# Exploring the Time Course of Lexical Access in Language Production: Picture–Word Interference Studies

H. Schriefers, A. S. Meyer, and W. J. M. Levelt

*Max-Planck-Institute for Psycholinguistics, Nijmegen, The Netherlands*

According to certain theories of language production, lexical access to a content word consists of two independent and serially ordered stages. In the first, semantically driven stage, so-called lemmas are retrieved, i.e., lexical items that are specified with respect to syntactic and semantic properties, but not with respect to phonological characteristics. In the second stage, the corresponding wordforms, the so-called lexemes, are retrieved. This implies that the access to a content word involves an early stage of exclusively semantic activation and a later stage of exclusively phonological activation. This seriality assumption was tested experimentally, using a picture–word interference paradigm in which the interfering words were presented auditorily. The results show an interference effect of semantically related words on picture naming latencies at an early SOA ($-150$ ms), and a facilitatory effect of phonologically related words at later SOAs (0 ms, $+150$ ms). On the basis of these results it can be concluded that there is indeed a stage of lexical access to a content word where only its meaning is activated, followed by a stage where only its form is activated. These findings can be seen as empirical support for a two-stage model of lexical access, or, alternatively, as putting constraints on the parameters in a network model of lexical access, such as the model proposed by Dell and Reich. © 1990 Academic Press, Inc.

One of the most influential models of language production has been proposed by Garrett (e.g., 1976, 1980, 1988). According to this model, the formulation of a sentence involves a sequence of processes generating different levels of representation. On the basis of a preverbal representation of what the speaker wants to express the functional level representation is generated. It encodes the meanings of the lexical items and the grammatical relationships between them. Based on the functional level representation, the positional level representation is constructed, which encodes the phonological forms of the words and their order in the surface structure of the sentence.

In this model, lexical access to content words involves two distinct stages. We will refer to this position as the two-stage theory of lexical access. The first stage, lexical selection, delivers the information about lexical items needed to construct the functional level representation. The second stage, retrieval of wordforms, delivers the information needed for the construction of the positional level representation. During the first stage, lexical items are retrieved that are specified with respect to their semantic and syntactic properties. At this stage, information about the phonological forms of the words is not yet available to the processor. These phonologically unspecified lexical items have been termed lemmas by Kempen and Huijbers (1983). Only in the second stage, the wordforms, or lexemes in Kempen and Huijber's terminology, are accessed. The syntactic and semantic properties of the words are no longer available to the processor. Thus, according to Garrett's model, lexical access

proceeds in two *serially ordered* and *independent* stages.[1]

The evidence for the distinction between the functional and the positional level representation comes primarily from the comparison of certain types of speech errors, namely word exchanges and sound exchanges (e.g., Garrett, 1988). Word exchanges usually occur between phrases and sometimes even between clauses. The exchanging words almost always belong to the same syntactic category and often fulfill similar grammatical roles in their respective phrases. They are usually not similar in their wordforms (but see Dell & Reich, 1981). By contrast, sound exchanges are mostly phrase-internal and span only one or two words. In most cases, the two words participating in a sound exchange belong to different syntactic categories. Often, they are phonologically similar. In summary, word exchanges are constrained by the syntactic category and the grammatical function of the exchanging words. Sound exchanges, on the other hand, are constrained by phonological factors and by adjacency of the participating words in the surface structure of the sentence. Under the assumption that constraints on speech errors reveal computational simultaneity of linguistic information, the differences between word and sound exchange errors indicate that they occur at different levels of processing. Word exchanges arise at the functional level and sound exchanges at the positional level.

The evidence for two independent and serially ordered stages of lexical access stems primarily from constraints on an-

other type of speech errors, namely word substitutions. Most of these errors belong to one of two classes. Errors of the first class involve words that are semantically related, but not phonologically. Errors of the second class, on the other hand, involve words that are similar in form, but not in meaning. These latter errors are usually called malapropisms (see Fay & Cutler, 1977). This distinction between semantically and phonologically conditioned errors supports the proposal of two independent stages of lexical access. Semantically motivated substitutions occur during the first stage of lexical access, where the semantic characteristics of the words are taken into consideration. Malapropisms, by contrast, occur during the second stage of lexical access, where the phonological characteristics of the words are considered.

Dell and Reich (1981), however, showed that semantically and phonologically conditioned word substitutions do not represent two disjoint sets. Their error corpus includes a number of "mixed errors" involving pairs of words that are both semantically and phonologically similar, such as, for instance, the pairs "read–write" and "lobster–oyster" (Garrett, 1988, p. 87). Moreover, these errors were found to be more frequent than to be expected on the basis of a chance estimate.

Another piece of evidence that seems to challenge the two-stage model derives from the so-called lexical bias effect in sound errors, i.e., from the observation that sound errors lead to existing words of the language more often than would be expected on the basis of a chance estimate (e.g., Baars, Motley, & MacKay, 1975; Dell & Reich, 1981). If lexical access proceeds in two independent and serially ordered steps, errors in the stage of retrieval of phonological form should not lead to existing words more often than to be expected by chance (Dell & Reich, 1981).

For both phenomena, mixed errors and the lexical bias effect, accounts within the framework of the two-stage model have

---

[1] Garrett's model (1988) is not the only one assuming two distinct stages of lexical access. Similar models have been proposed by Butterworth (1980, 1989), Fay and Cutler (1977), Fromkin (1971), and Levelt (1989). Although these models differ from Garrett's account with respect to certain details (such as the access mechanisms for grammatical affixes and closed class words, see Butterworth, 1989), they basically give the same account of lexical access to content words.

been proposed. In order to account for the lexical bias effect Baars et al. (1975) have proposed a so-called output editor that checks for the correctness of the output of the second stage of lexical access. The editor is more likely to accept a phonological string that is not the intended target sequence but forms a word of the language than a sequence that does not form a word. The occurrence of mixed errors can be explained in a similar way. The language production system's monitoring system (e.g., Levelt, 1989) is presumably more likely to overlook errors which resemble the target in both form and meaning than errors that resemble the target in only one respect.

Dell and Reich (1981), however, have taken a different route, proposing a modification of Garrett's model, which we will call the network model. As in Garrett's model there are independent sets of processes generating the functional and the positional level representation. This is reflected in the differing constraints on sound versus word exchange errors, such as the syntactic category constraint on word errors and the phrasal membership constraint on sound errors. However, in contrast to Garrett's model, the lexicon is not assumed to consist of two corresponding independent sections, storing lemmas and lexemes, respectively. Instead, it is viewed as a network of interconnected nodes representing linguistic units, such as lemmas, phonological segments, and so on. In this network the flow of information is bidirectional. Lemma nodes activate phonological nodes, which in their turn feed back some of their activation to the lemma nodes to which they connect. This contrasts with a two-stage model, where the flow of information is strictly unidirectional, from lemmas to lexemes. Due to the feedback between lemmas and lexemes, phonological information can affect processing decisions at the functional level. Thus, the network model assumes independent sets of processes responsible for the generation of specific levels of representation. But there is no corresponding organization of the lexicon into independent sections. Information can, as Dell and Reich (1981, p. 627) put it, leak between different stages of language production via the lexicon. This explains the simultaneous phonological and semantic influences on word substitution errors as well as the lexical bias effect in sound errors.

Other network models of language production have recently been proposed, for instance, by Dell (1986), MacKay (1987), and Stemberger (1985a, 1985b). In order to simplify the discussion we will focus on the comparison of one two-stage model, namely the one proposed by Garrett (1988), and one network model, namely the one proposed by Dell and Reich (1981).

Both the two-stage model and the network model imply certain assumptions about the time course of semantic and phonological activation during lexical access. According to the two-stage model, lemma selection precedes wordform retrieval, and there is no feedback from the wordform to the lemma level. That is, there should be an early stage of lexical access with exclusively semantic activation and a late stage with exclusively phonological activation.[2]

On the basis of the network model, it is more difficult to derive predictions concerning the time course of lexical access, since it depends crucially on the activation and decay rate parameters used. Since activation flows from the conceptual level to the lemma and from there to the lexeme and

___

[2] Whether a two-stage theory predicts an overlap of semantic and phonological activation stages or not appears to depend on how one implements the retrieval mechanisms in these models. If one assumes that the output of lemma retrieval is an address in the phonological lexicon, then one would presumably predict no overlap (see Butterworth, 1989, for such a version of a two-stage theory). If the link between lemma and sound form is given by spreading activation (without feedback between the levels), one would predict an intermediate stage of semantic and phonological activation (see also Humphreys, Riddoch, & Quinlan, 1985).

since the activation takes some time to spread, the lemma of a word reaches its activation maximum before the lexeme. Whether the decay of activation for the lemma and the lexeme occur at about the same time, or whether, for example, the activation of the lemma decays prior to the activation of the lexeme depends on the activation and decay rate parameters of the model. A plausible prediction is that there should be an early stage at which only semantic activation can be measured. Following this, there should be a stage of both semantic and phonological activation because the feedback from the lexeme to the lemma keeps the lemma's activation "alive." Finally, there might or might not be a late stage of pure phonological activation, depending on how the activation and decay parameters are set.

To summarize, for almost all speech error data that at first sight appear to challenge the assumption of two independent levels of lexical organization and that seem to support the network model, construals of the two-stage model have been proposed that explain these findings without introducing feedback relations between lexemes and lemmas (e.g., Cutler, 1988; Garrett, 1988). But the two-stage model and the network model make differing predictions about the time course of semantic and phonological activation. A result showing sequentiality of semantic and phonological activation would support the two-stage model. Such a finding would be difficult to explain within the network model, though the model could presumably be designed in such a way that a late stage of pure phonological activation would be predicted. But such a result would at least add to the constraints on the activation and decay parameters, and the question would arise whether the network model can account for the time course functions of semantic and phonological activation without losing its power to explain the speech error evidence. The experiments reported below put the assumption of sequentiality of semantic and phonological activation to test.

## EXPERIMENTAL PARADIGM

In the experiments reported in the following sections we employed a picture–word interference paradigm. In this type of experiments, the subjects are presented with line drawings of common objects. Superimposed on each picture appears a printed interfering stimulus (IS, hereafter), which can be a word. The subjects' task is to name the pictures as quickly as possible, ignoring the printed stimuli. The naming latencies are the main dependent variable.

The experimental conditions differ in the types of relationships obtaining between the names of the pictures and the interfering stimuli. In a SEMANTIC condition, the picture names and the interfering words are semantically related to each other (as, for instance, in the IS-picture pair "cat–dog"). In a PHONOLOGICAL condition the ISs and the picture names are related in their wordforms (as in "fog–dog"). The effect of semantic or phonological relationships between the ISs and the picture names are estimated by comparing the reaction times in these conditions to those of an UNRELATED condition, in which the ISs are words that are not related to the corresponding picture names in either form or meaning (as, for example, in "roof–dog"). This latter condition can in its turn be compared to a NEUTRAL condition in which a non-linguistic sequence is presented as IS, such as a row of "x." Finally, the NEUTRAL condition can be compared to a condition without any interfering stimulus (SILENCE condition hereafter) to determine the unspecific interference effect due to the presentation of any stimuli in addition to the pictures.

In this paradigm, the presentation of semantically related ISs typically leads to an increase in the mean reaction time relative to the unrelated ISs (see Lupker, 1979; Rosinski, 1977). As Glaser and Düngelhoff

(1984) have shown, the strength of this interference effect depends on the exact timing of word and picture presentation. Glaser and Düngelhoff varied the stimulus onset asynchrony (SOA), defined as the time interval between the onset of the picture and the onset of the IS, covering a range from −400 ms (i.e., the IS onset preceded the picture onset by 400 ms) to +400 ms (i.e., the IS onset followed the picture onset by 400 ms) in steps of 100 ms. Four types of interfering stimuli were used at each SOA, namely words that were semantically related to the picture names, unrelated words, a row of "x" as neutral IS, and the picture names themselves. The semantically related words were members of the same basic category as the picture names. In an SOA range from −300 ms to +200 ms, longer naming times were obtained in the UNRELATED condition and in the SEMANTIC condition than in the NEUTRAL condition. These interference effects were most pronounced at SOA = +100 ms. Furthermore, in a time window from SOA = −100 ms to SOA = +100 ms, the interference effect was stronger in the SEMANTIC than in the UNRELATED condition.

A number of studies have investigated the effects of graphemically and/or phonologically related interfering stimuli on the picture naming latencies (e.g., Lupker, 1982). In these studies, the presentation of interfering stimuli that were phonologically related or graphemically *and* phonologically related to the picture names tended to speed the reactions relative to unrelated ISs.

Thus, both semantically and phonologically related interfering stimuli lead to specific effects relative to unrelated ones. Moreover, at least for the semantically related condition, the strength of the effect has been shown to be sensitive to the timing of the presentation of the interfering stimulus relative to the picture onset. In the experiments reported below these two characteristics of the paradigm were exploited in order to trace the time course of the activation of the lemmas and the lexemes of the picture names. Interfering stimuli were presented that were either semantically or phonologically related or unrelated to the picture names. In addition, the SOA was varied systematically within a range from −150 ms to +150 ms. We assume that a semantically related interfering stimulus only leads to a specific interference effect relative to an unrelated one if it is presented slightly before or while the lemma of the picture name is being activated. Likewise, a phonologically related interfering stimulus is taken to be effective only if presented slightly prior or during the time interval devoted to the retrieval of the lexeme of the picture name. In other words, the pattern of semantic and phonological effects across a range of SOAs should allow us to trace the time course of the activation of meaning and form of a picture name.

According to the two-stage model, lexical access to a content word includes a stage where only its meaning has been retrieved, followed by a stage where only its form is available to the processor. Thus, at early measurement points, there should only be a semantic, but no phonological, effect. At later measurement points, on the other hand, there should only be a phonological effect. The network model also predicts that at the earliest moments of lexical access to a word, only its meaning should be activated, but later the activation of the wordform should be "added" to it, rather than replacing it. Thus, at later measurement points, both a semantic and a phonological effect should be obtained. Depending on parameter settings, a network model may or may not predict a purely phonological effect at very late measurement points.

## EXPERIMENT 1

In picture–word interference experiments, the interfering stimuli are usually presented visually. For the investigation of

the issue outlined above, it seemed crucial that the subjects generated a *phonological* representation of the interfering stimulus rather than a graphemic one. Therefore, it was decided to present the interfering stimuli auditorily.

In order to test whether auditory presentation of the ISs would yield any interference at all, Experiment 1 was run, in which the pictures to be named were either presented alone (SILENCE condition) or in combination with unrelated words (UNRELATED condition). Longer naming latencies were expected for the latter than for the former condition.

The first experiment also served to select an appropriate neutral interfering stimulus. In experiments in which the ISs are presented visually, the neutral stimulus usually consists of a row of "*x*." Two potential neutral conditions were included in Experiment 1. In one of them (BLANCO condition), all pictures were coupled with the word "blanco" ("blank"). In the other condition (NOISE condition), each picture was accompanied by a stretch of white noise, whose duration matched the length of the unrelated word combined with the respective picture. A neutral condition in the auditory paradigm that would be comparable to the neutral condition in the visual paradigm should yield a mean reaction time in between the SILENCE and the UNRELATED condition.

One might expect that when the interfering stimuli are presented auditorily, the subjects might await the end of the IS presentation before initiating the naming response. In this case, the picture–word interference paradigm would be unsuitable to our purposes. In order to investigate whether or not the subjects would employ such a strategy, Experiment 1 included a condition that was identical to the NOISE condition described above, except that the length of each interfering stimulus was increased by 200 ms (NOISE200 condition). If the subjects adopt a waiting strategy, the interference effect relative to the condition

SILENCE should be stronger in the condition NOISE200 than in the condition NOISE.

A final goal of the first experiment was to test whether the strength of the interference effects would depend on the temporal coordination of picture and word presentation. The experiment included one condition in which picture and word onset coincided (SOA = 0 ms) and one condition where the presentation of the IS began 300 ms after the onset of the picture (SOA = +300 ms). On the basis of the findings from picture–word interference experiments with visual IS presentation, more pronounced interference effects were expected in the former than in the latter condition.

*Method*

*Subjects.* The experiment was run with 16 subjects, who were undergraduate students at Nijmegen University and native speakers of Dutch. They were paid Dfl. 8.50 for their participation in the experiment.

*Materials.* Thirty-two line drawings of common objects served as experimental pictures. In addition, there were 10 practice pictures. The drawings were selected from a pool of pictures we had used in other experiments on lexical access in language production. From these experiments it was known how the pictures were usually named. The names of the experimental pictures are listed in Appendix A.

*Design.* The experimental design included three crossed within-subjects factors, namely IS-TYPES with five levels (SILENCE, BLANCO, UNRELATED, NOISE, and NOISE200), SOAs with two levels (SOA = 0 ms and SOA = +300 ms), and ITEMS with 32 levels. Each subject saw each picture 10 times, five times under SOA = 0 ms and five times under SOA = +300 ms. For each occurrence of a given picture under a given SOA, the picture was paired with a different IS. Thus, each picture was presented once under each

SOA and under each of the five IS condi-
tions.

SOAs were blocked. One group of eight
subjects started with the 160 trials of the
condition SOA = 0 ms and then proceeded
to the condition SOA = +300 ms, while a
second group of subjects began with the
condition SOA = +300 ms and then was
tested under the condition SOA = 0 ms.
Hence, there was one between-subjects
factor, GROUPS, with two levels.

The pictures were randomly grouped into
five sets, four sets of six items and one set
of eight items. The pictures of each set were
combined with the five types of ISs in the
same order. They were, for instance, first
combined with an unrelated word, then
with the word "blanco," and so on. The
items of different sets were combined with
the various types of interfering stimuli in
differing orders. Across the complete set of
pictures, the order of administering the five
types of interfering stimuli was counterbal-
anced. No picture was used in two consec-
utive trials. Otherwise, the order of the pic-
tures was random. Different random se-
quences were employed in the two blocks
testing different SOAs.

*Apparatus.* The pictures were mounted
on slides and presented by means of a
Kodak slide-projector. The interfering
stimuli were recorded and digitized using a
sampling frequency of 20 kHz. They were
stored on the disk of a PDP 11/55 computer
and presented via Sennheiser MD211N
headphones. The computer controlled the
presentation of the pictures and ISs and re-
corded the naming times, measured from
the onset of a picture to the triggering of a
voicekey by the subject's response. The en-
tire session was taped using a Revox A700
recorder.

*Procedure.* The subjects were tested in-
dividually. At the beginning of the experi-
ment, they were given a booklet including
all drawings. Next to each object, the noun
was printed that the subjects in the earlier
experiments had used spontaneously most
frequently to name the depicted object. The

subjects were asked to use only these
names to refer to the pictures.

As soon as the subjects indicated that
they had seen all drawings and studied their
names, the practice block was adminis-
tered, followed by the two test blocks.
There were short pauses between the
blocks. The entire experimental session
lasted approximately half an hour.

A single test trial had the following struc-
ture. A picture was presented for 400 ms.
The onset of the IS either coincided with
the picture onset (SOA = 0 ms) or was de-
layed by 300 ms relative to the picture on-
set (SOA = +300 ms). The subject named
the picture as quickly as possible, and the
naming time was recorded. The picture pre-
sentation was followed by a pause of
2600 ms. Then the next trial began with the
presentation of a new picture.

*Results*

All responses where the subjects had
used picture names other than the expected
ones, had repaired their utterances, had
stuttered or produced mouth clicks (i.e.,
clicking or smacking non-speech sounds
produced by the lips or the tongue) were
excluded from the analyses. The same held
for response latencies from trials where
technical errors had occurred. Finally, la-
tencies longer than 1500 ms and latencies
deviating by more than two standard devi-
ations from a subject's and an item's mean
were likewise excluded. The excluded data
were replaced by estimates following Winer
(1971). A total of 285 data points (5.57% of
all data) was treated in this way.

Subsequently, the reaction times were
submitted to an ANOVA with GROUPS,
SOAs, IS-TYPES, and ITEMS as experi-
mental factors and SUBJECTS as replica-
tion dimension. Table 1 lists the mean nam-
ing latencies per IS-TYPE and SOA. Figure
1 displays for each SOA the differences be-
tween the naming latencies in the SI-
LENCE condition and each of the other
conditions. A positive value indicates inter-

TABLE 1
RESULTS OF EXPERIMENT 1: MEAN REACTION TIMES
PER SOA AND IS-TYPE (MILLISECONDS)

| IS-Type | SOA | | |
|---|---|---|---|
| | 0 | 300 | Mean |
| SILENCE | 588 | 594 | 614 |
| BLANCO | 638 | 611 | 624 |
| UNRELATED | 658 | 606 | 632 |
| NOISE | 599 | 585 | 592 |
| NOISE200 | 586 | 603 | 594 |
| Mean | 614 | 600 | 607 |

ference in a given IS condition relative to SILENCE and a negative value facilitation.

Based on the ANOVA, $F'$-ratios were computed (see, for example, Winer, 1971). A significant main effect of SOAs was obtained ($F'(1,22) = 4.9$, $MS_e = 52755$, $p < .05$). The main effect of IS-TYPES was also significant ($F'(4,173) = 10.7$, $MS_e = 38159$, $p < .01$), as was the interaction of IS-TYPES and SOAs ($F'(4,151) = 4.2$, $MS_e = 68250$, $p < .05$). The analysis of simple effects revealed that the effect of IS-TYPES



FIG. 1. Results of Experiment 1. Mean reaction time differences (ms) for NOISE, NOISE200, and BLANCO from SILENCE at SOA = 0 ms and SOA = +300 ms.

was only significant at SOA = 0 ms ($p < .01$). Newman-Keuls tests showed that, at SOA = 0 ms, significantly longer reaction times were obtained in the conditions BLANCO and UNRELATED than in NOISE, NOISE200, and SILENCE ($p < .05$). The latter three conditions did not differ significantly from each other. At SOA = +300 ms the five IS-TYPES did not differ significantly from each other.

*Discussion*

At SOA = 0 ms, the mean reaction time was by 70 ms longer in the UNRELATED condition than in SILENCE. This finding indicates that the auditory presentation of unrelated words interfered with the naming of simultaneously presented pictures, as has been demonstrated for visual presentation of the interfering stimuli.

The presentation of the word "blanco" also led to significant interference at SOA = 0 ms, but this effect was less pronounced than the effect in the UNRELATED condition. The mean reaction times in the two NOISE conditions and in SILENCE, on the other hand, did not differ significantly from each other. On the basis of this pattern of results, it was decided to include BLANCO as a neutral condition in the following experiment.

At SOA = +300 ms, no interference effects were obtained. Thus, the pattern of interference effects depended on the timing of the ISs relative to the pictures. This pattern of findings is in agreement with results from picture–word interference studies with visual IS presentation (e.g., Glaser & Düngelhoff, 1984). On the basis of these findings it was concluded that SOAs of +300 ms or longer should not be tested in the following experiments.

Since neither of the two NOISE conditions yielded any reliable interference effect relative to the SILENCE condition, these conditions cannot be referred to in determining whether or not the subjects
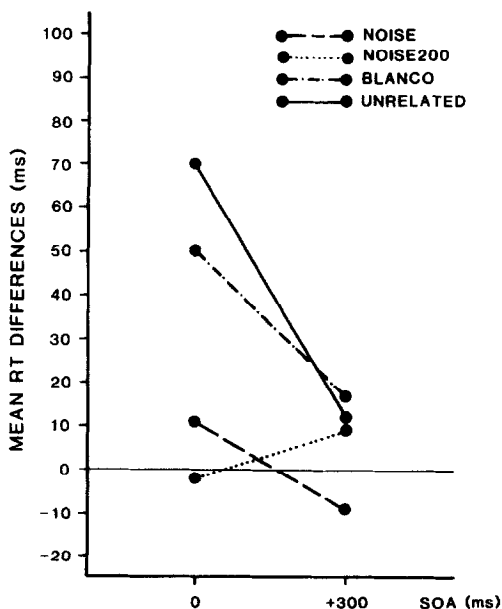
awaited the end of the IS before initiating their response. But there is other evidence suggesting that they probably did not pursue such a strategy. If the subjects postponed their response until after the end of the IS, the mean naming times for the pictures in the UNRELATED condition should correlate positively with the length of the interfering stimuli. This was, however, not the case. The correlation coefficients were $r = -.13$ and $r = .04$ for SOA $= 0$ ms and SOA $= +300$ ms, respectively. Furthermore, if the subjects awaited the end of the IS before naming the picture, the latencies should be longer at SOA $= +300$ ms than at SOA $= 0$ ms. However, the opposite was true. In the UNRELATED as well as in the BLANCO condition, the mean latencies were *shorter* at SOA $= +300$ ms than at SOA $= 0$ ms.

The results of Experiment 1 indicate that auditory presentation of interfering stimuli yields results similar to visual presentation. This holds for the relative strength of the interference effects obtained from unrelated and neutral stimuli as compared to a picture-alone condition and for the dependency of the effects on the temporal coordination of IS and picture presentation. The response latencies were found to be independent of the length of the interfering stimuli, indicating that the subjects did not await the end of the IS presentation before initiating their response. Thus, the paradigm seems well suited to investigate the time course of semantic and phonological activation in lexical access.

## EXPERIMENT 2

Experiment 2 tested the prediction derived from the two-stage model that lexical access to content words proceeds in two successive stages, a semantic and a phonological stage.

### Method

*Subjects.* The experiment was run with 48 paid subjects, who were undergraduate students at Nijmegen University.

*Materials.* Sixteen of the drawings used in Experiment 1 served as experimental pictures in Experiment 2. These were the pictures that in earlier studies had been spontaneously labeled using one and the same name by the largest proportions of subjects. Another six drawings were used as practice pictures.

Each experimental picture was presented under five conditions, namely without any interfering stimulus (SILENCE condition), together with the word "blanco" (BLANCO condition), with an unrelated word (UNRELATED condition), with a phonologically related word (PHONOLOGICAL condition), and with a semantically related word (SEMANTIC condition).

In an earlier experiment, using the same pictures, the SEMANTIC ISs had been nouns that were high associates of the picture names. Some of them were abstract, some of them concrete. Six SOAs were tested, ranging from $-450$ ms to $+300$ ms. In that experiment, there was no significant interference effect of the SEMANTIC ISs compared to the UNRELATED ones at any of the SOAs tested. Using a visual–visual picture–word interference task, Lupker (1979) also failed to obtain semantic interference effects for ISs that were associates of the target names and/or abstract nouns. By contrast, he found interference effects for concrete nouns that denoted members of the same semantic category as the target names. On the basis of these findings we decided to use only concrete nouns as interfering stimuli in the UNRELATED, PHONOLOGICAL, and SEMANTIC conditions of Experiment 2. The words in the UNRELATED condition were the same as in Experiment 1, except for one word that was now used as a semantically related IS for another picture and was replaced by a new word. The phonologically related ISs shared two or more word-initial phonological segments with the picture names and, if

possible, had the same number of syllables and the same stress pattern. Finally, the words in the SEMANTIC condition were the names of members of the same semantic category as the corresponding target names. The stimulus materials are listed in Appendix B.

*Design.* The interfering stimuli were presented at three SOAs. The IS preceded the picture onset by 150 ms (SOA = −150 ms), or IS and picture onset coincided (SOA = 0 ms), or the presentation of the IS began 150 ms after the onset of the picture (SOA = +150 ms). There were three groups of 16 subjects. Each group was tested under one SOA condition. Each subject was presented five times with each experimental picture, each time combined with a different type of interfering stimulus. Across the complete set of pictures, the order of testing the interference conditions was counterbalanced as described above (see Experiment 1). The experimental pictures were presented in the same order to all subjects.

*Apparatus.* The pictures were presented on an electronic display (Vector General), rather than being mounted on slides. Otherwise, the same equipment was used as in Experiment 1.

*Procedure.* The procedure was the same as in Experiment 1, except that the experimental session only included one instead of two test blocks. The session took about 15 min.

## Results

190 data points (4.95% of all data) were eliminated and replaced by estimates following the same criteria as in Experiment 1. The reaction times were submitted to an ANOVA with SOAs as between-subjects factor and IS-TYPES and ITEMS as within-subjects factors. Table 2 lists the mean naming latencies per IS-TYPE and SOA. Figure 2 displays the differences between the naming latencies in the SILENCE condition and each of the other conditions.

TABLE 2
RESULTS OF EXPERIMENT 2: MEAN REACTION TIMES
PER SOA AND IS-TYPE (MILLISECONDS)

| IS-Type | SOA | | | |
|---|---|---|---|---|
| | −150 | 0 | +150 | Mean |
| SILENCE | 584 | 579 | 566 | 577 |
| BLANCO | 595 | 612 | 593 | 600 |
| UNRELATED | 629 | 656 | 609 | 631 |
| SEMANTIC | 668 | 668 | 614 | 651 |
| PHONOLOGICAL | 629 | 609 | 547 | 595 |
| Mean | 621 | 625 | 586 | 611 |

Based on the ANOVA, $F'$-ratios were computed (see Winer, 1971). Significant main effects were obtained for SOAs ($F'(2,48)$ = 3.3, $MS_e$ = 182546, $p < .05$) and IS-TYPES ($F'(4,85)$ = 13.7, $MS_e$ = 50106, $p < .001$). The interaction of SOAs and IS-TYPES was also significant ($F'(9,300)$ = 6.1, $MS_e$ = 13962, $p < .01$). The analysis of simple effects revealed that the effect of IS-TYPES was significant for each SOA.

Newman-Keuls tests showed that at SOA = −150 ms the presentation of unrelated, phonological, and semantic interfering stimuli led to significant interference relative to the SILENCE condition ($p < .01$). Also, significant interference relative to the BLANCO condition was obtained in the UNRELATED, PHONOLOGICAL, and SEMANTIC condition ($p < .01$). Finally, the mean reaction time was significantly longer in the SEMANTIC condition than in the UNRELATED and in the PHONOLOGICAL conditions ($p < .01$), for which identical means were obtained.

At SOA = 0 ms, the mean reaction time in SILENCE was significantly shorter than the means in all other conditions ($p < .01$). The mean latencies in the UNRELATED and in the SEMANTIC conditions were significantly longer than the mean in the BLANCO condition ($p < .01$), but the former two conditions did not differ significantly from each other anymore. The mean reaction time in the PHONOLOGI-
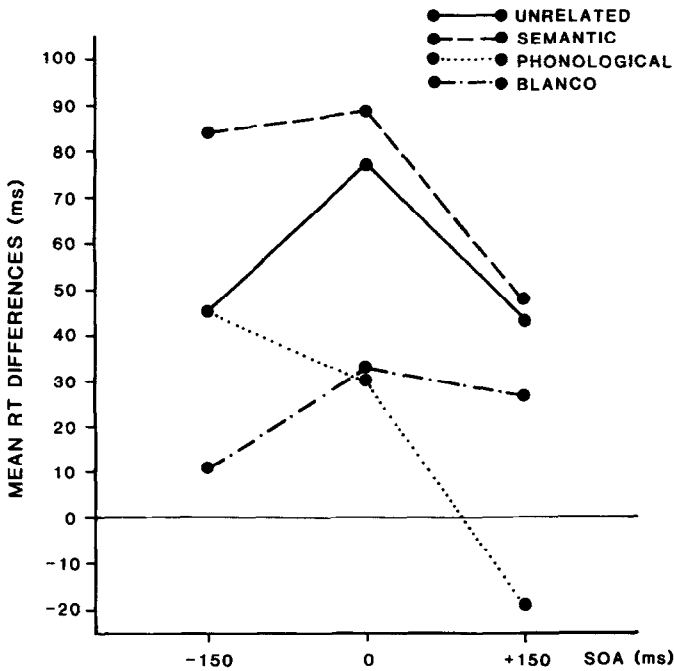
FIG. 2. Results of Experiment 2. Mean reaction time differences (ms) for UNRELATED, SEMAN-TIC, PHONOLOGICAL, and BLANCO from SILENCE at three SOAs.

CAL condition was significantly shorter than the mean in the UNRELATED condition ($p < .01$). In other words, at this SOA, the phonologically related ISs led to facilitation relative to the unrelated ones.

Finally, at SOA = +150 ms, facilitation in the PHONOLOGICAL condition was observed relative to the UNRELATED, SEMANTIC, and BLANCO conditions ($p < .01$). The latter three conditions yielded significant interference effects relative to SILENCE ($p < .05$). Finally, the UNRE-LATED and SEMANTIC conditions did not differ significantly from each other.

*Discussion*

At SOA = −150 ms, the mean reaction time was significantly longer in the SE-MANTIC condition than in the UNRE-LATED condition, whereas the PHONO-LOGICAL and the UNRELATED condi-tions did not differ from each other. This pattern can be taken as indication of an early stage of pure semantic activation. At SOA = 0 ms and at SOA = +150 ms, no semantic interference effect was obtained any more. At those SOAs, a significant fa-cilitatory effect in the PHONOLOGICAL condition relative to the UNRELATED condition was observed. These results sup-port the notion of a stage of pure phonolog-ical activation, following the stage of pure semantic activation.

Thus, the findings are in agreement with the two-stage model of lexical access. There is, however, one caveat. The effect of a phonologically related interfering stim-ulus can be attributed fairly unambiguously to the processing stage during which the wordform of a picture name is retrieved. The effect of a semantically related stimu-lus *can* be assigned to the preceding stage, the retrieval of the lemma. However, it could also be argued that this effect arose earlier, namely during the preverbal stage, in which the depicted objects were per-ceived and categorized. In this case, our findings would not allow us to draw any

conclusions concerning the time course of lemma and lexeme activation.

## EXPERIMENT 3

In order to determine the locus of the semantic interference effect Experiment 3 was run. In this experiment, a task was chosen that required the perception and categorization of the pictures, but no naming. In the presentation phase at the beginning of the experiment, the subjects saw a set of pictures, which they should try to remember. In the subsequent recognition phase, they were presented with a series of pictures, half of which had been shown in the presentation phase. The subjects' task was to indicate for each picture whether or not they had seen it before by pressing one of two buttons. The dependent variable was the reaction time from the onset of a picture to the subject's push-button response. During the recognition phase each picture was paired with an interfering word that was either semantically related to the picture name or unrelated to it. The interfering stimuli were presented auditorily with an SOA of − 150 ms. At this SOA, a specific interference effect from semantically related ISs relative to unrelated ones had been obtained in Experiment 2. If such an effect is not obtained in the present experiment, we can conclude that the semantic interference effect obtained in the previous experiment can be attributed to the stage of lemma retrieval, rather than to the visual and conceptual processing of the pictures.

### Method

*Subjects.* Twenty undergraduate students of Nijmegen University served as subjects.

*Materials.* Thirty-two line drawings were used in Experiment 3, most of which had also been used in Experiment 1. A set of 16 pictures was presented both in the presentation and in the recognition phase (set I), while 16 other pictures were only shown in the recognition phase (set II). The latter set, to which the correct response was to

push the "no" button, was composed of the experimental pictures from Experiment 2. They represented the critical items, which were included in the statistical analyses. The critical items were not shown in the presentation phase in order to rule out any effect of implicit verbal encoding of the pictures in this phase.

Each picture was combined with an unrelated and a semantically related interfering stimulus. The pictures of set II were paired with the ISs of the SEMANTIC and UNRELATED conditions of Experiment 2. For the pictures of set I, new unrelated and semantically related interfering stimuli were selected.

In addition, there was a practice set of six pictures. Three of them were shown in the presentation and in the test phase and three only in the test phase. All practice pictures were combined with unrelated words as ISs.

*Design.* As mentioned before, an SOA of − 150 ms was used. The experiment was run with two groups of subjects. The set of critical pictures was split up into two subsets of eight pictures each. For the first group of subjects, the first subset of critical pictures was paired with SEMANTIC ISs and the second subset with UNRELATED ones. Conversely, for the second group, the first subset was paired with UNRELATED ISs, and the second with SEMANTIC ones. Each subject saw each picture only once during the recognition phase in order to avoid a situation in which a picture that had not been shown in the presentation phase, but had already appeared earlier during the recognition phase would have to be categorized as "new."

*Procedure.* In the presentation phase, the pictures of set I were displayed, one after the other, for 1500 ms each with pauses of 1000 ms. After a short break, the recognition phase began. The structure of the test trials was the same as in the preceding experiments, except that, rather than naming the pictures, the subjects in each trial pressed one of two buttons, thereby indi-

cating whether or not they had seen the respective picture before.

*Results*

Only the reaction times from the critical trials were evaluated. Latencies from incorrect responses were eliminated. Furthermore, extreme reaction times were excluded according to the same criteria as in the preceding experiments. The eliminated latencies were again substituted by estimates. Sixty-one data points (1.9% of all reaction times to critical pictures) were treated in this way.

For the subject ANOVA within-subject means were computed across the eight items of a subset. These means were analyzed in an ANOVA with GROUPS as between-subjects factor and IS-TYPES as within-subjects factor. For the item ANOVA the means over subjects for each picture in a given condition (UNRELATED vs. SEMANTIC) were computed and submitted to an ANOVA with IS-TYPES as a within-item factor. The mean reaction times were 645 ms and 640 ms for the UNRELATED and the SEMANTIC conditions, respectively. This difference was not significant (both $Fs < 1$). The effect of GROUPS and the interaction of IS-types and GROUPS, which could only be evaluated in the subject analysis, were not significant either ($F1(1,18) = 1.3$, $MS_e = 14263$, and $F1(1,18) < 1$, $MS_e = 634$, respectively).[3]

---

[3] In Experiments 1 and 2, each critical item was presented several times under varying experimental conditions. In the recognition phase of Experiment 3, on the other hand, each critical item was presented only once. In order to investigate the effect of the repetition of critical items in the recognition experiment, the experiment was repeated with the same subjects right after the end of the experimental session described above. Both the presentation and the recognition phase were repeated, and in each phase the same set of pictures was presented as in the first presentation and recognition phase. Again, no effect of IS-TYPES was obtained ($F1(1,18) < 1$, $MS_e = 1985$; $F2(1,15) = 2.2$, $MS_e = 615$, $p > .10$), and the interaction of IS-TYPES and GROUPS was not significant

*Discussion*

The semantic interference effect obtained in Experiment 2 completely disappeared in Experiment 3, where the subjects engaged in a different type of task, namely categorizing the pictures as "seen before" vs. "new" rather than naming them. In other words, the semantic interference effect is only obtained if the subjects name the pictures. This supports the attribution of the semantic interference effect of Experiment 2 to the stage of lemma retrieval.

GENERAL DISCUSSION

Three experiments were run in order to trace the time course of semantic and phonological activation during lexical access in language production. We employed a picture–word interference task with auditory presentation of the interfering stimuli. Experiment 1 established some basic characteristics of this paradigm. In Experiment 2 we found a specific semantic interference effect when the ISs were presented early (SOA = −150 ms), but not when they were presented later (SOA = 0 ms and SOA = +150 ms; see also Smith & Magee, 1980, for results suggesting that a picture activates the lemma more rapidly than the phonological form). A different pattern of results was obtained for the phonologically related ISs. There was no specific phonological effect when the ISs were presented at the earliest SOA (−150 ms), but at the remaining two SOAs facilitatory effects were obtained. The finding that the phonological effect was confined to late SOAs had also been found in a different experi-

either ($F1(1,18) = 2.2$, $MS_e = 1985$, $p > .10$). In the joint analysis of variance for both replications of the experiment, a main effect of REPLICATIONS was obtained (means: 642 and 596 ms for replications 1 and 2, $F1(1,18) = 8.5$, $MS_e = 4918$, $p < .01$; $F2(1,15) = 32.9$, $MS_e = 1020$, $p < .001$). The effect of IS-TYPES ($F1(1,18) < 1$, $MS_e = 1919$; $F2(1,15) = 1.8$, $MS_e = 707$, $p > .10$) and the interaction of IS-TYPES with REPLICATIONS ($F1(1,18) < 1$, $MS_e = 699$; $F2(1,15) < 1$, $MS_e = 1691$) were not significant.

ment, which was not reported here, and in which SOAs ranging from −450 ms to +300 ms had been tested. In Experiment 3, the subjects categorized the pictures as to whether or not they had seen them before, rather than naming them. The specific semantic interference effect obtained before disappeared, indicating that the presentation of the ISs affected the retrieval of the lexical items for the naming responses rather than the recognition of the pictures (see also Smith & Magee, 1980, for similar results).

Thus, the presentation of semantically related ISs inhibited the retrieval of the picture names, while the presentation of phonologically related ISs facilitated it. This pattern of results agrees with other findings reported in the literature (e.g., Lupker, 1979, 1982; Rosinsky, 1977). There is as yet no satisfactory explanation for why phonological and semantic ISs yield such strikingly different effects. Our proposal concerning this issue is based on the assumption that a lemma is retrieved as one unit. This does not mean necessarily that the representation of lemmas is non-compositional. A lemma might consist of components like semantic features. However, lemmas are not constructed by assembling these semantic features. Rather, the semantic features of a lemma serve as the conditions that have to be fulfilled by the preverbal input for the language production system in order to activate the lemma (Levelt, 1989). Lexemes, by contrast, are created by selecting and combining sublexical units. At least the second part of this supposition is supported by independent evidence. First, one of the most frequent types of speech errors are sound errors, i.e., errors in which a segment or a string of segments not corresponding to a complete morpheme is misselected or misplaced. It would be difficult to explain how such errors could arise if wordforms were retrieved as entities (see, for instance, Fromkin, 1971; Dell, 1986; Levelt, 1989). Second, there is experimental evidence

suggesting that lexemes are constructed in a series of processing steps, in each of which one phonological segment or a cluster is selected and assigned to its position in a syllable frame (Meyer, 1988). Thus, we assume that in lexical access to a content word first a single lemma is selected on the basis of conceptual information and then a number of phonological segments and/or clusters are retrieved and combined to form the corresponding lexeme.

Under this view, the ISs in picture–word interference experiments represent competitors to the target lemmas, rendering the selection of the correct lemmas more difficult than in a situation without IS presentation. The specific semantic interference effect reflects the fact that a target lemma and a semantically related IS consist partly of the same semantic features. Therefore, more detailed semantic information has to be available for successful selection of the target lemma in the presence of an IS constituting a semantically related competitor than in the presence of an unrelated IS. This should make semantically related ISs more potent competitors than unrelated ones. A different situation arises when a phonologically related IS is presented. Again there will be competition at the lemma level, just as in the case of the presentation of an unrelated IS. However, phonological related ISs share some segments with the corresponding picture names; hence, some of the units out of which the target wordforms are to be constructed already become activated when the IS is presented. The facilitatory effect of phonologically related ISs can be explained by assuming that due to the preactivation these segments can be selected more rapidly than in the control condition in which unrelated ISs are presented, so that the generation of the lexeme is speeded up and the response latency is reduced. Clearly, more evidence is needed to resolve the riddle of the opposing effects of phonological and semantic similarity in interference tasks.

With respect to our main point of interest, the time course of semantic and phonological activation, our data are in agreement with a two-stage model of lexical access. In such a model, lexical access proceeds in two serially ordered steps, retrieval of lemmas and retrieval of phonological forms, which do not affect each other. If such a view is adopted, the occurence of mixed errors, the lexical bias effect, and other findings that seem to suggest interactions between the lemma level and the level of phonological form must be explained by reference to additional theoretical notions, such as an output editor or a monitor.

It is an open issue whether or not the present data can also be accounted for in the network model. Butterworth (1989) has argued that evidence for strict seriality of semantic and phonological activation poses a serious problem for the network model of lexical access because in that model the levels of processing are "inextricably linked in time" due to the bidirectional links between the lemma level and the level of phonological form. Thus, according to this view our data would rule out such models. We would, however, like to be somewhat more conservative in our estimate of the implications of the present findings for the network model. As was mentioned in the Introduction, it might be possible to set the parameters of a network model in such a way that a stage of pure semantic activation and a stage of pure phonological activation result. Between these two stages, however, there must be a stage at which both form and meaning of a word are activated to some extent. Our data do not provide any evidence for such a stage; but this obviously does not rule out its existence. One might argue that if we had varied the SOAs in smaller steps we might have found both semantic and phonological activation at some point. From the perspective of the network model the present data can be viewed as an additional set of constraints on the setting of the activation and decay parameters. The crucial question is whether these parameters can be set in such a way that the present set of data as well as other evidence, such as the results from speech error analyses, receive a satisfactory explanation.

## APPENDIX A: MATERIALS FROM EXPERIMENT 1

| Picture name | Interfering stimulus |
|---|---|
| bureau | muts |
| cactus | tas |
| fietspomp | wip |
| geweer | koets |
| hark | bel |
| ijsje | brief |
| klok | film |
| knoop | zwaard |
| krokodil | muur |
| radio | kerk |
| schroef | taart |
| sigaar | poes |
| thermometer | noot |
| veer | slot |
| vinger | kwast |
| zak | koe |
| tractor | monnik |
| varken | raam |
| pakje | tunnel |
| bank | schaats |
| fluit | wapen |
| pijl | step |
| iglo | worst |
| hoefijzer | brug |
| caravan | stempel |
| lucifer | peddel |
| paperclip | hoed |
| pan | molen |
| spijker | krans |
| spons | kies |
| zeilboot | spin |
| badkuip | appel |

*Translation*

| Picture name | Interfering stimulus |
|---|---|
| desk | cap |
| cactus | bag |
| bicycle pump | seesaw |
| rifle | carriage |
| rake | bell |
| icecreamcone | letter |
| clock | film |
| button | sword |

| Picture name | Interfering stimulus | Picture name | Interfering stimulus |
|---|---|---|---|
| crocodile | wall | flute | weapon |
| radio | church | arrow | step |
| screw | cake | igloo | sausage |
| cigar | cat | horseshoe | bridge |
| thermometer | nut | caravan | stamp |
| feather | lock | match | pedal |
| finger | paintbrush | paperclip | hat |
| sack | cow | pan | windmill |
| tractor | monk | nail | wreath |
| pig | window | sponge | tooth |
| parcel | tunnel | sailboat | spider |
| bench | skate | bathtub | apple |

## APPENDIX B: MATERIALS FROM EXPERIMENT 2

| Picture name | IS-TYPE | | |
|---|---|---|---|
| | UNRELATED | SEMANTIC | PHONOLOGICAL |
| bureau | muts | kast | buurman |
| cactus | doek | dadel | kakkerlak |
| fietspomp | wip | ventiel | file |
| geweer | koets | pistool | gewei |
| hark | bel | gieter | harp |
| ijsje | brief | vla | ijzer |
| klok | film | horloge | klos |
| knoop | zwaard | gesp | knook |
| krokodil | muur | nijlpaard | krokus |
| radio | kerk | televisie | radijs |
| schroef | taart | moer | schroot |
| sigaar | poes | pijp | citroen |
| thermometer | noot | windvaan | terrier |
| veer | slot | pels | veen |
| vinger | kwast | teen | vink |
| zak | koe | tas | zalf |

*Translation*

| Picture name | IS-TYPE | | |
|---|---|---|---|
| | UNRELATED | SEMANTIC | PHONOLOGICAL |
| desk | cap | closet | neighbor |
| cactus | cloth | date | cockroach |
| bicycle pump | seesaw | valve | queue |
| rifle | carriage | pistol | horns |
| rake | bell | watering can | harp |
| icecreamcone | letter | custard | iron |
| clock | film | watch | chock |
| button | sword | buckle | knuckle |
| crocodile | wall | hippopotamus | croquette |
| radio | church | TV set | radish |
| screw | cake | nut | scrap |
| cigar | cat | pipe | lemmon |
| thermometer | nut | weathervane | terrier |
| feather | lock | fur | peat |
| finger | paintbrush | toe | finch |
| sack | cow | bag | ointment |

### REFERENCES

BAARS, B. J., MOTLEY, M., & MACKAY, D. G. (1975). Output editing for artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior,* 14, 382–391.

BUTTERWORTH, B. (1980). Some constraints on models of language production. In B. Butterworth (Ed.), *Language production: Vol. 1. Speech and talk.* New York: Academic Press.

BUTTERWORTH, B. (1989). Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical representation and process.* Cambridge, MA: MIT Press.

CUTLER, A. (1988). The perfect speech error. In L. M. Hyman & C. N. Li (Eds.), *Language, speech and mind: Studies in honour of Victoria A. Fromkin.* London: Routledge.

DELL, G. (1986). A spreading activation theory of retrieval in language production. *Psychological Review,* 93, 283–321.

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

FAY, D., & CUTLER, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry,* 8, 505–520.

FROMKIN, V. (1971). The non-anomalous nature of anomalous utterances. *Language,* 47, 27–52.

GARRETT, M. (1976). Syntactic processes in language production. In R. Wales & E. Walker (Eds.), *New approaches to language mechanisms. A collection of psycholinguistic studies.* Amsterdam: North-Holland.

GARRETT, M. (1980). Levels of processing in language production. In B. Butterworth (Ed.), *Language production: Vol. 1. Speech and talk.* London: Academic Press.

GARRETT, M. (1988). Processes in language production. In: F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey. Vol. III. Biological and psy-*

*chological aspects of language.* Cambridge, MA: Harvard University Press.

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.

HUMPHREYS, G. W., RIDDOCH, M. J., & QUINLAN, P. T. (1988). Cascade processing in picture identification. *Cognitive Neuropsychology,* 5, 67–103.

KEMPEN, G., & HUIJBERS, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition,* 14, 185–209.

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

LUPKER, S. J. (1979). The semantic nature of response competition in the picture–word interference task. *Memory and Cognition,* 7, 485–495.

LUPKER, S. J. (1982). The role of phonetic and orthographic similarity in picture word interference. *Canadian Journal of Psychology,* 36, 349–367.

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

MEYER, A. S. (1988). *Phonological encoding in language production. A priming study.* Unpublished doctoral dissertation, University of Nijmegen.

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

SMITH, M. C., & MAGEE, L. E. (1980). Tracing the time course of picture-word processing. *Journal of Experimental Psychology: General,* 109, 373–392.

STEMBERGER, J. P. (1985a). *The lexicon in a model of language production.* New York: Garland.

STEMBERGER, J. P. (1985b). An interactive model of language production. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 1). London: Erlbaum.

WINER, J. B. (1971). *Statistical principles in experimental design.* New York: McGraw–Hill.