
- Damian, M. F., & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 345–361.
- Dell, G. S. (1986). A spreading-activation model of retrieval in sentence production. *Psychological Review*, *93*, 283–321.
- Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review*, *104*, 801–838.
- Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, *42*, 287–314.
- Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey, Vol. 3: Language: Psychological and biological aspects* (pp. 69–96). New York: Cambridge University Press.
- Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 824–843.
- Jescheniak, J. D., & Schriefers, H. (1998). Serial discrete versus cascaded processing in lexical access in speech production: Further evidence from the co-activation of near-synonyms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1256–1274.
- Jescheniak, J. D., Schriefers, H., & Hantsch, A. (2001). Semantic and phonological activation in noun and pronoun production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1058–1078.
- La Heij, W., Mak, P., Sander, J., & Willeboordse, E. (1998). The gender congruency effect in picture-word tasks. *Psychological Research*, *61*, 209–219.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Science*, *3*, 223–232.
- Levelt, W. J. M., & Meyer, A. S. (2000). Word for word: Multiple lexical access in speech production. *European Journal of Cognitive Psychology*, *12*, 433–452.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75.
- Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, *98*, 122–142.
- Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture-word interference experiments. *Journal of Memory and Language*, *35*, 477–496.
- Miozzo, M., & Caramazza, A. (1999). The selection of determiners in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 907–922.
- Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological

- encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 539–557.
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107–142.
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, 64, 249–284.
- Roelofs, A., Meyer, A. S., & Levelt, W. J. M. (1996). Interaction between semantic and orthographic factors in conceptually driven naming: Comment on Starreveld and La Heij (1995). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 246–251.
- Schriefers, H. (1992). Lexical access in the production of noun phrases. *Cognition*, 45, 33–54.
- Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 841–850.
- Schriefers, H., & Jescheniak, J. D. (1999). The representation and processing of grammatical gender in language production: A review. *Journal of Psycholinguistic Research*, 28, 575–600.
- Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in production: Picture–word interference studies. *Journal of Memory and Language*, 29, 86–102.
- Schriefers, H., & Teruel, E. (1999). Phonological facilitation in the production of two-word utterances. *European Journal of Cognitive Psychology*, 11, 17–50.
- Vigliocco, G., Vinson, D. P., Martin, R. C., & Garrett, M. F. (1999). Is “count” and “mass” information available when noun is not? An investigation of tip of the tongue states and anomia. *Journal of Memory and Language*, 40, 534–558.

(Appendix follows)

## Appendix

## Object Names and Distractors Used in the Experiment

Object name	Distractor			
	SEM-REL	SEM-UNR	PHO-REL	PHO-UNR
Fuß [foot—m]	Nase [nose—f]	Säge [saw—f]	Fuge [seam—f]	Menge [quantity—f]
Hobel [plane—m]	Feile [file—f]	Villa [villa—f]	Hose [trousers—f]	Aula [assembly hall—f]
Kamm [comb—m]	Bürste [brush—f]	Decke [blanket—f]	Kanne [jug—f]	Hose [trousers—f]
Mond [moon—m]	Sonne [sun—f]	Bürste [brush—f]	Mole [mole—f]	Kiste [box—f]
Rock [skirt—m]	Kleid [dress—n]	Bein [leg—n]	Roß [horse—n]	Hirn [brain—n]
Mantel [coat—m]	Jackett [jacket—n]	Becken [basin—n]	Manko [deficit—n]	Luder [beast—n]
Hirsch [stag—m]	Pferd [horse—n]	Kleid [dress—n]	Hirn [brain—n]	Archiv [archive—n]
Arm [arm—m]	Bein [leg—n]	Pferd [horse—n]	Archiv [archive—n]	Roß [horse—n]
Geige [violin—f]	Baß [bass—m]	Hut [hat—m]	Geist [ghost—m]	Müll [garbage—m]
Mütze [cap—f]	Hut [hat—m]	Wurm [worm—m]	Müll [garbage—m]	Glanz [brightness—m]
Raupe [maggot—f]	Wurm [worm—m]	Baß [bass—m]	Rauch [smoke—m]	Geist [ghost—m]
Zange [pliers—f]	Hammer [hammer—m]	Pulli [pullover—m]	Zander [pike perch—m]	Henkel [handle—m]
Kanone [canon—f]	Gewehr [rifle—n]	Flugzeug [airplane—n]	Kanu [canoe—n]	Manko [deficit—n]
Lupe [magnifying glas—f]	Fernglas [binoculars—n]	Gewehr [rifle—n]	Luder [beast—n]	Wapper [heraldic figure—n]
Rakete [rocket—f]	Flugzeug [airplane—n]	Fernglas [binoculars—n]	Raster [screen—n]	Kanu [canoe—n]
Wanne [tub—f]	Becken [basin—n]	Jackett [jacket—n]	Wappen [heraldic figure—n]	Raster [screen—n]
Eis [ice cream—n]	Lutscher [lolipop—m]	Hammer [hammer—m]	Eimer [bucket—m]	Zander [pike perch—m]
Hemd [shirt—n]	Pulli [pullover—m]	Becher [mug—m]	Henkel [handle—m]	Eimer [bucket—m]
Schaf [sheep—n]	Hund [dog—m]	Lutscher [lolipop—m]	Schal [scalf—m]	Rauch [smoke—m]
Glas [glas—n]	Becher [mug—m]	Hund [dog—m]	Glanz [brightness—m]	Schal [scalf—m]
Messer [knife—n]	Säge [saw—f]	Nase [nose—f]	Menge [quantity—f]	Haube [bonnet—f]
Auge [eye—n]	Zunge [tongue—f]	Feile [file—f]	Aula [assembly hall—f]	Kanne [jug—f]
Kissen [pillow—n]	Decke [blanket—f]	Zunge [tongue—f]	Kiste [box—f]	Fuge [seam—f]
Haus [house—n]	Villa [villa—f]	Sonne [sun—f]	Haube [bonnet—f]	Mole [mole—f]

*Note.* Approximate English translations and gender of the German words are given in brackets. m = masculine; f = feminine; n = neuter; SEM-REL = semantically related to the object name; SEM-UNR = semantically unrelated to the object name; PHO-REL = phonologically related to the object name; PHO-UNR = phonologically unrelated to the object name.

Received November 20, 2001  
Revision received July 15, 2002  
Accepted September 13, 2002 ■