



Gender congruency effects in picture naming

F.-X. Alario *, Rafael E. Matos, Juan Segui

CNRS and Université René Descartes, Paris, France

Received 2 October 2003; received in revised form 1 June 2004; accepted 2 June 2004

Available online 28 July 2004

Abstract

The analysis of speech error corpora in various gender-marked languages has shown that noun substitutions tend to preserve grammatical gender. This result has been taken as an indication that grammatical gender could play a constraining role during the process of lexical selection. To gain insights on the status of grammatical gender in the speech production system, we discuss this hypothesis and we report three picture naming experiments. We attempted to observe gender-marked context effects in the course of error-free speech production. Participants named pictures shortly after processing a prime that was or was not gender marked and that was or was not congruent with the name of the picture. A clear congruency effect was observed, involving both facilitation in the gender congruent conditions and inhibition in gender incongruent conditions. Different interpretations of this effect and of previously reported gender context effects are discussed in the context of current models of speech production.
© 2004 Elsevier B.V. All rights reserved.

PsycINFO classification: 2720

Keywords: Language production; Speech; Grammatical gender; Picture; Naming

* Corresponding author. Address: CNRS, Laboratoire de Psychologie Cognitive, Case 66, Université de Provence, 3 place Victor Hugo, 13 331 Marseille Cedex, France. Fax: +33 491 106255.

E-mail address: alario@up.univ-mrs.fr (F.-X. Alario).

1. Introduction

The oral production of words is generally thought to entail various levels of processing. First, the message to be communicated is elaborated at the conceptual level; then, the relevant words are selected and encoded at the so-called lexicalization level; finally, the encoded sequence is outputted at the articulatory stage (Bock & Levelt, 1994; Caramazza, 1997; Dell, 1986; Levelt, Roelofs, & Meyer, 1999). In this type of model, the message to be communicated is the main source of activation of lexical nodes. For example, the representation of the word “cat” is activated at the lexicalization level whenever the speaker wants to convey a message about that animal.¹ The analysis of speech error corpora has suggested that other types of information might also contribute to this process. According to these studies, the presence of activated grammatical gender in the production system could under some circumstances play a role in the activation and/or selection of lexical items (see below). The validity of such a conclusion is important, as it provides potential insights about the architecture of the speech production system. The goal of this article is to report three experiments in which we sought to observe the effect of grammatical gender during the course of error-free speech production.

2. The grammatical gender feature

Grammatical gender is a feature present in many languages. In a gender marking language, the form of certain items—such as determiners, adjectives, verbs, etc.—is dependent on properties of other lexical items—typically nouns. For example in French, the language used in this study, nouns can either be masculine (masc.) or feminine (fem.). This property surfaces for example when nouns are used with determiners, such as the definite article: compare “*la*_{fem} *table*_{fem}” (*the table*), to “*le*_{masc} *piano*_{masc}” (*the piano*). In some cases, grammatical gender correlates with semantic properties of the lexical items. For example, in French, words denoting female entities tend to be feminine (e.g., “*la jument*_{fem}” [the mare]; for more complex semantic correlates, see Corbett, 1991). However, the grammatical gender of nouns denoting non-sexuated entities in French is to a large extent semantically arbitrary. A correlation also exists in certain languages between grammatical gender and phonology: nouns of a given gender will have certain specific phonological features. Although there are some regularities of this kind in French, there is no deterministic mapping between gender and phonology (Tucker, Lambert, & Rigault, 1977).

For these reasons, French grammatical gender is likely to be specified explicitly in the lexical system as a word-specific property. This feature of grammatical gender, found in a number of other languages, has attracted much attention in the field of psycholinguistics (Friederici & Jacobsen, 1999; Schriefers & Jescheniak, 1999). Since gen-

¹ The production of certain kinds of words depends on the message to be communicated and on other sources of activation. See below for more details.

der is a genuinely lexical property, the study of its processing might provide crucial insights about the lexical processes involved in speech production and perception.

3. Gender effects in speech error corpora

One of the lines of study of grammatical gender has come from the analysis of speech error corpora. These corpora are collections of self-elicited speech errors, collected from natural speech productions. One much studied type of error is the word substitution error, as for example in “All I want is something for my *shoulders*” where *elbows* was intended (Garrett, 1975, 1984). In the vast majority of cases, word-substitutions involve words of the same grammatical category: nouns exchange with nouns (see previous example), verbs with verbs, etc. More recently, in studies of gender marking languages, a further property of noun substitutions has been reported. Generally, noun substitutions preserve grammatical gender. That is to say, when a noun is used in place of another, the two words tend to be of the same gender (Arnaud, 1999; Berg, 1992; Marx, 1999). For instance, a speaker uttered “qu’est ce que j’ai fait de ma / sErv.../ | de ma cravate?” (*what have I done with my napkin | with my necktie?*) where both *napkin* and *necktie* are feminine nouns in French. In Berg’s (1992) German corpus, gender was preserved in 81.4% of the word substitutions. In the French corpus studied by Arnaud (1999), gender was preserved for 86% of the word substitutions. Furthermore, in Marx’s (1999) corpus, the rate of gender preservation in noun substitutions was independent of the presence in the sentential context of an explicitly gender-marked word (e.g., an article). These figures of gender preservation are very high compared to what would be expected if grammatical gender was randomly selected according to its distribution in the French or German lexicons. For French, chance would roughly correspond to preservations in 50% of the cases, whereas for German it would roughly be 36.5%. This observation has therefore suggested that gender is significantly preserved in noun substitutions in these languages.

Note, however, that evaluating chance level in the occurrence of gender preservations might not be so simple. It is well-known that word substitutions are constrained by semantic factors: In many cases, words that are substituted tend to have similar meanings. Such constraints have to be taken into account when the baseline rate of gender preservation is calculated. To do so, the probability of preserving gender must be assessed conditionally to the probability of preserving semantics.² This issue could in principle be dealt with by evaluating the distribution of gender among potential candidate words. However, just as it has been shown for the analysis of the occurrence of mixed semantic and phonological errors, it is very difficult to determine a priori what candidates are to be considered as potential competitors in natural production settings (see for example Gagnon, Schwartz, Martin, Dell, & Saffran, 1997). Although semantically related items such as the members of a

² In the extreme (but illustrative) case where all semantic candidates to a word substitution have the same grammatical gender, that substitution will preserve gender without any need for a gender-based mechanism.

given semantic category (i.e., coordinates) could be considered as the only legitimate competitors, things need not be so simple. The semantic similarity driving speech errors does not need to be strictly based on semantic coordination (see the previous *napkin-necktie* example). Overall, then, determining the appropriate baseline for assessing the magnitude of grammatical gender preservations is not without problems.

In spite of these difficulties, let us assume at this point that the preservation of gender over and above chance levels is a real phenomenon in slips-of-the-tongue corpora. The explanation that has been proposed for this observation is that when an error occurs, the target word and its gender are at least partially activated. The presence of this activated gender would produce a constraint on the activation and/or selection of a noun for the utterance (Arnaud, 1999; Berg, 1992). One implementation of this hypothesis is to postulate an activation link between gender representations and lexical representations. By favoring nouns of a particular gender, this activation increases the probability that the intruder noun is of the same gender as the intended target word. Vigliocco, Vinson, Indefrey, Levelt, and Hellwig (2004) have questioned this proposal on the basis of experimentally elicited speech error patterns. In their interpretation of gender preservation effects, it is not the representation of grammatical gender per se that produces the feedback but rather a morpho-syntactic frame that is retrieved on the basis of the properties of the noun (see also Vigliocco, Lauer, Damian, & Levelt, 2002). Irrespective of this aspect, this interpretation of gender preservations makes two critical related assumptions. First of all, it assumes that information other than semantic (lexical information) could contribute to the processes of activation and/or selection during lexical access. Secondly, as a correlate of the previous assumption, activation is hypothesized to flow bidirectionally between lexical representations and syntactic representations. This type of hypothesis is not allowed in language production models that allow feed forward activation only (e.g., Levelt et al., 1999). It is more in line with interactive activation architectures where a two-way information flow between levels (e.g., lexical and phonological) is allowed (e.g., Dell, 1986, and subsequent models).

In the interpretation we have discussed, a word might be sufficiently activated for its gender to impose a constraint on the selection of other lexical items (hence the gender preservation), without being sufficiently activated to be selected and produced. Although this proposal may seem paradoxical, the distinction between activation and selection has been used to account for another well-known situation in speech production. Speakers of a gender marking language who are in the tip of the tongue (TOT) state have been repeatedly reported to be able to retrieve the grammatical gender of words they could not produce (Caramazza & Miozzo, 1997; Vigliocco, Antonini, & Garrett, 1997). The interpretation of this observation is that some aspects of the word representation (e.g., gender) have passed an activation threshold, whereas others (e.g., word form) have not. If a somewhat similar situation occurred in the flow of speech, a possible consequence would be a lexical mis-selection influenced by the activated gender in some cases (and a TOT in other cases).

The reported preservation of grammatical gender in noun substitutions happens in the course of erroneous speech production. As is the case with other observations

made in speech error corpora, it is important to evaluate whether the effects of gender observed during the production of erroneous sequences are also observed in cases where speech is fluent and error-free (Meyer, 1992). This is especially true for the effect of gender, as we have seen that there are inherent limitations to the detailed assessment of the baseline of the effect.

4. Some gender effects in speech production

One line of research has investigated gender processing in the context of determiner+noun phrase (NP) production (Alario & Caramazza, 2002; Costa, Sebastián-Gallés, Miozzo, & Caramazza, 1999; Miozzo & Caramazza, 1999; Schiller & Caramazza, 2003; Schriefers, 1993). Schriefers (1993) observed a gender congruency effect in NP production (e.g., producing the Dutch version of “the green table” where “the” is gender marked). His experiments used the picture word interference paradigm. He observed that participants were slower to produce an NP in the presence of a distractor word of a different gender than in the presence of a distractor of the same gender. Further research has shown that the gender congruency effect is only observed in languages where determiner form is not dependent upon the phonological properties of the context of its occurrence (e.g., in Dutch and German). In languages where such a dependency exists (e.g., many Romance languages), the congruency effect is not observed. This discrepancy is attributed to a cross-linguistic difference in the process of determiner selection. In those languages where phonology is required to specify determiner form, their selection will be delayed until phonology is available (i.e. the “late selection hypothesis”; see details in Miozzo & Caramazza, 1999). Furthermore, it has been determined that, when observed, the gender congruency effect in NP production is located at the level of determiner selection, rather than at the level of gender feature selection. In other words, participants are delayed in the incongruent conditions because of conflicting evidence about what determiner form they have to use, rather than about what grammatical gender they need to consider. A consequence of this is that the investigation of gender processing in this paradigm informs us about the role of grammatical gender in the selection of determiners rather than the role of grammatical gender in noun retrieval. Since we are interested in the possible role of grammatical gender in the selection of nouns we will now focus on studies of single noun production.

Some experiments reported by Bentrovato, Devescovi, D’Amico, and Bates (1999), Jacobsen (1999), and Jescheniak (1999) have addressed the issue of the influence of a gender-marked context on the picture naming task with the purpose of investigating noun production. In a typical trial of the experiments reported in these studies, participants named a picture shortly after processing a linguistic prime that could comprise a gender-marked word. The gender-marked word was of the same gender (gender congruent condition) or of a different gender (gender incongruent condition) than that of the picture to be named. In a third type of trials, the primes did not carry gender information (neutral condition). In this type of design, slower reaction times in the incongruent condition than in the neutral condition are

generally thought to involve post-access processing, for example prime–target integration strategies. By contrast, faster reaction times in the congruent than in the neutral condition could reflect activation processes within the lexicon (Friederici & Jacobsen, 1999). Following this rationale, and given our interest for the effect of the activation of grammatical gender in the lexical selection process, we will concentrate our discussion on the facilitation effects reported in the congruent conditions.

In Jacobsen's (1999) study, participants had to name a picture or a word that was preceded by an auditory prime. The prime was a short sentence fragment without semantic content that ended in a gender-marked determiner (e.g., the German version of "This is the..." where *the* is gender marked) or a sentence fragment that did not carry gender information (e.g., the German version of "Please say now..."). Results show both facilitatory and inhibitory effects in the picture naming task, and only a small inhibition effect in the word naming task. In Jescheniak's (1999) study, only pictures were used as targets. The pictures were preceded by single word contexts—determiners, presented in the visual or auditory modality. In the neutral condition, a determiner that was compatible with both genders was used. The results were largely independent of the modality of the prime: both modalities show a clear congruency effect with slower responses given in the gender incongruent condition than in the gender congruent condition. However, no clear facilitation effect emerged: only a small non-reliable difference between the gender-congruent and the gender-neutral conditions was observed. Finally, Bentrovato et al. (1999) manipulated the grammatical and semantic congruency of sentence primes in a similar paradigm. In this study, context effects were both facilitative and inhibitory with respect to a neutral baseline. Note, however, that the facilitatory effect was only observed when both dimensions—grammatical and semantic—were congruent and the inhibitory effect was only observed when both dimensions were incongruent. This result leaves some uncertainty about the origin of the effect.

From this review, it appears that facilitatory gender context effects in error-free speech production are not totally robust. Of the three studies we have reviewed, one found the effect, the other did not, and in the third the effect was dependent on the semantic nature of the context. This apparent lack of robustness is in contrast with the repeated observation of the gender congruency effect in speech-error corpora. To explain this difference, one could argue that the observation in speech-error corpora reflects in fact some specific phenomena whose effects are directly tied to the production of an error. If this were true, the occurrence of gender effects of this type would be closely tied to the very specific circumstances that lead to the production of the error. Before reaching such a conclusion, however, it might be relevant to consider the evidence gathered in other tasks that require lexical access, for example in speech *perception* studies.

5. Lexical access in speech perception

Clearly, the processes subtending lexical access during the perception and the production of speech differ in important respects. Still, it is notable that gender-marked

context effects have repeatedly been found in studies of lexical access in speech perception that were parallel to the production experiments described earlier, as well as in studies using different experimental paradigms (Bates, Devescovi, Hernandez, & Pizzamiglio, 1996; Dahan, Swingley, Tanenhaus, & Magnuson, 2000; Grosjean, Dommergues, Cornu, & Besson, 1994; Spinelli & Alario, 2002). For example, Dahan et al. (2000) conducted a speech perception study where participants' eye-movements were recorded while they heard auditory instructions related to a visual display composed of four pictures. In the absence of grammatical gender in the instructions, participants fixated pictures whose names shared initial sounds with the picture name used in the instruction more than pictures with phonologically unrelated names (a "cohort" effect). When a gender-marked article immediately preceded the noun in the instruction, participants no longer fixated cohort-consistent pictures whose name was gender-inconsistent with the target noun. This result was taken as an indication that the gender-marked determiner information, in conjunction with the first sounds of the noun to be heard, had an effect on the set of hypotheses considered for the recognition of that noun (see Dahan et al., 2000, for details).

6. The current study

Gender context effects in speech production are of potential interest for our understanding of lexical access in speech production. Given the repeated observation of gender context effects in speech error corpora and in perception studies, we decided to collect more evidence on this specific topic in a series of picture naming experiments. In the usual experimental settings, picture naming involves the production of fluent, error-free speech. In Experiment 1, gender-congruent and gender-incongruent conditions were created by pairing a gender-marked determiner presented as a prime with a picture whose name was of the same or of a different grammatical gender. Because we were interested in determining whether gender effects are facilitatory or not, an important issue was that of defining a neutral baseline to evaluate gender effects. In Experiment 2, we used three prime types: gender congruent, gender incongruent, and non-gender marked. In Experiment 3, we concentrated on the facilitatory effects by excluding the incongruent condition. This was done to ensure that the observations made in the congruent condition of Experiment 2 were not dependent on the presence of an incongruent condition (see below).

6.1. Experiment 1

In this experiment, we compared a condition where the target stimulus to be named was preceded by a gender congruent prime with a condition where the target was preceded by a gender incongruent prime. There were three types of primes: definite determiners ("le" or "la", both meaning *the*), indefinite determiners ("un" or "une", both meaning *a*), and possessive adjectives ("mon" or "ma", both meaning *my*). The targets were either pictures or written words.

On the basis of the results of Jescheniak (1999), Jacobsen (1999) and Bentrovato et al. (1999), we predicted longer picture naming latencies for the incongruent than for the congruent condition. The word naming conditions were included to provide a comparison with the theoretically relevant picture naming conditions. Contrary to picture naming, where lexical selection is required to articulate the response, word naming can be largely conducted on the basis of orthography to phonology conversion procedures. In principle, these conversion procedures should not be sensitive to grammatical information such as gender. Therefore, no congruency effect was expected for the word naming task.

6.2. Method

6.2.1. Participants

Forty native speakers of French volunteered for this experiment. They were students at René Descartes university. All reported normal or corrected to normal vision.

6.2.2. Materials

We selected 48 simple black and white drawings of common objects and their corresponding names as experimental materials. Concepts with arbitrary (non-semantic) grammatical gender were used. The pictures were taken from the French norms developed by Alario and Ferrand (1999). All of the pictures had a 100% name agreement (Alario & Ferrand, 1999). Half of the picture names were masculine and half of them feminine. Their mean written frequency was 30 occurrences per million (masc. 28; fem. 33; Content, Mousty, & Radeau, 1990) and their mean number of syllables 1.6 (masc. 1.8; fem. 1.5). In addition to the experimental stimuli, 10 training stimuli were used (these were pictures from the same source and their corresponding names). The pictures appeared in black on a 245×245 pixels white square, in the center of the computer screen. A complete list of the stimuli is provided in the Appendix.

Three kinds of determiners were used. These determiners follow a grammatical gender agreement rule that yields a masculine and a feminine form, with distinct orthography and phonology. They were: definite articles (masc. “le”, fem. “la”), indefinite articles (masc. “un”, fem. “une”), and possessive adjectives (masc. “mon”, fem. “ma”). They were written black on white using uppercase letters on the middle of a 50×245 pixels rectangle.

The experimental lists also included five “catch trials”. In these trials, the determiner was not followed by a target to name but by question mark (“?”), in which case participants had to report the determiner they had just seen. The inclusion of catch trials is not a standard procedure in picture naming priming paradigms. However, since the primes were incongruent in half of the trials, we included such catch trials so that participants were not tempted to ignore them altogether. We reasoned that the presence of catch trials would enhance the processing of the prime.

6.2.3. Design

Half of the participants received picture targets and the other half received word targets. Gender congruency was manipulated within participants and items. Target

format (either picture or word) was manipulated between participants and within items. During the experiment, every participant saw 48 determiner–target pairs (either det.–picture or det.–word), 24 of which were in a congruent context and 24 in an incongruent context. To achieve this, the 48 targets were divided in two groups of 24 each (12 masc. and 12 fem.), with paired characteristics. A first list of materials (List 1) was created by assigning gender congruent determiners to the first group of targets and gender incongruent determiners to the second group (an equal number, 16, of determiners of each kind was used). List 2 was created by a symmetric procedure. Participants were assigned to one of the two formats (picture or word) and to one of the two lists, alternatively, in their order of appearance in the laboratory.

6.2.4. Apparatus and procedure

Participants were tested individually. Context words and targets were presented on the center of the screen of a personal computer with a 70 Hz refresh rate. Participants in the picture condition saw the following sequence of stimuli: (1) a fixation cross appeared for 500 ms; (2) then, a white rectangle (245×50 pixels) on which the determiner prime appeared in black uppercase letters on the rectangle for 250 ms; (3) the rectangle and the word subsequently disappeared: the screen was blank for 64 ms; (4) the picture to be named replaced the rectangle and remained on the screen until the participant responded. The next sequence followed after a delay of 2.5 s. Participants in the word condition saw a similar sequence where the target picture was replaced by a target word. Participants were asked to concentrate on the middle of the screen (location of all stimuli), to silently read the determiner, and to name as rapidly and as accurately as possible the target. They were told that the determiner might or might not agree in gender with the target.

The experiment began with the presentation of 12 training trials, which included two catch trials. After that, the probe trials (including the catch trials) were presented in a randomized order different for each participant.

The computer recorded the naming times, measured from the target onset to the triggering of the voice key by the participants' response (Sennheiser MD211N microphone). The experimenter sat in the same room as the participant in order to check and note down the responses of the participant.

6.2.5. Results

Trials on which participants incorrectly named the picture (for example, if they did not use the modal name), hesitated noisily, repeated their utterances, or stuttered were regarded as errors. Three pictures (DUSTPAN, SKIRT, and STOOL) and one word ("nest") produced more than 25% of errors and therefore were excluded from the analysis altogether. Among the remaining items, we observed 69 errors (3.9% of the data). Latencies that were more than 2.5 standard deviations away from the mean of each participant were considered as outliers and were also excluded from further treatment (36 data points, 2.0% of the data). The catch trials elicited correct responses in all cases.

Mean naming latencies and error rates for each experimental condition are given in Table 1. Two analyses of variance (ANOVA) were conducted, one by

Table 1

Mean picture and word naming latencies (M , in ms) and mean error rates (% ERs) per condition in Experiment 1

Target	Prime	M	% ERs
Pictures	Congruent	733 (65)	4.8% (5.0)
	Incongruent	806 (110)	7.1% (4.8)
	Congruency effect	+73	+2.4%
Words	Congruent	482 (57)	2.6% (4.3)
	Incongruent	482 (62)	1.3% (2.4)
	Congruency effect	0	-1.3%

Standard deviations per condition are given in parenthesis.

participants (F_1) and one by items (F_2). These analyses show an effect of the target format ($F_1(1,38)=158.3$, $MSE=10400$, $p<0.01$; $F_2(1,43)=513$, $MSE=7429$, $p<0.01$); naming latencies to words were significantly shorter (482 ms) than naming latencies to pictures (769 ms). There was also an effect of congruency, with faster responses in the congruent (608 ms) condition than in the incongruent condition (644 ms; $F_1(1,38)=19.8$, $MSE=1338$, $p<0.01$; $F_2(1,43)=22.9$, $MSE=3161$, $p<0.01$). The interaction between the two factors was significant ($F_1(1,38)=20.1$, $MSE=1338$, $p<0.01$; $F_2(1,43)=20.3$, $MSE=3543$, $p<0.01$). Planned comparisons show that the congruency effect is present for picture targets (73 ms effect; $F_1(1,19)=21.2$, $p<0.01$; $F_2(1,43)=40.7$, $p<0.01$) but not for the word targets (F_1 and $F_2<1$).

Similar analyses were conducted on the error rates. There was an effect of target format ($F_1(1,38)=15.0$, $MSE=0.002$, $p<0.01$; $F_2(1,43)=15.1$, $MSE=0.004$, $p<0.01$), no effect of congruency (F_1 and $F_2<1$) and an interaction between the two factors ($F_1(1,38)=4.68$, $MSE=0.001$, $p=0.04$; $F_2(1,43)=3.69$, $MSE=0.004$, $p=0.06$). The interaction reflects a bigger difference between error rates in the congruent and incongruent conditions for the picture targets than for the word targets. However, planned comparisons failed to reach significance for both the pictures ($F_1(1,19)=2.86$, $MSE=0.002$, $p=0.1$; $F_2(1,43)=3.09$; $p=0.09$) and the words ($F_1(1,19)=1.85$, $MSE=0.001$, $p=0.19$; $F_2<1$).

6.3. Discussion

This experiment produced two basic results. First, word naming latencies were insensitive to the gender congruency of the prime. Secondly, a clear gender congruency effect was observed in the picture naming task where naming latencies were faster when the picture to be named was preceded by a gender congruent determiner (e.g., “la”—TABLE, “*the_{fem}*”—TABLE) than when it was preceded by an incongruent determiner (e.g., “le”—TABLE, “*the_{masc}*”—TABLE). These results indicate that the process of lexical retrieval required for performing picture naming could be influ-

enced by the availability of grammatical gender information. By contrast, the conversion procedures that allow the production of responses in the word naming task are insensitive to this type of context. Note, however, that in Experiment 1 we have compared congruent and incongruent conditions, but no baseline was included. Therefore, it is possible that the reported effect is due to the difficulty in processing the ungrammatical incongruent condition. As was said earlier, this type of inhibitory effect is generally attributed to post-lexical processes. The following experiments were designed to determine whether the gender congruency effect would be observed when baseline priming was used.

6.4. *Experiment 2*

In this experiment, we compared picture naming latencies when the pictures were preceded by a gender-congruent prime (e.g., “la”—TABLE, “*the_{fem}*”—TABLE), by a gender-neutral linguistic baseline (e.g., “chaque”—TABLE, “*each_{not marked}*”—TABLE), or by a gender-incongruent prime (e.g., “le”—TABLE, “*the_{masc}*”—TABLE). The main issue was whether the congruent condition would lead to faster times than the gender-neutral condition. Observing faster naming latencies in the congruent than in the neutral baseline would provide an indication that the effect of gender is not solely due to difficulties in processing the ungrammatical gender-incongruent condition, but possibly to activation induced by the correct gender information in the congruent condition.

6.5. *Method*

6.5.1. *Participants*

Twenty-eight participants of the same population than before volunteered for this experiment.

6.5.2. *Materials*

The same 48 pictures that were used in the preceding experiment were used here. The pictures could either be preceded by a gender congruent determiner (the same three types that were used previously), by a gender-neutral closed-class word (“chaque”, meaning *each*) or by an incongruent determiner (again, the same that were used previously).

6.5.3. *Design*

Only the picture naming task was used in this experiment. The design included one factor, prime–target relatedness, with three levels: congruent, gender-neutral, and incongruent. The factor was manipulated within participants and within items. Parallel to what was done in the previous experiment, three experimental lists including all conditions were created, and participants were assigned randomly to one of them. In this way, every participant saw all experimental conditions, but saw each target picture only once.

6.5.4. Apparatus and procedure

They were similar to those of the previous experiment.

6.6. Results and discussion

Errors and outliers were identified as in the previous experiment. Three items produced more than 25% of errors (**RING**, **SHIRT**, and **SKIRT**) and were excluded from the analysis. Among the remaining items, there were 52 errors (5.8% of the data) and 20 outliers (2.2% of the data). Catch trials elicited correct responses in (97 out of 100 trials). The results are presented on Table 2.

One way ANOVAs conducted on the naming latencies showed that there was a main effect of priming condition ($F_1(2, 52) = 7.25$, $MSE = 2387$, $p < 0.01$; $F_2(2, 86) = 8.79$, $MSE = 2641$, $p < 0.01$). Planned comparison (bilateral Student *t*-tests) showed that the gender-congruent condition lead to faster responses (22 ms) than the non-gender-marked condition ($t_1(26) = 2.30$, $p = 0.03$; $t_2(43) = 2.4$, $p = 0.02$). Also, the non-gender-marked condition lead to faster responses than the incongruent condition (29 ms; $t_1(26) = 1.98$, $p = 0.06$; $t_2(43) = 1.96$, $p = 0.06$). Similar analysis conducted on the error rates showed no effect of priming condition (F_1 and $F_2 < 1$).

Picture naming latencies were faster in the congruent than in the neutral and the incongruent conditions. This result indicates that the facilitation is most likely due to the difference of gender information present/absent in the three types of determiner primes. According to the rationale presented in Section 4, these results suggest that the activation of grammatical gender information can have an influence on lexical access in the course of error-free speech production.

One possible concern with the results of Experiment 2 is that the presence of the incongruent condition in the experiment induces participants in certain task-specific strategies. For example, they might compare or attempt to consciously “match” the prime with the retrieved lexical item and trigger their response on the basis of this comparison. This type of mechanism is generally postulated in accounts of the effects observed in the incongruent condition itself (Friederici & Jacobsen, 1999). However, the mere presence of this type of trials in the experiment might induce the use of such strategy in all experimental trials. If this were to be the case, the matching strategy

Table 2
Mean picture and word naming latencies (*M*, in ms) and mean error rates (% ERs) per condition in Experiment 2

Picture targets	<i>M</i>	% ERs
Congruent	703 (53)	6.7% (5.9)
Non-gender-marked	725 (63)	7.3% (6.0)
Incongruent	754 (101)	7.6% (6.9)
Congruency effect	+22	+0.6
Incongruency effect	−29	−0.3

Standard deviations per condition are given in parenthesis.

might affect in some undetermined way the triggering of the response even for congruent trials. To test this hypothesis, we conducted Experiment 3 where the incongruent condition was excluded. In this experiment, we compared picture naming latencies when the pictures were preceded by a gender congruent prime (e.g., “la”—TABLE, “*the_{fem}*”—TABLE) or by a gender neutral linguistic baseline (e.g., “chaque”—TABLE, “*each_{not marked}*”—TABLE).

6.7. Experiment 3

6.7.1. Participants

Twenty participants of the same population than before volunteered for this experiment.

6.7.2. Materials

The same pictures as before were used. They could either be preceded by a gender congruent determiner (the same three types that were used previously; e.g., “la”—TABLE, “*the_{fem}*”—TABLE) or by a gender neutral linguistic stimulus (“chaque”).

6.7.3. Design, apparatus, and procedure

The design, apparatus, and procedure paralleled those of the Experiment 1. The basic differences are the following: the incongruent condition was now replaced by the neutral condition and only pictures were used as targets.

6.8. Results and discussion

Errors and outliers were identified as in the previous experiments. Two items produced more than 25% of errors (DUSTPAN, and SHIRT) and were excluded from the analysis. Among the remaining items, there were 38 errors (4.1% of the data) and 25 outliers (2.7% of the data). The detail of the results is presented on Table 3. Bilateral student *t*-tests showed that the gender-congruent condition lead to faster responses (740 ms) than the neutral condition (775 ms; $t_1(19)=2.34$, $p<0.05$; $t_2(45)=2.67$, $p=0.01$). Similar analysis conducted on the error rates showed that the difference between the congruent (7.2%) and the neutral condition (6.3%) failed to reach significance ($t_1(19)<1$; $t_2(45)<1$).

Table 3
Mean picture and word naming latencies (*M*, in ms) and mean error rates (% ERs) per condition in Experiment 3

Picture targets	<i>M</i>	% ERs
Congruent	740 (109)	7.2% (7.9)
Non-gender-marked	775 (104)	6.3% (4.6)
Congruency effect	+35	−1.0%

Standard deviations per condition are given in parenthesis.

Experiment 3 showed that picture naming latencies were faster after a gender-marked article than after a neutral stimuli (a non-gender-marked determiner, 'chaque'). This result can be taken as an indication that the gender congruency effect comprises a robust and significant facilitatory component. This effect is independent of the presence of an incongruent condition in the experiment that might induce participants into specific checking mechanisms to license production.

7. General discussion

This study investigated the impact of a gender-marked context on the process of lexical selection during error-free speech production. We conducted three picture naming experiments in which the pictures to be named were preceded by linguistic or non-linguistic contexts that could either be gender marked or not. When the context was gender marked it could be congruent or incongruent with the target that followed it. Experiment 1 showed that participants were slower at naming pictures after an incongruent determiner (e.g., "le"—TABLE, "*the*_{masc}"—TABLE_{fem}) than after a congruent determiner (e.g., "la"—TABLE, "*the*_{fem}"—TABLE_{fem}). This result is a clear demonstration of a gender-marked context effect in error-free picture naming. However, it does not inform us on whether the effect is due to difficulties in processing the incongruent condition, or to a facilitation of performance in the congruent condition, or to a combination of these effects. Experiment 2 included a baseline condition in which the non-gender-marked determiner was used (e.g., "chaque"—TABLE, "*each*_{not marked}"—TABLE). Results showed that naming times were faster for the gender congruent priming condition than for the gender-neutral condition. The gender incongruent priming condition was the slowest of all three. Experiment 3 replicated the facilitation effect when no incongruent condition was used. Overall then, these results show a reliable facilitation effect of gender context in error-free picture naming.

The observation of a robust facilitatory gender context effects in the picture naming task is in line with the effects reported in the speech perception literature where gender-marked contexts have been shown to reliably affect lexical access (see references in the Introduction). However, our effect stands in contrast with some observations previously made in language production experiments. In experiments similar to those reported here, [Jescheniak \(1999\)](#) did not observe a reliable facilitation effect neither with visual nor with auditory primes (note, however, the study by [Jacobsen \(1999\)](#) where auditory primes produced significant facilitatory effects). In [Jescheniak's](#) study, participants were familiarized with the experimental stimuli: before the experiment proper participants saw all the pictures and were given feedback about their intended names. In our experimental procedure, participants were directly involved in the experimental trials, without a familiarization phase.³ The ab-

³ This was done because we selected pictures with high name agreement (100%) according to a norming study ([Alario & Ferrand, 1999](#)). The 100% name agreement figure was observed in an offline study where participants had up to 5 s to identify and name the depicted objects.

sence of a familiarization phase potentially increases the amount of processing required for identifying the depicted object and activating its name—it can increase uncertainty during the process of lexical selection. This characteristic of our experiments could allow the gender information activated by the prime to play a visible role. By contrast, in experiments where participants are trained with the experimental pictures, the activation of lexical items by semantic information would be more effective and less prone to the influence of gender.⁴

The hypothesis that the effect of the gender-marked context is inherently limited fits well with the observation of a gender preservation effect in speech error corpora. As discussed in the Introduction, noun substitutions likely occur in cases where the intended target noun does not reach a level of activation high enough to be selected, while various alternative candidate words are activated. In this type of situation the occurrence of the error can be attributed to the fact that semantic information alone does not eliminate the uncertainty on which item should be selected. A gender representation that has been sufficiently activated by the intended word could bias the selection process towards nouns of that same gender. Thus, the erroneously selected word will tend to be of the same gender than the intended target word. A critical point in this proposal is the fact that the gender effect is visible provided that semantic information—the major source of activation for noun retrieval—does not override it.

In the interpretation of the facilitatory effects reported here, the processing of the prime activates a given gender and this gender in turn has an influence on the process of lexical selection. How different, then, are the accounts of the gender effect observed in our experiments and the gender effect reported in speech error corpora? As discussed earlier, in the case of errors, it is the partial activation of the representation of the target word which allows the activation of a gender representation (e.g., there can be activation of the gender in the absence of activation of the phonological form). This activated gender in turn influences the activation levels of other candidate words. In the experimental effects, however, the activation of gender which influences lexical selection originates in the processing of the prime. In other words: the activation of gender information is thought to influence lexical selection in both cases, but the origin of this activation differs between the spontaneous speech situation and the experimental setting.

The interpretation we have suggested for the experimental effect is based on the idea that the representations activated by the prime directly contribute to the process of lexical selection, on top of the activation produced by the semantic system. One way of implementing this hypothesis would be to postulate independent representations for grammatical gender. The activation of such a representation—e.g., the representation for feminine—feeds activation directly to the nodes that are compatible

⁴ Uncertainty is probably not the only factor affecting how much influence a gender-marked context can have on the process of lexical access. In Jacobsen's (1999) experiment, where a facilitation effect was found, participants went through a familiarization phase. In his experiment, sentence fragments rather than single words were used as primes: as the author notes, this could have enhanced the influence of the context on the naming task.

with it in the lexical system. This feedback can be directed to all nouns of the activated gender: in this case, gender information contributes to the activation of lexical candidates. Alternatively, the feedback could be limited to the cohort of candidates activated by the semantic system: gender information would play a role in the process of lexical selection.

There is a clear parallel between this proposal and previous proposals in which activation flows bi-directionally between levels of representation. For example, Dell (1986) hypothesized that the activation of lexical candidates leads to the activation of their phonological properties. This activation in turn feeds back to the lexical level and can influence the process of lexical selection. Such proposal was made to account for the observation that mixed errors (lexical mis-selections that share both semantic and phonological features with the target word) often have a probability of occurrence that is higher than chance if lexical selection were made independently of the phonological properties of the words (Gagnon et al., 1997). The interpretation of gender effects in terms of feedback from gender nodes to lexical nodes is based on the exact same principles than the interpretation of the mixed error effect.

Some authors, for example Vigliocco et al. (2002, 2004), question the need for postulating a direct link from the gender representation to the lexical nodes. These authors conducted a series of speech production experiments designed to elicit errors such as word substitutions (Vigliocco et al., 2004). The gender information carried by the context (determiner) with which the responses had to be given was manipulated. Word substitutions occurring in the context of a noun phrase for which the determiner form was different for each gender showed a clear gender preservation effect, similar to the one observed in naturally occurring speech errors. By contrast, the effect was not observed when the determiner forms that had to be produced with the response nouns were similar across genders (i.e., a determiner form that can be used with nouns of two genders). According to the authors, this result indicates that gender information alone cannot be responsible for the context effects: the form of determiners through which gender surfaces could be involved in the occurrence of the effect. They further propose that the gender preservation effect could be produced by the retrieval of a morphophonologically specified frame that carries determiner information. Dahan et al. (2000) also distinguish between two interpretations of the gender-marked context effect in the discussion of their speech perception study. They propose that this type of effect could be due to a representation of the grammatical gender property of the prime (as we have argued earlier to account for our own results) or, alternatively, to the likelihood of co-occurrence of the phonological forms of the prime and the target. In the second interpretation, a higher co-occurrence between the gender-marked determiners and the target nouns than between the non-gender-marked determiners and the nouns would make them more compatible and hence produce better performance.

These alternative interpretations have in common the involvement of the *phonological form* of the gender-marked context words in the production of the effects. One method for evaluating this type of proposal requires the computation, in an appropriate corpus, of the frequencies of co-occurrence of the determiner–noun pairs used in the experiments. The relevance of that variable, and hence of the dimension “co-

occurrence of phonological forms”, would be established if that frequency measure systematically modulates the facilitation effect reported here. A more coarse measure would simply consider the single-word frequency of the different determiner forms used in our experiments: A more frequent determiner form is overall more likely to co-occur with just any word. Since we do not have access to reliable measures of co-occurrence, we conducted a post hoc analysis based on the frequencies of determiner forms. The average frequencies in a written corpus (New, Pallier, Ferrand, & Matos, 2001) of the determiner forms used the experiments reported here were as follows: definite articles (‘le’, ‘la’) 20895 occurrences per million;⁵ indefinite articles (‘un’, ‘une’) 10220, possessive adjectives (‘mon’, ‘ma’) 1121, indefinite determiner (‘chaque’) 483. In Experiment 2, we observed a 22 ms facilitation effect. This effect was homogeneously distributed across determiner forms: definite articles 25 ms, indefinite articles 18 ms, and possessive adjectives 26 ms. In Experiment 3, we observed a 35 ms facilitation effect. This effect was not homogeneous across determiner types. It was the strongest for definite articles (66 ms), smaller for indefinite articles (29 ms) and the smallest for possessive adjectives (12 ms). This ordering of effect sizes follows the ordering of determiner frequency. To summarize: in one experiment we found a facilitation effect which was independent of determiner frequency, whereas in the other determiner frequency might have modulated the size of the effect.

This post hoc observation suggests that it might be the case that frequency of co-occurrence determines, at least partially, the amount of facilitation that the primes induce in this paradigm. Our conclusions on this point can only be speculative, however. This is because of various reasons. First, there was no systematic pattern across experiments. Secondly, by lack of it, we did not use a measure of the frequency of the actual co-occurrence of determiner forms. The measure of prime single-word frequency is only a coarse indication of the likelihood of co-occurrence between prime and target. Furthermore, it is important to note that even if it turned out that more frequent determiners produce stronger effects, it would still be possible to argue for a gender activation based effect. This is because it is reasonable to postulate that the degree of activation of the gender information carried by the primes (and, hence, its effect) is dependent on its frequency. In the light of these considerations, we are at this point inclined to favor the gender activation based account above the phonological form based account.

A final point concerning the relation between the co-occurrence of prime and target in language use and the observation of the priming effects is the following. In the congruent and neutral conditions of our experiments, prime and target pairs always formed syntactically appropriate noun phrases. In other words, it was never the case that the prime was a gender-marked word (e.g., ‘she’) that cannot be followed by a

⁵ Note that the frequency reported for the definite articles is a notable overestimation because these orthographic forms are also used as pronouns and the two lemmas are counted as one entry in all available databases.

noun in French. The interpretation of the effect based in feedback from the gender node predicts that, provided that other gender-marked words such as ‘she’ activate their gender as much as determiners do, facilitation should be observed irrespective of the syntactic constraint. The data that we have collected do not allow evaluating whether the syntactic compatibility of prime and target is a requisite for the observation of the facilitation effect. If it turned out that the facilitation effect is modulated by this syntactic factor, the account given here should be modified. For example, it could be proposed that the processing of the prime not only activates its gender but also a minimal syntactic frame with which the compatibility of the noun is assessed prior to its selection. Clearly, the details of such a mechanism go beyond the objectives of the current article.

To conclude: In this study we report three picture naming experiments in which we observed a robust grammatical gender context effect in the picture naming task. When assessed against an appropriate baseline, the effect was shown to have both a facilitatory and an inhibitory component. These results are in line with previously reported gender context effects in speech error corpora, in some studies of error-free speech production, as well as in studies of lexical access during speech perception. We have suggested that the observation of a gender context effect in speech production could be closely tied to the presence of uncertainty during the process of lexical selection. When semantic information is not sufficient to clearly single out a target lexical item, gender information could play a constraining role in the process of lexical access. Although several interpretations of this context effect are possible, our discussion based on the available evidence leads us to favor the view that postulates a direct link from gender representations to lexical items. This previously proposed hypothesis provides a simple and efficient account of the data.

Acknowledgments

The authors wish to thank Johan Wagemans (editor) as well as Delphine Dahan (reviewer) and an anonymous reviewer for their insightful comments on previous versions of this work.

Appendix A. Materials used in the Experiments. In Experiment 1, both the words and their corresponding pictures were presented. In Experiments 2 and 3 only the pictures were used. The materials are drawn from [Cycowicz, Friedman, Rothstein, and Snodgrass \(1997\)](#), normed in French by [Alario and Ferrand \(1999\)](#).

French	Gender	English	French	Gender	English
Bougie	Fem	Candle	Camion	Masc	Truck
Bouteille	Fem	Bottle	Champignon	Masc	Mushroom
Cage	Fem	Birdcage	Chapeau	Masc	Hat
Casserole	Fem	Pot	Crayon	Masc	Pencil

Appendix A (continued)

French	Gender	English	French	Gender	English
Chemise	Fem	Shirt	Gant	Masc	Glove
Couronne	Fem	Crown	Serpent	Masc	Snake
Jupe	Fem	Skirt	Téléphone	Masc	Telephone
Pipe	Fem	Pipe	Verre	Masc	Glass
Bague	Fem	Ring	Balai	Masc	Broom
Chaussette	Fem	Sock	Banc	Masc	Bench
Fusée	Fem	Rocket	Croissant	Masc	French Croissant
Montre	Fem	Watch	Nid	Masc	Bird-Nest
Pelle	Fem	Dustpan	Panier	Masc	Basket
Pomme	Fem	Apple	Piano	Masc	Piano
Pyramide	Fem	Pyramid	Puits	Masc	Well
Selle	Fem	Saddle	Tabouret	Masc	Stool
Chaise	Fem	Chair	Bouton	Masc	Button
Chaussure	Fem	Shoe	Cigare	Masc	Cigare
Cheminée	Fem	Chimney	Couteau	Masc	Knife
Flèche	Fem	Arrow	Drapeau	Masc	Flag
Lampe	Fem	Lamp	Moulin	Masc	Windmill
Plume	Fem	Feather	Noeud	Masc	Bow
Règle	Fem	Ruler	Peigne	Masc	Comb
Robe	Fem	Dress	Tambour	Masc	Drum

References

- Alario, F.-X., & Caramazza, A. (2002). The production of determiners: Evidence from French. *Cognition*, 89, 179–223.
- Alario, F.-X., & Ferrand, L. (1999). A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments and Computers*, 31, 531–552.
- Arnaud, P. J. L. (1999). Target-error resemblance in French word substitution speech errors and the mental lexicon. *Applied Psycholinguistics*, 20, 269–287.
- Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Perception & Psychophysics*, 58, 992–1004.
- Bentrovato, S., Devescovi, A., D'Amico, S., & Bates, E. (1999). Effect of grammatical gender and semantic context on lexical access in Italian. *Journal of Psycholinguistic Research*, 28(6), 677–693.
- Berg, T. (1992). Productive and perceptual constraints on speech-error correction. *Psychological Research*, 54, 114–126.
- Bock, K., & Levelt, W. J. M. (1994). Language production. Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). New-York: Academic Press.
- Caramazza, A. (1997). How many levels of processing are there in lexical access. *Cognitive Neuropsychology*, 14, 177–208.
- Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: Evidence from the tip-of-the-tongue phenomenon. *Cognition*, 64, 309–343.
- Content, A., Mousty, P., & Radeau, M. (1990). BRULEX. Une base de données lexicales informatisée pour le français écrit et parlé. *L'Année Psychologique*, 90, 551–566.
- Corbett, G. (1991). *Gender*. Cambridge, UK: Cambridge University Press.

- Costa, A., Sebastián-Gallés, N., Miozzo, M., & Caramazza, A. (1999). The gender congruity effect: Evidence from Spanish and Catalan. *Language and Cognitive Processes*, *14*, 381–391.
- Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology*, *65*, 171–237.
- Dahan, D., Swingle, D., Tanenhaus, M. K., & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language*, *42*, 465–480.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283–321.
- Friederici, A. D., & Jacobsen, T. (1999). Processing grammatical gender during language comprehension. *Journal of Psycholinguistic Research*, *28*, 467–484.
- Gagnon, D. A., Schwartz, M. F., Martin, N., Dell, G. S., & Saffran, E. M. (1997). The origins of formal paraphasias in aphasics' picture naming. *Brain & Language*, *59*(3), 450–472.
- Garrett, M. F. (1975). The analysis of sentence production. In G. Bower (Ed.), *Psychology of learning and motivation* (Vol. 9) (pp. 133–177). New York: Academic Press.
- Garrett, M. F. (1984). The organization of processing structure for language production: Applications to aphasic speech. In D. Caplan, A. R. Lecours, & A. Smith (Eds.), *Biological perspectives on language* (pp. 172–193). Cambridge, MA: MIT Press.
- Grosjean, F., Dommergues, J.-Y., Cornu, E. D. G., & Besson, C. (1994). The gender-marking effect in spoken word recognition. *Perception & Psychophysics*, *56*, 590–598.
- Jacobsen, T. (1999). Effects of grammatical gender on picture and word naming: Evidence from German. *Journal of Psycholinguistic Research*, *28*, 499–514.
- Jescheniak, J. D. (1999). Gender priming in picture naming: Modality and baseline effects. *Journal of Psycholinguistic Research*, *28*, 729–737.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75.
- Marx, E. (1999). Gender processing in speech production: Evidence from German Speech errors. *Journal of Psycholinguistic Research*, *28*(6), 601–621.
- Meyer, A. S. (1992). Investigation of phonological encoding through speech error analyses: Achievements, limitations, and alternatives. *Cognition*, *42*, 181–211.
- Miozzo, M., & Caramazza, A. (1999). The selection of determiners in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 907–922.
- New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur Internet: LEXIQUE [A lexical database for contemporary French on the Internet: LEXIQUE]. *L'Année Psychologique*, *101*(3), 447–462.
- Schiller, N. O., & Caramazza, A. (2003). Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory & Language*, *48*(1), 169–194.
- Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 841–850.
- Schriefers, H., & Jescheniak, J. D. (1999). Representation and production of grammatical gender in language production: A review. *Journal of Psycholinguistic Research*, *28*, 575–600.
- Spinelli, E., & Alario, F.-X. (2002). Gender context effects on homophone words. *Language and Cognitive Processes*, *17*(5), 457–469.
- Tucker, G. R., Lambert, W. E., & Rigault, A. (1977). *The French speaker's skill with grammatical gender*. The Hague, Paris: Mouton.
- Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, *8*, 314–317.
- Vigliocco, G., Lauer, M., Damian, M. F., & Levelt, W. J. M. (2002). Semantic and syntactic forces in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 46–58.
- Vigliocco, G., Vinson, D. P., Indefrey, P., Levelt, W. J. M., & Hellwig, F. (2004). Role of grammatical gender and semantics in German word production. *Journal of Experimental Psychology: Learning Memory and Cognition*, *30*(2), 483–497.