



ELSEVIER

Cognition 69 (1999) 313–335

COGNITION

Lexical access in the production of pronouns

Bernadette M. Schmitt^{a,*}, Antje S. Meyer^b, Willem J.M. Levelt^b

^a*Department of Psychology/Neurocognition, Maastricht University, P.O. Box 616,
6200 MD Maastricht, The Netherlands*

^b*Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands*

Received 26 May 1998; accepted 13 November 1998

Abstract

Speakers can use pronouns when their conceptual referents are accessible from the preceding discourse, as in ‘*The flower is red. It turns blue*’. Theories of language production agree that in order to produce a noun semantic, syntactic, and phonological information must be accessed. However, little is known about lexical access to pronouns. In this paper, we propose a model of pronoun access in German. Since the forms of German pronouns depend on the grammatical gender of the nouns they replace, the model claims that speakers must access the syntactic representation of the replaced noun (its lemma) to select a pronoun. In two experiments using the lexical decision during naming paradigm [Levelt, W.J.M., Schriefers, H., Vorberg, D., Meyer, A.S., Pechmann, T., Havinga, J., 1991a. The time course of lexical access in speech production: a study of picture naming. *Psychological Review* 98, 122–142], we investigated whether lemma access automatically entails the activation of the corresponding word form or whether a word form is only activated when the noun itself is produced, but not when it is replaced by a pronoun. Experiment 1 showed that during pronoun production the phonological form of the replaced noun is activated. Experiment 2 demonstrated that this phonological activation was not a residual of the use of the noun in the preceding sentence. Thus, when a pronoun is produced, the lemma and the phonological form of the replaced noun become reactivated. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Speech production; Lexical access; Nouns and pronouns; Lexical decision; Picture naming

* Corresponding author. Tel.: +31-43-3881934; fax: +31-43-3671850;
e-mail: b.schmitt@psychology.unimaas.nl

1. Introduction

Speaking, most of the time, means transforming an idea or a message into a structured pattern of sounds. The speaker's goal is to produce words so that a listener can understand the message. In many theories, the transfer from ideas to words involves three main steps: message encoding, grammatical encoding, and phonological encoding (Garrett, 1975, 1980, 1988; Stemberger, 1985; Dell, 1986; Levelt, 1989, 1992; Dell and O'Seaghdha, 1991, 1992; Bock and Levelt, 1994; Bock, 1995; Levelt et al., in press). During message encoding the intended meaning of the utterance is specified prelinguistically. The message serves as input to the grammatical level of processing. During grammatical encoding, words are retrieved from the mental lexicon via a process called lexical access. As a result of lexical access, the syntactic properties of words, i.e. their lemmas, become available. The selected elements are combined into a syntactic structure. This structure is the input to the stage of phonological encoding, during which the phonological (segmental, prosodic) and phonetic shape of the utterance are generated. After the phonetic form has been computed, articulation can take place. Empirical evidence consistent with the existence of several levels of speech planning comes from speech-error data (Garrett, 1975, 1980, 1988; Dell, 1986; Martin et al., 1996), picture-naming studies (for a review see Schriefers et al., 1990; Glaser, 1992; Levelt et al., 1991a,b, in press; Peterson and Savoy, 1998), and electrophysiological data (Van Turenout et al., 1997, 1998). Critical discussions of the notion of a separate syntactic (lemma) level are offered by Caramazza (1997) and Starreveld and La Heij (1995, 1996).

Most experiments on lexical access have studied the access to single nouns (e.g. Schriefers et al., 1990; Levelt et al., 1991a,b; Roelofs, 1992a,b; Starreveld and La Heij, 1995, 1996; O'Seaghdha and Marin, 1997; Peterson and Savoy, 1998). A few have investigated slightly more complex utterances, such as adjective noun phrases (Schriefers, 1993) and noun phrase coordinations (Meyer, 1996, 1997; Meyer et al., 1998). In the present study, we addressed the generation of pronouns in utterances such as '*The flower is red. It turns blue*'. The participants in our experiments generated such sentence pairs during a picture description task. As will be explained below, our main goal was to determine whether the phonological form of the antecedent noun is activated when it is replaced by a pronoun. We will, however, first propose a working model of lexical access to pronouns.

2. Four steps from intention to pronoun access

We propose that lexical access to a pronoun involves the following four steps.

2.1. From intention to preverbal concept

The generation of any utterance begins with the transformation of a communicative intention into a preverbal conceptual message (Levelt, 1989). As part of this process, the speaker must fit each message fragment into the current discourse. This

involves, among other things, marking certain concepts as old information, or in focus, and others as new or not in focus (e.g. Chafe, 1976). Focused information and information that is not in focus can be referred to linguistically in different ways, for instance by full noun phrases (for information that is not in focus) or pronouns (for information that is in focus). Such marking helps speakers and listeners maintain a coherent discourse record (e.g. Marslen-Wilson et al., 1982; Levelt, 1989; Ericsson and Kintsch, 1995). A formal way to describe how linguistic devices, such as pronouns, affect discourse coherence has been developed within the framework of centering theory in computational linguistics (Joshi and Weinstein, 1981; Grosz et al., 1983; Gordon et al., 1993; Gordon and Scarsee, 1995; see also Gordon and Hendrick, 1997).

How speakers decide which concepts are in focus and which are not, and which discourse devices should be used to signal this to their listeners is beyond the scope of the present paper. As part of our working model of lexical access to pronouns we simply assume that, depending on the discourse context, lexical concepts are marked as being in focus or not being in focus. A pronoun is selected if, and only if, the corresponding lexical concept is active and is marked as in focus within the discourse. Otherwise, a noun is selected. In our working model, we represent lexical concepts as a network of conceptual nodes that get input from a variety of sources, including the visual system (e.g. Roelofs, 1992a,b). Discourse information is represented as nodes that get input from the conceptual layer. Both, the conceptual and the discourse layer are connected to the lemma layer (for details see Schmitt, 1997).

2.2. *Lexical selection*

The encoded message (activated concepts and discourse information) serves as the input to the grammatical encoding stage (Levelt, 1989). One main step during grammatical encoding is lexical selection. Lexical selection is the retrieval of lemmas, i.e. syntactic words, from the mental lexicon (Kempen and Huijbers, 1983).

The message activates the corresponding lemmas, which in turn give access to syntactic information about words. Lemmas specify the syntactic categories (noun, verb, etc.) of words. Many lemmas carry diacritic parameters that must be set. English verbs, for instance, have parameters for number, person, tense, and mode. German nouns are specified for grammatical gender; they are either masculine, feminine, or neuter.

2.3. *Gender access*

Though there are certain correlations between conceptual properties of the referents and the grammatical gender of their German names, and between grammatical gender and phonological form of nouns, the gender of a noun cannot be reliably predicted on the basis of these properties. Therefore, theories of lexical access commonly assume that the grammatical gender is part of the stored syntactic information about the nouns (Roelofs, 1992a,b; Schriefers, 1993; Jescheniak and Levelt,

1994; Van Berkum, 1996, 1997). In our working model, gender is represented in gender nodes, which are linked to lemmas. As explained elsewhere, we assume that gender information becomes activated when a noun lemma is activated, but is selected only when it is required to generate a well-formed utterance (Roelofs et al., 1996, 1998).

2.4. Pronoun access

One of the cases where a gender node must be selected is the generation of German pronouns. Access of the gender node is necessary because the gender of the pronoun must correspond with the gender of the antecedent. If the antecedent is feminine, the appropriate pronoun in nominative case is 'sie', as in 'Die Blume ist rot. Sie wird blau.' (*The flower is red. It turns blue*). A masculine noun, such as 'der Klee' (*the clover*), must be referred to by 'er' (*he*), and a neuter noun, such as 'das Veilchen' (*the violet*), by 'es' (*it*). We assume that pronoun nodes are stored at the lemma level, and that they are accessed from gender nodes (see Fig. 1).

Fig. 1 illustrates how the pronoun 'sie' (*it*, feminine) is accessed during the generation of the utterance 'Die Blume ist rot. Sie wird blau.' Such sentence pairs were produced by the participants of our experiments on successive trials. The first sentence was said when the object was shown for the first time and the second sentence on the next trial, when the object appeared in a different color. The figure illustrates the moment when an object (e.g. a flower) is presented for the second time. The concept is active and is marked as 'in focus' in the discourse record, which should lead to the selection of a pronoun. How can this be realized? The activated concept activates the corresponding lemma, which in turn activates the gender node. The activated concept activates the corresponding lemma, which in turn activates the gender node.

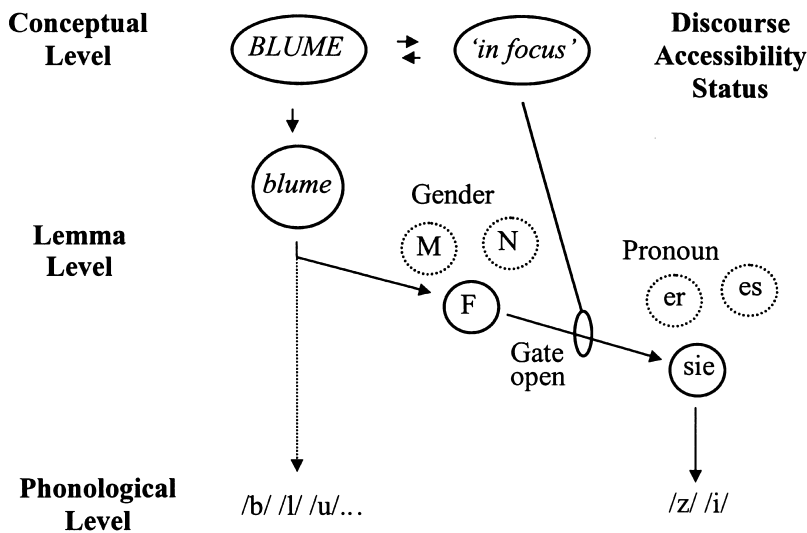


Fig. 1. A lexical access view of the generation of the pronoun *sie* in *Die Blume ist rot. Sie wird blau.* (The flower is red. It turns blue). M, masculine; F, feminine; N, neuter.

At this point, there is no difference in activation of the information necessary to overtly produce a noun, as in *‘die Blume’*, or its corresponding pronoun *‘sie’*. However, because the discourse record includes the feature *‘in focus’*, selection of the phonological form of the noun should be prevented for overt naming, and the pronoun should be selected instead. This discourse-dependent switch in the processing mode is depicted in the figure as a gate in the connection between the lexical gender and pronoun nodes. Note that the gate mechanism is not the main concern of the present paper. It is a theoretical assumption that has been successfully implemented in a computational model that simulated the dual task of lexical decision during picture naming, for both noun and pronoun generation (Schmitt, 1997). The gate is mentioned here as part of our working model of lexical access. It can be either open or closed. If the accessibility status in the discourse record is *‘in focus’*, the gate is open and gives access to the pronoun. The pronoun is selected and subsequently phonologically encoded, just like a noun would be. If the discourse record signals that the current concept is not accessible from previous discourse (*‘not in focus’*), the gate is closed and access of a pronoun is impossible.

The binary on/off nature of the gate may appear strict for theories that in general assume graded spread of activation through a lexical network. The gate is implemented as a binary feature to capture the binary coding of concepts as being or not being in focus. The gating function reflects a procedural rule that works in a discrete rather than continuous fashion.

The gating function is, of course, speculative in nature. It is one way to capture how the speech system generates different kinds of speech output for the same visual input. Obviously, speakers must have long-term memory representations of words and concepts. But if naming relied exclusively on stored knowledge, a given object would always be named in the same way and could not be referred to by a noun on one occasion and by a pronoun on another occasion. The speakers’ ability to use variable discourse information during utterance planning should interact with the stored knowledge. The proposed gating mechanism enables such an interplay between procedural rules and stored conceptual and linguistic information within the same network architecture.

3. Is the antecedent phonologically activated during pronoun generation?

The aim of the experiments reported below was to investigate the phonological activation of a noun antecedent during the generation of the corresponding pronoun. In one condition, speakers produced utterance pairs such as *‘Die Blume ist rot. Sie wird blau.’* (The flower is red. It turns blue.) In another condition, they produced pairs such as *‘Die Sonne ist rot. Die Blume ist blau.’* (The sun is red. The flower is blue.) In order to produce the noun *‘Blume’*, its phonological representation must obviously be activated. We examined whether the phonological form of the noun would also be activated when the concept was referred to anaphorically. As discussed above (see also Fig. 1), in our model the noun’s lemma and the gender node must be selected to choose the correct pronoun, but information about the phono-

logical form of the noun is not required. Hence, one might predict that the form of the noun is only activated when the noun itself is produced, but not when the pronoun is used. In order to generate such a pattern a lexical access model would have to assume a strict cut-off of information flow between the noun lemma and its phonemes when the phonemes are not needed. Alternatively, the activation of a lemma may always lead to some activation of the corresponding word form, even when the form itself is not produced. This view is in line with lexical access models that assume cascading spread of activation (Dell, 1986, 1988; Dell and O'Seaghdha, 1991, 1992; Peterson and Savoy, 1998). Phonological activation of the noun's phonemes during pronoun generation is also assumed by strict serial stage models, according to which the form of a noun is only activated after the lemma has been selected (Levelt et al., 1991a,b, in press; Roelofs, 1992a,b; Roelofs et al., 1998). In these models, mere activation of a lemma would not lead to phonological activation of corresponding phonemes.

Thus, if we find evidence for form activation of the noun during pronoun production, the existence of a strict cut-off of information flow has to be rejected. The evidence for form activation cannot be used to decide between the cascading and the strict-stage view because both predict form activation. Most importantly, however, a phonological effect would support the hypothesis of our working model that pronoun production requires lemma access. This is because the cascading view, as well as the strict stage view, state that the lemma must at least be activated (cascading models) or selected (strict stage models) before access to phonological information is possible.

4. The experimental paradigm

To test whether the phonological form of the antecedent noun is activated during the production of a pronoun, we used a novel version of the lexical decision during naming paradigm introduced by Levelt et al. (1991a), which is based on Stroop-like interference effects (Stroop, 1935; for a review see MacLeod, 1991). In the experiment carried out by Levelt et al., the participants' default task was to name pictures of objects. However, on some trials the naming process was disturbed by the presentation of an acoustic probe. In these cases, the participant had to postpone picture naming and first classify the probe as a word or pseudoword by pressing one of two response buttons. Levelt et al. (1991a) varied the relationship between the lexical decision probe and the target (identical, semantically related, phonologically related, or unrelated) and the timing of the probe onset relative to the picture onset (the stimulus onset asynchrony, SOA). They found, among other things, that at certain SOAs the lexical decision latencies were longer for phonologically related than for unrelated probes. Their account for this finding was that in the related condition, the phonological representation of the picture name, which had begun to be built up, boosted the activation of phonological competitors to the lexical decision probe. This phonological competition had to be resolved, which took extra processing time and delayed the lexical decision to phonologically related probes relative to the unrelated probes.

According to Levelt et al. (1991a), phonological inhibition only arises if the phonological processing of probe and picture name overlap in time. Phonological inhibition indicates that the form of the picture name is activated at the moment when the lexical decision task is carried out. Thus, the lexical decision paradigm can be used to tap whether the form of a word is activated at a given moment in time. Exactly this information is needed to test whether the phonological form of the antecedent noun is activated when a pronoun is produced. Recall that in our model, access to a word form presupposes access to the word's lemma. If we find that the phonological form of the antecedent noun is activated during the generation of a pronoun, it can be concluded that the lemma of the antecedent is also activated.

The materials of our experiments consisted of pairs of pictures shown on successive trials. In the pronoun condition, the same object appeared on both trials for each pair, but in different colors. For instance, the participants first saw a red flower, and then saw the same flower in blue (see Appendix B for an illustration). They had to describe the sequence by saying '*The flower is red. It turns blue.*' On half the trials (the fillers), this was their only task. On the remaining trials, an acoustic probe, a word or pseudo-word of German, was presented time-locked to the onset of the second picture (the target picture). In these cases, the participants first performed the lexical decision task and then described the second picture. On all critical trials the probes were words. They could be either phonologically related or unrelated to the form of the pronominal antecedent. Thus, while the participants prepared the utterance '*Sie wird blau*', referring to a flower (German '*Blume*'), they either heard the related probe '*Bluse*' (*blouse*) or the unrelated probe '*Kelle*' (*ladle*). If the form of the noun antecedent is activated during the production of the pronoun, the mean lexical decision latency to the related probes should be longer than that to the unrelated probes, as discussed above.

In addition to this pronoun condition, we included a noun condition, in which participants also saw object pairs, but the two objects were different. They saw, for instance, first a red sun and then a blue flower (see Appendix B). They were asked to describe such sequences in utterances such as '*The sun is red. The flower is blue.*' Again, describing the pictures was the participants' only task on half the trials. On the remaining trials words or pseudowords of German were presented for lexical decision, time-locked to the onset of the second picture (target picture). The probe was either phonologically related or unrelated to the name of the second picture.

Note that for both the noun and the pronoun condition the same target objects in the same colors were presented, and that they were combined with the same probes. The only difference between the two conditions was whether the object shown on the preceding trial was identical to the target object or different and, therefore, whether participants referred to the target object by a pronoun or a noun. Based on the results obtained by Levelt et al., we strongly expected phonological inhibition (i.e. longer decision latencies to related than to unrelated probes) in the noun condition. The question of main interest was whether such an effect would also be obtained in the pronoun condition.

5. Experiment 1

5.1. Method

5.1.1. Participants

Thirty-two participants, aged 17–38 years, were recruited through a newspaper advertisement. They were native speakers of German. The data from three participants were excluded from further analyses because they made errors on more than 20% of the trials.

5.1.2. Materials

A set of unambiguous pictures was selected, such that most of the participants labeled them in the same way in the naming task. This ensured that the participants would retrieve the intended lexical item for a particular picture. In addition, the picture set had to be homogenous with respect to naming latencies, in order to guarantee that the constant SOA in the main experiment between the onset of the picture and the acoustic probe was comparable for all picture-probe pairs. To create a set of pictures meeting the criteria of unambiguous naming and homogenous naming latencies, we pretested a large set of pictures in a picture-naming experiment.

5.1.2.1. Picture-naming pretest. Eighty-seven black-on-white line drawings were presented for 800 ms each, with an inter-trial interval of 2000 ms. Fifteen German participants were paid for taking part in the experiment. They were instructed to name the pictures as quickly and as correctly as possible. For each participant a different random order of the pictures was created. Pictures were shown on a NEC/Multisync 4FG screen. Naming responses were recorded by a Sennheiser HMD 224 Headphone-Microphone combination and a SONY TCD D3 DAT recorder. The naming latencies were measured by a voice key. The selection criteria for pictures to be used in the main experiment were a mean naming latency below 1000 ms, a standard deviation below 200 ms, and an error rate below 20%. Forty-eight pictures fulfilled the criteria, from which 16 were selected as the target picture set (for details see Schmitt, 1997).

5.1.2.2. Lexical-decision pretest. The acoustic word probes should also be as homogenous as possible in terms of lexical decision latencies and errors. Therefore, we pretested a large set of words and pseudowords in a lexical decision experiment. The same 15 participants from the naming pretest took part in this experiment.

The acoustic stimuli were 87 nouns and 87 pseudowords. The words consisted of one to three syllables. In order to avoid possible gender-incongruency effects (Schriefers, 1993) between probe words and picture names in the main experiment, the word probes had the same grammatical gender as the target picture name with which they were combined later in the main session. Phonologically related probes shared the word onset consonant or consonant cluster and the following vowel with

the target. Unrelated probes did not share any segments in corresponding word positions with the targets. For example, the probe *Bluse* (blouse_{FEM}) was phonologically related to the picture name *Blume* (flower_{FEM}), whereas the probe *Kelle* (ladle_{FEM}) was unrelated. Unanimous rating by five judges had established that the probes and targets were not related in meaning. In addition to these words, the materials included 87 pseudowords. They were generated out of the probe words by exchanging consonants so that the stimuli was still pronounceable, but meaningless. The stimuli were spoken by a male speaker and recorded using a Sony 59ES DAT recorder. They were digitized with a sampling frequency of 16 kHz. The stimuli were presented in a random sequence for each participant, with an inter-trial interval of 2000 ms. The participants heard the stimuli via a Sennheiser HMD 224 Headphone-Microphone combination. Participants were asked to press one of two push-buttons as quickly and as correctly as possible to indicate whether a stimulus was a word or pseudoword. The decision latencies were registered. Criteria for including acoustic probes in the main experiment were a mean reaction time below 1100 ms, a standard deviation below 200 ms, and an error rate below 20% (see Schmitt, 1997, for details).

5.1.3. *The resulting set of pictures and probes*

The resulting materials for the main experiment were a set of 16 pictures and their related and unrelated probe words, plus 16 pseudowords. In addition 32 filler pictures were chosen (see Appendix A). In order to realize color naming in the main experiment, the pictures were colored. Four colors were chosen so that they were easily recognizable (red, yellow, green, and blue).

5.1.4. *Design*

The experiment included two sessions, carried out on separate days: a main session, in which participants carried out the lexical decision during naming task, and a baseline session in which they classified the probes as words or pseudowords, just like the participants of the second pretest had done.

The main session included equal numbers of lexical decision and filler trials (128 trials in total). The lexical decision trials consisted of the presentation of the target pictures at the second position of the picture sequence, paired with the presentation of word and pseudoword probes (16 related, 16 unrelated word probes, and 32 pseudoword probes). The 64 filler trials consisted of sequences of filler pictures and target pictures, in random order. The filler trials were included in order to keep participants in the naming mode.

In the experimental trials, the word probes were either phonologically related or unrelated to the name of the second object (the target) of a pair. Crossed with this variable ‘probe conditions’ was the variable ‘utterance conditions’ (pronoun versus noun). To elicit nouns, two different objects were shown on a pair of trials. To elicit pronouns, the same object was shown twice in different colors. The same target pictures (second pictures of the pairs) and the same probes were used in both utterance conditions. For each of the probe presentations per target picture, the color change from the first to the second picture was the same. The color changes were counterbalanced across target pictures.

Each target picture was presented twice to each participant as a second picture of a pair, once paired with the related probe, once with the unrelated probe. The experimental picture set was split into two groups. For eight of the 16 experimental pictures, a participant received the phonological and unrelated probe words in the noun condition. For the remaining eight pictures the same participant received the related and unrelated probe in the pronoun condition. The order of presenting related and unrelated probes was randomized. The 32 participants were randomly divided into two groups of 16. One subject group received the first eight experimental pictures in the noun condition, and the second eight experimental pictures in the pronoun condition. The reverse held for the second subject group. All participants received the same materials on filler trials.

The order of trials was random, with the constraint that no more than two word probe presentations followed each other. In addition, no more than five acoustic probes, and no more than five successive trial pairs of the same naming format were presented. Each participant obtained a different randomized order of trials.

Although the related and unrelated probes were matched for mean lemma frequency, mean word length, and mean response latency in the pretest, participants varied, of course, in their decision latencies to the probes. To control for inter-individual variation, a baseline experiment was carried out with the participants of the main experiment. In that experiment, all probes of the main experiment were presented without pictures, and the participants only performed the lexical decision task. The probes were presented in a different random order to each participant. This experiment was carried out about one week after the main experiment. As will be explained below, the analyses of the experimental results were based on differential scores between the decision latencies in the baseline and the main experiment.

5.1.5. *Apparatus*

The same equipment was used as in the pretests.

5.1.6. *Procedure*

The entire experiment included a practice session, a main session, and a baseline session, carried out in this order. In the practice session, the participants first saw all pictures with their names printed below the picture. They were asked to name the pictures as shown. Then they were presented with 32 picture pairs (16 in the noun condition and 16 in the pronoun condition) and practiced describing them using either two nouns or a noun and a pronoun. Next, the lexical decision task was introduced. The participants were asked to describe the picture pairs as practiced before, but whenever a probe was presented, they should first make the lexical decision and then name the second picture. This practice block consisted of 80 trials, on half of which a probe was presented. Words and pseudowords occurred equally often in the noun and pronoun condition. The trial timing was the same as in the experimental blocks and is described below. The practice session lasted about 40 min and was followed by the main session. The main session consisted of two blocks, each of which included three warming-up and 64 experimental trials. The

baseline session was carried out about a week after the main session, testing the lexical decision latencies on the word probes and pseudoword probes that were used in the main session.

5.1.7. *Trial structure*

A fixation cross was presented for 500 ms, followed by a blank-screen interval of 300 ms. Then, the first picture of a pair was presented for 1500 ms. When a participant started to describe the picture, the voice key was triggered, and 1500 ms later the second picture was presented. Thus, the time interval between the onset of the first picture naming and onset of the second picture was constant. This will be relevant in the comparison of the first and the second main experiment (as discussed below).

On pure naming trials, the participant described the second picture, and 2500 ms after the offset of the second picture the next trial began. On lexical decision trials, a probe was presented 100 ms after the onset of the second picture. This SOA was chosen because Levelt et al. (1991a) found phonological inhibition at a similar SOA (approximately 73 ms). The push-button device was activated from probe onset for 1500 ms. After the lexical decision was carried out, the participants gave the picture description. The next trial began 2500 ms after offset of the second picture. Each trial lasted about 7 s. Each experimental block lasted about 7 min. The trial structure of the baseline session was identical to that of the lexical decision pretest.

5.2. *Results*

We were primarily interested in the differential scores, i.e., the decision latency differences between the baseline and main session for the two types of probes (related and unrelated) and the two types of utterances (nouns versus pronouns). We will only consider the results for the word probe trials and report separate subject and item analyses. ‘Item’ was defined as being the target object name of the second picture in the main session. Each item was paired with related and unrelated probes in different naming conditions. To be able to interpret the differential scores we must first examine the results for the probes in the baseline session, which will therefore be described first.

5.2.1. *Baseline session*

In the baseline session each participant made lexical decisions to the same target word probes as in the main session. Because of the grouping in the design of the main experiment (see design section), each participant was presented with four different types of word probes: (1) the words that were related to a target picture name in the main session, and the picture had to be described using a pronoun (in short: related probes, pronoun condition), (2) unrelated probes, pronoun condition (3) related probes, noun condition, and (4) unrelated probes, noun condition. This grouping also holds for the item analysis. In order to see whether the four groups of word probes in the baseline differ *per se*, errors and lexical decision latencies for the

four groups were analyzed by 1×4 repeated measurement ANOVAs with ‘group of word probes’ as the repeated measures factor.

5.2.1.1. Errors. Errors consisted of responses with latencies above 1500 ms and incorrect decisions. The error percentages across subjects in the baseline experiment were 3.5% (for probes that were used in the main session as related words in the pronoun condition), 1% (unrelated probes in pronoun condition); 3.8% (related probes in noun condition), and 4.3% (unrelated probes in noun condition). The arcsin transformed error proportions were analyzed in a repeated measures ANOVA with the four groups of probes as a within factor (Winer, 1971). Statistically the baseline groups did not differ in error rates, $F_1(3,84) = 2.2$, $Ms_e = 0.011$; $F_2(3,45) = 1.3$, $Ms_e = 0.02$.

5.2.1.2. Lexical decision latencies. The mean reaction times for correct lexical decisions on word probes in the baseline are shown in Table 1. The four groups are statistically equal in terms of reaction times. This was shown by repeated measures ANOVAs with the factor ‘group’ (four groups of word probes), $F_1(3,84) = 2.2$, $Ms_e = 2657$; $F_2(3,45) = 0.72$, $Ms_e = 3985$.

5.2.2. Main session

The main results are shown in Table 1 and Fig. 2. The table displays mean lexical decision latencies for the critical word probes in the baseline and main session, separately for the two naming conditions. The statistical analyses are based on differential scores. They were obtained by pairwise subtracting each participant’s main session reaction time from that of the same item in the baseline session for each utterance condition. Because of the longer lexical decision latencies in the dual task session in comparison to the baseline session, this subtraction resulted in negative values. The mean differential scores are also presented in Table 1. Missing values were defined pairwise. That is, whenever a participant’s reaction time in the baseline

Table 1

Results of Experiment 1: Mean lexical decision latencies (by subjects) and differential scores (in ms), standard deviation (SD) of the differential scores, and percentages of missing values for the baseline and the main session for the two word probe conditions (phonologically related and unrelated) per utterance format (pronoun and noun generation).

Utterance format	Pronoun generation		Noun generation	
	Phon.	Unrel.	Phon.	Unrel.
Baseline latency	972	952	972	943
Main session latency	1036	995	1086	1007
Difference (baseline – main)*	–64	–44	–112	–67
SD	107	150	108	149
% missing	13	7	18	12

*Any apparent inaccuracies in means are the result of rounding.

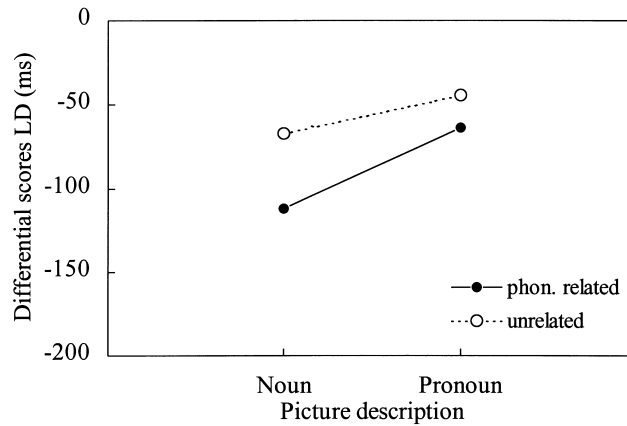


Fig. 2. Mean differential scores of lexical decision latencies in Experiment 1: Baseline session (without picture description) minus main session (during picture description) for related and unrelated probes during noun and pronoun generation. Negative values indicate inhibition.

or the main session was missing, the differential score was also treated as a missing value. The analyses are based on the remaining data.

A missing value in the baseline session was defined as mentioned above. A missing value in the main session arose if no lexical decision was made, if an incorrect decision was made, or if the lexical decision latency was longer than 1500 ms. In addition, incorrect naming or naming onset before the onset of the acoustic probe was defined as missing.

5.2.2.1 Errors. The percentage of missing values for differential scores are presented in Table 1. The error analysis was carried out on the arcsin transformed error proportions in a 2×2 repeated measures ANOVA with ‘utterance condition’ (pronoun vs. noun) and ‘probe condition’ (related vs. unrelated probes) as within-subject and within-item factors. Only the factor ‘probe condition’ was significant: $F_1(1,28) = 5.5$, $Ms_e = 0.048$, $P = 0.03$; $F_2(1,15) = 8.6$, $Ms_e = 0.032$, $P = 0.01$. This result indicated a higher percentage of errors in the related condition than in the unrelated condition. Because the related condition also had the longest decision latencies in the main experiment, a speed-accuracy trade-off could be excluded. Furthermore, because the error analysis in the baseline did not reveal differences between groups, the differences in error percentages in the main session can be attributed to the relationship between probes and target pictures.

5.2.2.2 Differential scores. Table 1 shows the mean differential scores and the raw means in the baseline and main session. Repeated measurements ANOVAs were performed on the differential scores with ‘utterance condition’ (noun and pronoun generation) and ‘probe condition’ (related and unrelated probes) as within-factors, separately for subjects and items. Five extreme outliers were excluded from the analysis, following Tukey (1977) (see also Kirby, 1993). The main effect ‘utterance condition’ was significant, $F_1(1,28) = 8.9$, $Ms_e = 4142$,

$P < 0.01$; $F_2(1,15) = 4.96$, $Ms_e = 2860$, $P = 0.04$. The lexical decisions were slower in the noun condition than in the pronoun condition. The main effect ‘probe condition’ was significant as well, $F_1(1,28) = 4.2$, $Ms_e = 7137$, $P = 0.05$; $F_2(1,15) = 7.3$, $Ms_e = 2851$, $P = 0.01$. Lexical decisions were slower to related probes than to unrelated probes. The phonological interference effect was 20 ms in the pronoun condition, and 45 ms in the noun condition. Importantly, the interaction of ‘utterance condition’ and ‘probe condition’ was not significant, $F_1(1,28) = 1.2$, $Ms_e = 3652$; $F_2(1,15) = 1.6$, $Ms_e = 820$. The lack of interaction showed that phonological inhibition was found in both the noun and the pronoun condition. It also showed that the inhibition was statistically of equal size for both utterance formats.

5.3. Discussion

In the noun naming condition, we obtained evidence for phonological inhibition; i.e., the lexical decision latencies were longer when probe and target noun were related in phonological form than when they were unrelated. This finding replicates a result obtained by Levelt et al. (1991a) for one-word utterances and shows that the lexical decision paradigm can fruitfully be employed to study phonological activation during more complex utterances than single nouns. As discussed above, the phonological inhibition can be ascribed to phonological competition between the phonological forms of the target pictures’ names and the probes.

More importantly, evidence for phonological inhibition was also obtained in the pronoun condition. This suggests that when the participants produced a pronoun, the phonological form of the antecedent was activated, just as it was when the noun itself was produced. This argues against the hypothesis that during pronoun production the flow of information is cut off at the lemma level. Instead it appears that the phonological form of the antecedent is activated. This finding is compatible with cascading models according to which activated lemmas automatically activate the corresponding word forms and with serial models according to which phonological information becomes activated as soon as a lemma has been selected.

Since the statistical analyses revealed no interaction between utterance condition and probe condition, we can argue that the phonological effect was the same for the noun and pronoun conditions. Nevertheless the effect sizes differed considerably (45 ms for nouns, 20 ms for pronouns). It should be noted here already that the phonological effect was replicated, and was much stronger, in the pronoun condition of Experiment 2.

It is an interesting open question how different models of lexical access could explain a weaker phonological effect for pronouns than for nouns. In a strict serial model, the selection of the lemma is the same for nouns and pronouns. Therefore, the initial phonological activation of the noun’s phonological segments should also be the same. However, in the pronoun condition, the segments of the pronoun also become activated and begin to compete with or inhibit those of the noun, such that eventually the form of the pronoun is selected. Thus, soon after the selection of the lemma, the phonological form of the noun should be less

strongly activated in the pronoun than in the noun condition, which may be reflected in a weaker phonological effect. Basically the same interpretation can be applied to a cascading model. But this account is speculative and needs further investigation.

In the next experiment, we addressed an alternative explanation for the observed phonological interference during pronoun generation. Recall that in Experiment 1 in the pronoun condition the context and target object were identical. Thus, the object that was referred to by a pronoun had been referred to by a noun in the immediately preceding utterance. It is therefore possible that the phonological inhibition did not arise because the lemma and phonological form of the noun antecedent were reactivated during the generation of the pronoun, but because they were still activated from the preceding utterance. Either phonological planning or articulation of the noun, or both could be responsible for the effect.

Experiment 2 was carried out to test this hypothesis. It used the same materials and design as Experiment 1, but the two nouns of the noun trials were reversed, while the probes remained the same as in the first experiment (see Appendix C for an illustration of the design). Thus, the phonologically related probes were now not related any more to the second noun of a pair, as they had been in the first experiment, but to the first noun. As before, the probe was presented 100 ms after the onset of the second picture and 1600 ms after the naming onset of the first noun. If the form of the first noun is still active during the presentation of the probe, phonological interference should be found. By contrast, if there is no residual activation left at the form level, the lexical decision latencies to related and unrelated probes should be identical.

Experiment 2 included the same pronoun trials as Experiment 1. This gave us the opportunity to test whether the phonological inhibition effect observed in the first experiment could be replicated.

6. Experiment 2

6.1. Method

6.1.1. Participants

Thirty-five participants were recruited through a newspaper advertisement, and paid for participation. Their age ranged from 18 to 36 years. The data from five participants were excluded from further analysis, because they made errors on more than 20% of the trials.

6.2. Material, design, apparatus, and procedure

The same materials were used as in Experiment 1, except that in the noun condition the order of the two pictures shown on successive trials was reversed (see Appendix C). Design, apparatus, and procedure were the same as in Experiment 1.

6.3. Results

As discussed in the Result Section of Experiment 1, in order to be able to interpret the differential scores we must first examine the results for the probes in the baseline session, which will therefore be described first. Again, we will only consider the results for the word probe trials, separately for subjects and items. ‘Item’ is defined as being the target object name of the first picture in the main session. The main results are shown in Fig. 3 and Table 2. Fig. 3 shows the differential scores. Table 2 displays mean lexical decision latencies for the critical test probes in the baseline session and main session for the two naming conditions, the differential scores, and the missing values for differential scores. Missing values were excluded from the latencies analysis (as defined in the Result Section of Experiment 1).

6.3.1. Baseline

6.3.1.1. Errors. Errors were defined in the same way as in Experiment 1. Error percentages across subjects in the baseline session were 2.5% (for probes that were used as related probes in the pronoun condition in the main session), 2.1% (used as unrelated probes in the pronoun condition), 3.7% (used as related probes in the noun condition), 1% (used as unrelated probes in the noun condition). The arcsin transformed error proportions of the baseline session were tested in a repeated measures ANOVA with the four conditions as the within subject or item factor. The four conditions in the baseline did not differ with regard to error rates, $F_1(3,87) = 1.7$, $Ms_e = 0.009$; $F_2(3,45) = 2.09$, $Ms_e = 0.01$.

6.3.1.2. Lexical decision latencies. The mean reaction times for correct responses on the four word probe conditions in the baseline (see Table 2) were statistically

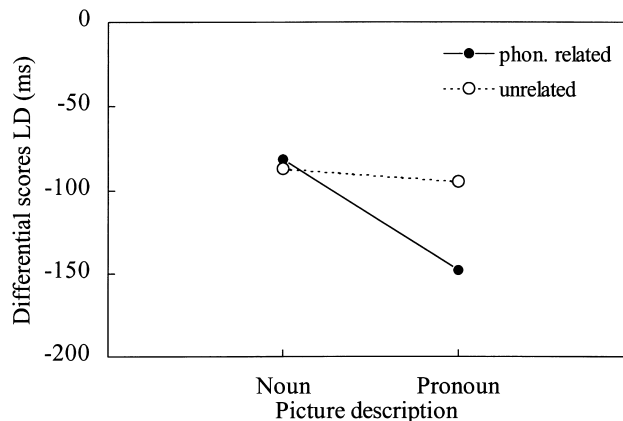


Fig. 3. Mean differential scores of lexical decision latencies in Experiment 2: Baseline session (without picture description) minus main session (during picture description) for related and unrelated probes during noun and pronoun generation. Negative values indicate inhibition.

Table 2

Results of Experiment 2: Mean lexical decision latencies (by subjects) and differential scores (in ms), standard deviation (SD) of the differential scores, and percentages of missing values for the baseline and the main session for the two word probe conditions (phonologically related and unrelated) per utterance format (pronoun and noun generation).

Utterance format	Pronoun generation		Noun generation	
	Phon.	Unrel.	Phon.	Unrel.
Baseline latency	910	894	921	905
Main session latency	1060	988	1003	991
Difference (baseline – main)*	–148	–95	–82	–87
SD	95	102	90	90
% missing	13	10	14	7

*Any apparent inaccuracies in means are the result of rounding.

equal. This was shown by repeated measures ANOVAs on the factor ‘group’ (four groups of word probes), $F_1(3,87) = 1.1$, $Ms_e = 3188$; $F_2(3,45) = 0.48$, $Ms_e = 4798$.

6.3.2. Main session

The main results of differential scores are shown in Fig. 3 and Table 2. Also shown in the table are the missing values, obtained by adding error trials in the baseline and in the main session (as defined in the result section of Experiment 1).

6.3.2.1 Errors. The error analysis was carried out on the arcsin transformed error proportions in a 2×2 repeated measures ANOVA with ‘utterance condition’ (noun and pronoun generation) and ‘probe condition’ (related and unrelated probes) as within subject and item factors. The effect of ‘probe condition’ was significant: $F_1(1,29) = 7.7$, $Ms_e = 0.056$, $P = 0.01$; $F_2(1,15) = 9.5$, $Ms_e = 0.037$, $P < 0.01$. More errors were produced in the phonologically related condition than in the unrelated condition. The related condition also had the longest decision latencies in the main session. Therefore, a speed-accuracy trade-off could be rejected. The possibility that the higher percentage of errors in the phonological conditions was due to item-specific effects could be ruled out because of the results in the baseline error analysis (see above). The error percentage across item groups did not differ in the baseline session. Therefore, the observed higher percentage of errors for related probes in the differential scores can be associated with a higher performance of errors in the main session. It indicates that the processing of related probes in a dual task situation is more difficult than that of unrelated probes.

6.3.2.2 Differential scores. For subjects and items, 2×2 repeated measures ANOVAs were carried out on the differential scores with ‘utterance condition’ (noun and pronoun naming) and ‘probe condition’ (related and unrelated) as repeated measures factors. Six outliers were excluded from the analysis, following Tukey (1977). The main effect ‘utterance condition’ was significant, $F_1(1,29) = 5.2$, $Ms_e = 7572$, $P = 0.03$; $F_2(1,15) = 6.03$, $Ms_e = 3845$, $P = 0.03$. This result showed that lexical decision became more inhibited in the

pronoun condition (about 121 ms) than in the noun condition (about 84 ms). The main effect of ‘probe condition’ was not significant, $F_1(1,29) = 2.9$, $Ms_e = 5772$; $F_2(1,15) = 1.5$, $Ms_e = 5903$. However, the interaction ‘utterance \times probe’ was significant, $F_1(1,29) = 5.2$, $Ms_e = 4405$, $P = 0.03$; $F_2(1,15) = 9.74$, $Ms_e = 1554$, $P < 0.01$. As the figure shows, there was substantial inhibition (53 ms) in the pronoun condition, but no inhibition in the noun condition.

6.4. Discussion

In Experiment 2, the phonological interference effect obtained in the pronoun condition of Experiment 1 was replicated. Thus, both experiments showed that the form of the noun antecedent is activated when it is replaced by a pronoun. In addition, Experiment 2 showed that the activation of the form of the antecedent was not due to residual activation from the use of the noun in the preceding utterance. This can be concluded from the finding that no phonological interference was found when the probe was related in form to the first noun of the sentence pair, and the probe was presented 1600 ms after the onset of the naming of the first picture (as in the present experiments). Thus, the phonological form of a noun apparently quickly decays after the noun has been said, but is reactivated if the noun’s corresponding pronoun is generated.

7. Conclusion

According to our working model, pronouns are used when lexical concepts are marked as ‘*in focus*’ in the discourse record. Since in German, and many other languages, pronouns agree in grammatical gender with the antecedent nouns, we assume that the choice of the correct pronoun requires the corresponding lemma and the attached gender node to be selected. On the assumption that a phonological form cannot become activated in the absence of activation of the corresponding lemma, the obtained phonological inhibition effect supports the hypothesis that pronoun selection requires lemma access. Furthermore, our results imply that lemma access of a noun automatically entails activation of the corresponding phonological form, even when the form is not uttered, because the noun is replaced by a pronoun.

In Experiment 2, we investigated the possible interpretation of this phonological effect in terms of residual activation of the description of the first picture (the overt naming of the target noun). The results of the second experiment showed no phonological effect in the noun condition. The null effect indicated that there was no measurable residual activation from the description of the first picture when the second picture and the probe were presented (1600 ms after speech onset of the first picture description). The null effect supported the conclusion that the phonological inhibition observed in Experiment 1 and 2 during pronoun generation must be due to re-accessing the form of the antecedent. Re-accessing the form, however, means that the lemma became re-

accessed during pronoun generation, supporting the proposed lexical access model of pronoun generation.

Acknowledgements

The research reported in this paper was part of a PhD project of B. Schmitt, supported by the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, München, Germany. We would especially like to thank Herbert Baumann, Inge Doehring, Gerd Klaas, John Nagengast, and Ad Verbunt for technical assistance; Jens Bölte and Pienie Zwitserlood, Department of Psychology, University of Münster, Germany, who helped us very much to carry out the experiments in German by offering one of their laboratories; and Sven Blankenberger, Kay Bock, Jos Van Berkum, and all members of the project ‘Lexical access in speech production’, from the Max-Planck-Institute for Psycholinguistics, Nijmegen, for helpful discussions of the issues considered in this paper.

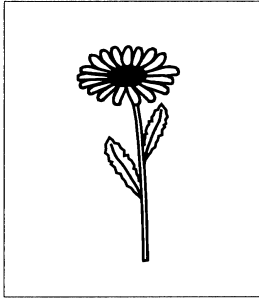
Appendix A. Materials for Experiments 1 and 2

Second picture in Experiment 1 or first picture in Experiment 2	Phonologically related probe	Unrelated probe	Gender
Apfel (apple)	Abfall (rubbish)	Tresor (safe)	Masculine
Bürste (brush)	Büchse (tin)	Polizei (police)	Feminine
Eimer (bucket)	Eiter (pus)	Kanal (canal)	Masculine
Ellipse (ellipse)	Elle (yardstick)	Palme (palm)	Feminine
Fass (barrel)	Fach (drawer)	Brot (bread)	Neuter
Fenster (window)	Fett (fat)	Schnitzel (escalope)	Neuter
Flasche (bottle)	Flamme (flame)	Scheune (barn)	Feminine
Frosch (frog)	Frost (frost)	Deckel (lid)	Masculine
Gans (goose)	Gasse (lane)	Hütte (hut)	Feminine
Gras (grass)	Grab (grave)	Bier (beer)	Neuter
Hammer (hammer)	Hammel (wether)	Vogel (bird)	Masculine
Kaktus (cactus)	Kasten (box)	Napf (bowl)	Masculine
Lineal (ruler)	Lied (song)	Nest (nest)	Neuter
Nase (nose)	Nadel (needle)	Weste (vest)	Feminine
Quadrat (square)	Quartett (quartet)	Benzin (gasoline)	Neuter
Zelt (tent)	Zentrum (center)	Lamm (lamb)	Neuter

Pseudowords used in Experiments 1 and 2: Banako, Eip, Holk, Humpf, Kator, Mosak, Natis, Nepf, Pidda, Pufke, Quesse, Schill, Stulp, Trekal, Vodil, Zars.

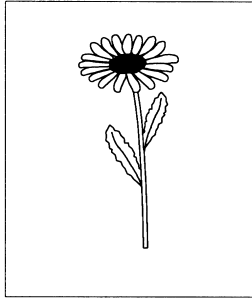
Appendix B. Design illustration of Experiment 1

First picture



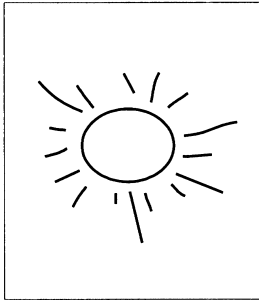
The flower is red...
Die Blume ist rot...

Second picture

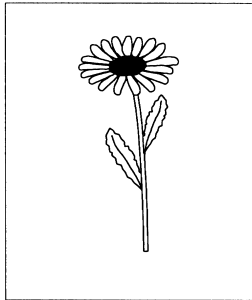


It turns blue.
Sie wird blau.

Pronoun generation

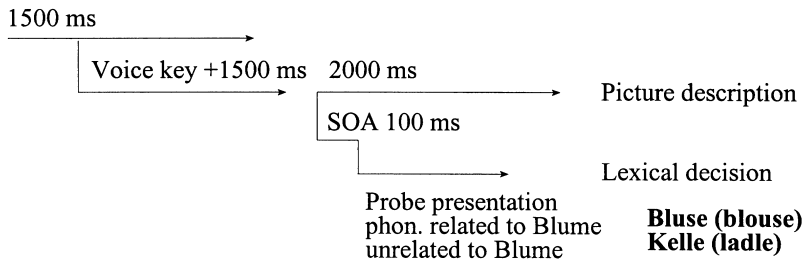


The sun is red.
Die Sonne ist rot.



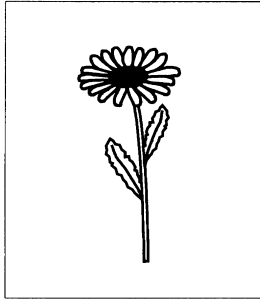
The flower is blue.
Die **Blume** ist blau.

Noun generation



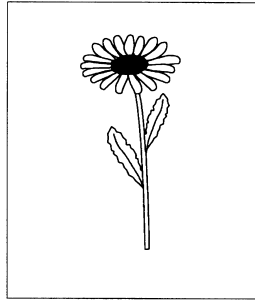
Appendix C. Design illustration of Experiment 2

First picture



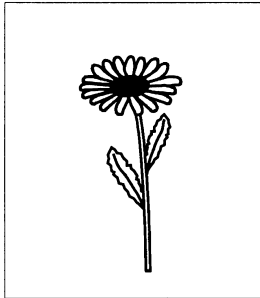
The flower is red...
Die Blume ist rot...

Second picture

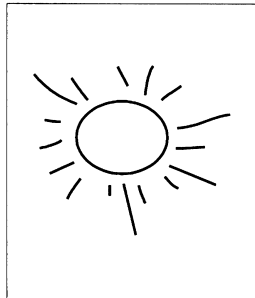


It turns blue.
Sie wird blau.

Pronoun generation

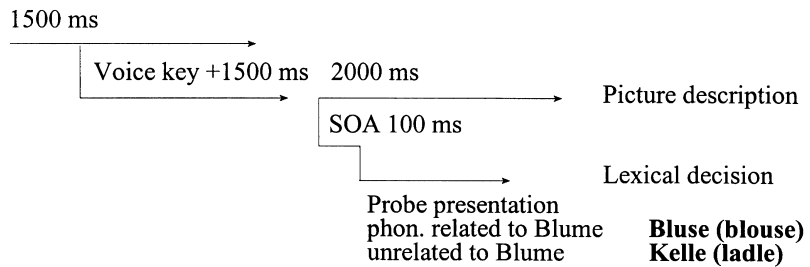


The flower is red.
Die **Blume** ist rot.



The sun is blue.
Die Sonne ist blau.

Noun generation



References

Bock, K., 1995. Sentence production: from mind to mouth. In: Miller, J.L., Eimas, P.D. (Eds.), *Handbook of Perception and Cognition: Vol. 11. Speech, language, and communication*. Academic Press, San Diego, pp. 181–216.

Bock, K., Levelt, W., 1994. Language production: Grammatical encoding. In: Gernsbacher, M.A. (Ed.), *Handbook of Psycholinguistics*. Academic Press, San Diego, pp. 945–984.

- Caramazza, A., 1997. How many levels of processing are there in lexical access? *Cognitive Neuropsychology* 14, 177–208.
- Chafe, W.L., 1976. Givenness, contrastiveness, definiteness, subjects, topics, and points of view. In: Li, C.N. (Ed.), *Subject and Topic*. Academic Press, New York, pp. 25–56.
- Dell, G.S., 1986. A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93, 283–321.
- Dell, G.S., 1988. The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language* 27, 124–142.
- Dell, G.S., O'Seaghdha, P.G., 1991. Mediated and convergent lexical priming in language production: a comment on Levelt et al. (1991). *Psychological Review* 98, 604–614.
- Dell, G.S., O'Seaghdha, P.G., 1992. Stages of lexical access in language production. *Cognition* 42, 287–314.
- Ericsson, K.A., Kintsch, W., 1995. Long-term working memory. *Psychological Review* 102, 211–245.
- Garrett, M.F., 1975. The analysis of sentence production. In: Bower, G.H. (Ed.), *The Psychology of Learning and Motivation*, Vol. 9. *Advances in Research and Theory*. Academic Press, New York, pp. 133–177.
- Garrett, M.F., 1980. Levels of processing in sentence production. In: Butterworth, B. (Ed.), *Language Production*, Vol. 1. Academic Press, London, pp. 177–220.
- Garrett, M.F., 1988. Processes in language production. In: Newmeyer, F.J. (Ed.), *Linguistics: The Cambridge Survey*, Vol. 3. *Language: Biological and Psychological Aspects*. Cambridge University Press, Cambridge, MA, pp. 69–96.
- Glaser, W.R., 1992. Picture naming. *Cognition* 42, 61–105.
- Gordon, P.C., Grosz, B.J., Gilliom, L.A., 1993. Pronouns, names, and the centering of attention in discourse. *Cognitive Science* 17, 311–347.
- Gordon, P.C., Hendrick, R., 1997. Intuitive knowledge of linguistic co-reference. *Cognition* 62, 325–370.
- Gordon, P.C., Searce, K.A., 1995. Pronominalization and discourse coherence, discourse structure and pronoun interpretation. *Memory and Cognition* 23, 313–323.
- Grosz, B.J., Joshi, A., Weinstein, S., 1983. Providing a unified account of definite noun phrases in discourse. *Proceedings of the 21st Annual Meeting of the Association of Computational Linguistics*. ACL, Cambridge, MA, pp. 41–50.
- 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.
- Joshi, A., Weinstein, S., 1981. Control of inference: role of some aspects of discourse-structured centering. *Proceedings of the International Joint Conference on Artificial Intelligence* 385–387.
- Kempen, G., Huijbers, P., 1983. The lexicalization process in sentence production and naming: indirect election of words. *Cognition* 14, 185–209.
- Kirby, K.N., 1993. *Advanced data analysis with SYSTAT*. Van Nostrand Reinhold, New York.
- Levelt, W.J.M., 1989. *Speaking: From Intention to Articulation*. MIT Press, Cambridge, MA.
- Levelt, W.J.M., 1992. Accessing words in speech production: stages, processes, and representations. *Cognition* 42, 1–22.
- Levelt, W.J.M., Roelofs, A., Meyer, A.S. (in press). A theory of lexical access in speech production. *Behavioral and Brain Sciences*.
- Levelt, W.J.M., Schriefers, H., Vorberg, D., Meyer, A.S., Pechmann, T., Havinga, J., 1991a. The time course of lexical access in speech production: a study of picture naming. *Psychological Review* 98, 122–142.
- Levelt, W.J.M., Schriefers, H., Vorberg, D., Meyer, A.S., Pechmann, T., Havinga, J., 1991b. Normal and deviant lexical processing: reply to Dell and O'Seaghdha. *Psychological Review* 98, 615–618.
- MacLeod, C.M., 1991. Half a century of research on the Stroop effect: an integrative review. *Psychological Bulletin* 109, 163–203.
- Marslen-Wilson, W., Levy, E., Tyler, L.K., 1982. Producing interpretable discourse: The establishment and maintenance of reference. In: Jarvella, R., Klein, W. (Eds.), *Speech, Place and Action*. *Studies in Deixis and Related Topics*. Wiley, Chichester, pp. 339–378.

- Martin, N., Gagnon, D.A., Schwartz, M.F., Dell, G.S., Saffran, E.M., 1996. Phonological facilitation of semantic errors in normal and aphasic speakers. *Language and Cognitive Processes* 11, 257–282.
- 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.
- Meyer, A.S., 1997. Conceptual influences on grammatical planning units. *Language and Cognitive Processes* 12, 859–863.
- Meyer, A.S., Sleiderink, A., Levelt, W.J.M., 1998. Viewing and naming objects: eye movements during noun phrase production. *Cognition* 66, B25–B33.
- O’Seaghdha, P.G., Marin, J.W., 1997. Mediated semantic-phonological priming: calling distant relatives. *Journal of Memory and Language* 36, 226–252.
- 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., 1992a. A spreading-activation theory of lemma retrieval in speaking. *Cognition* 42, 107–142.
- Roelofs, A., 1992b. Lemma retrieval in speaking: a theory, computer simulations, and empirical data. Doctoral dissertation, NICI Technical Report 92-08, University of Nijmegen, The Netherlands.
- 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.
- Roelofs, A., Meyer, A.S., Levelt, W.J.M., 1998. A case for the lemma/lexeme distinction in models of speaking: comment on Caramazza and Miozzo (1997). *Cognition* 66, 219–230.
- Schmitt, B.M., 1997. Lexical access in the production of ellipsis and pronouns. Doctoral dissertation, University of Nijmegen, The Netherlands.
- Schriefers, H., 1993. Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19, 841–850.
- Schriefers, H., Meyer, A.S., Levelt, W.J.M., 1990. Exploring the time course of lexical access in language production: picture-word interference studies. *Journal of Memory and Language* 29, 86–102.
- Starreveld, P.A., La Heij, W., 1995. Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21, 686–698.
- Starreveld, P.A., La Heij, W., 1996. Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22, 896–918.
- Stemberger, J.P., 1985. An interactive activation model of language production. In: Ellis, W.W. (Ed.), *Progress in the Psychology of Language*. Erlbaum, Hillsdale, NJ, Vol. 1, pp. 143–186.
- Stroop, J.R., 1935. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* 18, 643–662.
- Tukey, J.W., 1977. *Exploratory data analysis*. Addison-Wesley, Reading, MA.
- Van Berkum, J.J.A., 1996. The psycholinguistics of grammatical gender. Doctoral dissertation, University of Nijmegen, The Netherlands.
- Van Berkum, J.J.A., 1997. Syntactic processes in speech production: the retrieval of grammatical gender. *Cognition* 64, 115–152.
- Van Turenout, M., Hagoort, P., Brown, C.M., 1997. Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23, 787–806.
- Van Turenout, M., Hagoort, P., Brown, C.M., 1998. Brain activity during speaking: from syntax to phonology in 40 milliseconds. *Science* 280, 572–574.
- Winer, J.B., 1971. *Statistical principles in experimental design*. McGraw-Hill, New York.