Determiner Selection in Noun Phrase Production

Herbert Schriefers
Nijmegen University

Jörg D. Jescheniak
Max Planck Institute of Cognitive Neuroscience
and University of Potsdam

Ansgar Hantsch
Max Planck Institute of Cognitive Neuroscience

In 3 experiments, native speakers of German named pictures of 1 or 2 objects by producing singular or plural noun phrases consisting of a definite gender-marked determiner and a noun. When singular and plural determiners differed (masculine and neuter gender), naming latencies were longer for plural utterances than for singular utterances. By contrast, when singular and plural determiners were identical (feminine gender), no such effect was obtained. When participants produced bare nouns, the Gender × Number interaction disappeared. This pattern indicates that during the production of plural definite-determiner noun phrases, singular and plural determiners compete for selection. The resulting constraints on number and gender processing in noun phrase production are discussed in the framework of models of language production.

In many languages, definite determiners (and other lexical elements) depend on the corresponding noun's grammatical gender. Current psycholinguistic models of language production assume that grammatical gender is an arbitrary syntactic feature of nouns that is stored in the mental lexicon (e.g., Levelt, Roelofs, & Meyer, 1999). More specifically, grammatical gender is assumed to be stored at an abstract level of representation, the so-called lemma level. At this level, each noun of a given gender type is connected to one common gender node (e.g., Jescheniak & Levelt, 1994; Jescheniak, Schriefers, & Hantsch, 2001; Roelofs, 1992; Schriefers, 1993; see Schriefers & Jescheniak, 1999, for a review). Because grammatical gender is a lexically stored syntactic property, it provides an interesting test case for studying the interface between lexical processing and grammatical encoding; grammatical gender first has to be retrieved from the mental lexicon and is then used to retrieve corresponding gender-marked elements, such as definite determiners. The latter process is also referred to as *indirect lexical selection* because the selection of the definite determiner depends on the retrieval of a grammatical property of another lexical element, in this case a noun.

Schriefers (1993) investigated these processes in a variant of the picture–word interference task. Native speakers of Dutch described colored line drawings of common objects by producing noun phrases consisting of a gender-marked definite determiner, an adjective, and a noun (e.g., *de rode tafel* [the red table]). In addition, participants were presented with distractor words. These distractors either had the same grammatical gender as the noun of the target utterance (gender-congruent condition) or a different grammatical gender (gender-incongruent condition). The results showed significantly longer naming latencies in the gender-incongruent condition than in the gender-congruent condition (see also La Heij, Mak, Sander, & Willeboordse, 1998; Schiller & Caramazza, 2002, and van Berkum, 1997, for replications in Dutch, and Schiller & Caramazza, 2002, and Schriefers & Teruel, 2000, for replications in German).[1] Schriefers (1993) interpreted this gender-congruency effect as resulting from a competition in selecting the abstract gender feature of the target noun; in the gender-congruent condition, target noun and distractor activate the same gender feature, whereas in the gender-incongruent condition, they activate different gender features. Therefore, in the latter condition, two gender features compete for selection, leading to prolonged naming latencies.

However, Miozzo and Caramazza (1999) pointed out that the effect could also result from competition among determiners rather than competition among abstract gender features. Recent evidence in support of such a determiner-competition account comes from a study by Schiller and Caramazza (2002) on German and Dutch noun phrase production. Schiller and Caramazza's experiments, as

[1] The gender-congruency effect has not been found for Italian, Spanish, and Catalan. For an overview and a discussion of the potential reasons for this difference between the two sets of languages see Caramazza, Miozzo, Costa, Schiller, and Alario (2001).

well as the present experiments, made use of the fact that singular definite determiners in German (and Dutch) are gender marked, whereas the gender marking is neutralized for plural definite determiners. For example, German, the language used in the present study, has three gender classes: masculine, feminine, and neuter, with the corresponding singular definite determiners *der*, *die*, and *das*, respectively. In the plural, however, the definite determiner for all three gender classes is *die*, which is identical in form to the singular feminine definite determiner.

In their experiments, Schiller and Caramazza (2002) observed a gender-congruency effect for singular noun phrases with gender-marked definite determiners, replicating the results mentioned above. However, for plural noun phrases, no gender-congruency effect was obtained. Schiller and Caramazza interpreted these results as follows: They argued that the target noun's abstract gender feature is retrieved irrespective of whether speakers eventually produce a singular noun phrase or a plural noun phrase. In their view, gender retrieval is a noncompetitive process and gender comes for free whenever the lemma of the target noun is selected (Schiller and Caramazza, 2002). By contrast, the subsequent selection of the corresponding definite determiner is assumed to be a competitive process. When a speaker produces a singular noun phrase in the gender-incongruent condition, two different determiners activated by the target noun and by the distractor compete for selection. No such competition is assumed to occur in plural noun phrases because the plural definite determiners for the three gender classes are identical.[2]

Evidently, this interpretation crucially depends on the assumptions one makes about which determiners a singular distractor will activate. That is, does a singular distractor activate its singular determiner only or also its plural determiner, and how does this interact with the number feature of the target utterance? Because there is no answer to this question available yet, we took a different approach to the issue of determiner competition in the present study. We used a simple object-naming task, in which participants produced singular or plural gender-marked definite noun phrases in response to one or two target objects. This task entailed production processes only, such that valid inferences should have been more easily achieved. The only assumption we needed to make was that the form-identical yet syntactically different determiners share a representation on at least one level in the mental lexicon (i.e., at the lemma level, the word form level, or both). It is important to note that with respect to the issue at hand, all these options lead to the same predictions. With this qualification in mind, we next develop our specific hypotheses and defer to the General Discussion the question of at which level the representational overlap is to be located. To anticipate the results, the shared representation assumption was supported by the experimental data.

In contrast to the lexically specified and hardwired gender feature, the number feature is variable and dependent on properties of the to-be-expressed message. That is, if the speaker wants to talk about one exemplar of a kind, the number feature has to be specified as singular. By contrast, if the speaker wants to talk about multiple exemplars of a kind, the number feature has to be specified as plural. Considering the production of gender-marked singular and plural noun phrases raises the question of how the lexically specified gender feature and the conceptually specified number feature are used for retrieving the correct definite determiner. In the following paragraphs, we discuss three hypotheses

and outline their predictions for a production task in which participants produce simple singular or plural noun phrases with definite determiners. These three hypotheses are different instantiations of the view that determiners compete for selection (e.g., Caramazza et al., 2001).

The first hypothesis is referred to as the *number-dominance hypothesis.* According to this hypothesis, the number feature dominates the gender feature. For a singular noun phrase, the number feature will restrict the set of possible determiners to the three singular determiners. The noun's gender feature is then used to select the target determiner from this restricted set. For plural noun phrases, the number feature will activate only one determiner, namely *die*. As a consequence, any difference in the speed of producing a singular noun phrase versus a plural noun phrase should be independent of the actual grammatical gender of the target noun. Furthermore, the selection of the plural determiner should be faster than the selection of the singular determiner, because the latter selection process comprises the selection of one out of three candidates via the gender feature, whereas the former selection process comprises only a one-step selection of the plural determiner via the number feature *plural*. Thus, when comparing utterance-onset latencies for singular and plural noun phrases from the three different gender classes, the number-dominance hypothesis predicts a main effect of number, with faster utterance-onset latencies for plural noun phrases than for singular noun phrases. There might also be a main effect of gender class because the three gender classes necessarily comprise different to-be-named pictures. Such an effect is not critical for the present hypothesis. But clearly the number-dominance hypothesis does not allow for any differential influence of gender class on the predicted singular–plural difference (i.e., there should be no Gender Class × Number interaction, provided that the ease of plural formation is comparable for the items used in the three gender class conditions). The latter qualification emphasizes the need for controlling for potential differences in the ease of plural formation independently of determiner selection. We return to this point below.

The number-dominance hypothesis can be contrasted with what is referred to as the *gender-dominance hypothesis.* According to the latter hypothesis, the gender feature of the target noun will activate the corresponding singular and plural determiners. The number feature will then select the appropriate target determiner. For target nouns of masculine and neuter gender, the gender feature will activate two determiners (*der* and *die* for singular and plural masculine nouns and *das* and *die* for singular and plural neuter nouns); the number feature will guide the eventual selection between these two determiners. For nouns of feminine gender, the gender feature will activate only one determiner (*die* for singular feminine and plural feminine). Thus, for singular and plural noun phrases with masculine or neuter nouns, two determiners will initially compete, and the number feature will guide the selection of the eventual determiner. For singular and plural noun phrases with feminine nouns, by contrast, no such competition will be

---

[2] Actually, the situation is more complex in German if one takes into consideration that articles also differentiate case. However, in the following, we focus on utterances probed in the present experiments, that is, noun phrases in nominative case. We return to the issue of case marking in the General Discussion.

present. Thus, the gender-dominance hypothesis predicts a main effect of gender class on utterance-onset latencies, with noun phrases with feminine nouns being initiated faster than noun phrases with masculine and neuter nouns. This main effect should be independent of the factor number (singular vs. plural). Again, a potential main effect of gender class is difficult to interpret on its own because it necessarily implies a comparison between different items. But because the presence or absence of this effect is crucial in the evaluation of the present hypothesis, one must obtain a baseline measure identifying possible differences in retrieval speed for the items in the three gender classes, independent of determiner retrieval. This baseline can be assessed by asking speakers to name pictures by simple singular or plural bare nouns (i.e., without determiners). This control task allowed us to determine which parts of a potential main effect of gender class are due to determiner selection and which parts are due to general between-item differences. In addition, it revealed whether the ease of plural formation is comparable for the items used in the three gender class conditions, which is important in evaluating the number-dominance hypothesis.

The third hypothesis is referred to as the *singular-as-default hypothesis*. According to this hypothesis, *singular* acts as the default value of the number feature. Irrespective of whether a picture eventually requires a singular or a plural noun phrase, it will always activate its lemma, its gender, and its corresponding singular definite determiner. Only when the number feature *plural* is activated (e.g., by presenting two exemplars of the same object as targets), will the number feature *plural* activate the corresponding plural determiner. Thus, for singular noun phrases, the correct singular determiner will become activated and selected without any competition or interference from other determiners. In the case of plural noun phrases with nouns of masculine or neuter gender, the noun will activate its corresponding singular determiner (*der* or *das* for masculine and neuter nouns, respectively), and the *plural* feature will activate the plural determiner, *die*. For plural noun phrases with feminine nouns, the activation from the noun lemma and from the plural feature will converge on the same definite determiner, *die*. As a consequence, production of noun phrases with masculine and neuter nouns in the plural will take longer than production of these noun phrases in the singular. For feminine nouns, by contrast, no such effect should obtain. Actually, for noun phrases with feminine nouns one might even expect an advantage of *plural* over *singular*, because in the former case the determiner *die* receives activation from the noun lemma and the number feature *plural*, whereas in the latter case it receives activation only from the noun lemma. Thus, in contrast to the other two hypotheses, the singular-as-default hypothesis predicts an interaction of gender class and number for utterance-onset latencies. This interaction should be due to higher costs associated with the transition from singular to plural noun phrases for masculine and neuter gender, relative to feminine gender.

The proposal that *singular* might function as the default specification of the number feature is related to results from studies on the computation of number agreement between subject and verb. In these studies, number-agreement errors were elicited experimentally by asking participants to complete sentence beginnings consisting of a subject noun phrase followed by a so-called local noun (e.g., a prepositional phrase). These experiments consistently showed that a sentence beginning with a singular subject noun followed by a plural local noun (e.g., "the baby on the blankets") elicited more verb-agreement errors (i.e., plural verbs) than did the corresponding control condition with the local noun in singular (e.g., "the baby on the blanket"). By contrast, sentence beginnings with a plural subject noun and a singular local noun did not elicit more verb-agreement errors than did the corresponding control condition with subject noun and local noun in plural. This asymmetry has been explained by assuming that the number feature *singular* functions as a default value (see Bock, 1995, for an overview).

In the following experiments, the predictions from the three hypotheses described above were put to an experimental test. In Experiment 1, native speakers of German named pictures of simple objects by means of singular or plural noun phrases with definite determiners. In this experiment, the frequency of occurrence of the three determiners (*der*, *die*, and *das*) was controlled. Experiment 2 was a replication of Experiment 1 with the only difference that participants produced bare nouns in place of definite noun phrases; this tested for gender-class differences in the ease of producing singular versus plural forms that might have contributed to the pattern of results of Experiment 1.

Given the German gender system, it is logically impossible to control the frequency of occurrence of the different determiners and the frequency of occurrence of singular and plural trials within one experiment. Therefore, in Experiment 3 we replicated Experiment 1, but we controlled the frequency of occurrence of singular and plural trials rather than the frequency of occurrence of the three determiners.

In all experiments, singular and plural trials were cued in a perceptually similar way. Rather than presenting one versus two objects, two exemplars of an object were presented side by side on each trial. The to-be-named object (in a singular trial) or objects (in a plural trial) were identified by their color. If only one object was foregrounded by its black color (whereas the other object was presented in medium gray), the participant was instructed to produce a singular noun phrase. If both objects were foregrounded by their black color, the participant was instructed to produce a plural noun phrase. This procedure was used to make the plural trials perceptually less salient, thus preventing participants from initiating the articulation of the plural determiner, after having detected the presence of two objects, on the basis of some superficial visual cues only and without having retrieved the object name from the mental lexicon yet.

Experiment 1

In Experiment 1, participants named one or two target objects by producing either a definite singular noun phrase or a definite plural noun phrase.

Method

Participants. Sixteen native speakers of German, most of them students from the University of Leipzig, Leipzig, Germany, took part in the experiment. In all experiments described here, participants were paid DM 13 (approximately U.S. $5). They had no known hearing deficit, and they had normal or corrected-to-normal vision.

Materials. We used line drawings of 60 different objects that each have one unambiguous name, and an equal number of line drawings of objects that have names with masculine, feminine, and neuter gender. With

one exception, the object names were identical to those used in a study by Schiller and Caramazza (2002); see the Appendix for a complete list. All line drawings were sized to fill a square of approximately 45 × 45 mm. During the main experiment, two exemplars of a picture were presented side by side as double-object pictures. Singular trials were cued by presenting one target exemplar in black (RGB: 0, 0, 0; in the stimulus-presentation system used, RGB color codes vary from 0 to 63) and a second exemplar in medium gray (RGB: 42, 42, 42) on a light gray background (RGB: 60, 60, 60). Plural trials were cued by presenting both exemplars in black. Each experimental picture appeared once in the singular condition and once in the plural condition, yielding the following distribution of determiners: 20 occurrences of *der*, 20 of *das*, and 80 of *die* (20 occurrences of *die* from singular noun phrases with feminine nouns and 60 of *die* from the plural noun phrases across all three gender classes). To equate the probability of occurrence for the three determiners, we included another 60 filler objects, half with names of masculine gender and half with names of neuter gender. Each of these filler objects was presented twice in a singular trial. This way, for the whole set of items all three determiners appeared equally often, namely 80 times each. For the construction of a practice block to be presented at the beginning of the main experiment, 4 additional objects with names of feminine gender were selected. Together with 8 of the filler objects with masculine names and 8 of the filler objects with neuter names, these items were used for constructing 24 practice trials.

Design. There were two crossed variables: the three-level variable gender (masculine vs. feminine vs. neuter) and the two-level variable number (singular vs. plural). Both variables were tested within participants. Gender was tested between items, and number was tested within items.

Each participant received each of the 60 experimental pictures exactly once in the singular condition and once in the plural condition, resulting in 120 experimental trials. For the singular trials, the position of the target object (left vs. right) was systematically counterbalanced. The same held for the presentation sequence of the singular and plural condition. Across groups of participants, the items were reassigned to these four conditions in a way such that, overall, each item occurred in each condition equally often. The 120 experimental trials were intermixed with the 120 filler trials.

Eight different experimental lists were created according to the following general criteria: (a) Repetitions of an object were separated by at least eight intervening trials, (b) semantically or phonologically related objects did not appear in adjacent trials, (c) no more than five trials from the same gender class or with the same determiner followed each other, and (d) no more than five plural trials or five singular trials with the same position of the target object followed each other. The eight experimental lists were used equally often.

Procedure. Each participant was tested individually in a session lasting about 45 min. The participant was comfortably seated in a dimly lit room and separated from the experimenter by a partition wall. The visual stimuli were presented centered on a 17-in. (43.18-cm) monitor. Viewing distance was about 60 cm. The presentation of the visual stimuli and the online collection of the data were controlled by a Pentium computer. Speech-onset latencies were measured to the closest millisecond with a voice key connected to the computer.

In the main experiment, a double-object picture was presented for 1 s. Participants named the picture as quickly as possible by producing either a singular noun phrase (if only one of the objects in the double-object picture appeared in black; e.g., "der Baum" [the tree]) or a plural noun phrase (if both objects appeared in black; e.g., "die Bäume" [the trees]). Speech-onset latencies were measured from the onset of the target picture; 750 ms after a response was registered, the next picture was presented. If no response was registered within 2 s, the next picture was presented with a delay of 750 ms.

The actual experiment consisted of three parts: a study phase, a training phase, and the main session. During the study phase participants studied a written instruction booklet that emphasized both the speed and accuracy of their responses. Participants also previewed all pictures and their names on the monitor. Each picture was presented in black in the center of the screen along with its name (as bare noun; i.e., without the definite determiner), and participants were instructed to only use these names during the experiment. During the training phase, each of the 120 experimental and filler items was presented once as a single object in black in the center of the monitor, and participants named them by producing singular bare nouns. The main experimental session started with the 24 practice trials followed by the 240 experimental and filler trials. After half of these trials, there was a short break.

Results and Discussion

Observations were discarded from the reaction-time analyses whenever any of the following conditions held: (a) A picture had been named with other than the expected name; (b) a nonspeech sound preceded the target utterance, triggering the voice key; or (c) a dysfluency occurred or an utterance was repaired. Observations deviating from a participant's and an item's mean by more than two standard deviations were considered outliers and discarded from the reaction-time analyses. According to these criteria, 115 observations (6.0%) were marked as erroneous and 38 observations (2.0%) as outliers.

Averaged reaction times were submitted to analyses of variance (ANOVAs), involving the two fixed variables: number (singular vs. plural) and gender (masculine vs. feminine vs. neuter). Two complementary analyses were computed, each treating participants and items, respectively, as a random variable (Clark, 1973). In the following, we adopt an alpha level of $p < .05$. The p value of a statistical test is reported only if it did not reach this level.

Table 1 displays mean reaction times and error rates for the experimental trials, broken down by number and gender class. Plural noun phrases were produced significantly slower than sin-

Table 1

Mean Reaction Times (in ms) and Error Rates (in Percentages) by Number and Gender for Experiment 1

	Gender											
	Masculine				Feminine				Neuter			
Number	M	SE	%	SE	M	SE	%	SE	M	SE	%	SE
Singular	653	12	4.7	1.2	655	17	3.1	0.9	647	15	5.6	1.5
Plural	673	18	10.9	2.0	649	15	4.1	1.3	681	17	7.5	1.5
Difference[a]	20	11	6.2	2.7	−6	7	1.0	1.7	34	9	1.9	2.2

[a] Difference = plural minus singular.

Table 2
Mean Reaction Times (in ms) and Error Rates (in Percentages) by Number and Gender for Experiment 2

| | Gender | | | | | | | | | | | |
| | Masculine | | | | Feminine | | | | Neuter | | | |
Number	M	SE	%	SE	M	SE	%	SE	M	SE	%	SE
Singular	688	10	4.1	1.2	672	9	2.2	0.9	701	9	4.7	1.0
Plural	696	10	10.6	1.6	693	12	4.1	0.9	712	13	7.8	2.1
Difference[a]	8	6	6.5	1.6	21	6	1.9	1.0	11	9	3.1	2.5

[a] Difference = plural minus singular.

gular noun phrases (668 vs. 652 ms), $F_1(1, 15) = 6.08$, $MSE = 974.20$; $F_2(1, 57) = 6.16$, $MSE = 1,163.00$. They were also associated with more errors (7.5% vs. 4.5%), $F_1(1, 15) = 3.51$, $MSE = 2.49$, $p = .081$; $F_2(1, 57) = 6.21$, $MSE = 1.13$. Whereas there was no reliable effect of gender in the analysis of reaction times, $F_1(2, 30) = 1.61$, $MSE = 832.76$, $p = .22$; $F_2 < 1$, masculine nouns elicited more errors than did either feminine or neuter nouns. This effect was reliable in the participant analysis but only marginally significant in the item analysis (7.8% vs. 3.6% vs. 6.6%), $F_1(2, 30) = 7.32$, $MSE = 0.82$, $p < .01$; $F_2(2, 57) = 2.37$, $MSE = 2.03$, $p = .10$. Most important, in the analysis of reaction times, the Gender Class × Number interaction was reliable, $F_1(2, 30) = 6.82$, $MSE = 479.76$; $F_2(2, 57) = 3.32$, $MSE = 1,163.00$. T tests revealed that the increase in naming latencies from singular to plural noun phrases was marginally significant for nouns of masculine gender (20-ms difference), $t_1(15) = 1.82$, $p = .09$; $t_2(19) = 2.02$, $p = .06$, and reliable for nouns of neuter gender (34-ms difference), $t_1(15) = 3.98$; $t_2(19) = 2.56$.[3] By contrast, for nouns of feminine gender, there was a trend in the opposite direction, which, however, was not significant (6-ms difference; both ts < 1).

The significant Gender Class × Number interaction supports the singular-as-default hypothesis and is in conflict with the predictions from the other two hypotheses. Furthermore, the interaction is precisely of the type predicted by the singular-as-default hypothesis. There were reaction-time costs for plural noun phrases with masculine and neuter nouns relative to the corresponding singular-noun phrases, whereas no such cost was obtained for noun phrases with feminine nouns.

However, as pointed out above, the gender-class comparisons are inevitably between-item comparisons. Therefore, we performed Experiment 2 as a control experiment. It was identical to Experiment 1, except for the utterance format used; in Experiment 2, participants were instructed to name the pictures by bare singular or plural nouns (i.e., without gender-marked determiners). The predictions for Experiment 2 were straightforward. If the Gender Class × Number interaction observed in Experiment 1 was due to determiner competition in the way proposed by the singular-as-default hypothesis, rather than gender-class differences in the ease of producing singular versus plural forms, this interaction should not obtain in Experiment 2.

Experiment 2

Experiment 2 compared the production of bare singular nouns with the production of bare plural nouns, in which no determiner

selection was required, using the same materials and the same procedure as Experiment 1. It allowed us to test whether the Gender Class × Number interaction obtained in Experiment 1 was due to differences between the items in the three gender classes, such as, for example, the relative ease of plural formation in the three gender classes.

Method

Participants. Sixteen students from the University of Leipzig took part in the experiment.

Materials. The materials were identical to those used in Experiment 1.

Design. The design was identical to the one used in Experiment 1.

Procedure. The procedure was the same as the one used in Experiment 1, with the only exception being that in the main experiment participants were instructed to produce singular nouns (e.g., "Baum" [tree]) if only one of the objects in the double-object picture was presented in black and to produce plural nouns (e.g., "Bäume" [trees]) if both objects were presented in black.

Results and Discussion

The raw data were treated as in Experiment 1. According to these criteria, 107 observations (5.6%) were marked as erroneous and 31 observations (1.6%) as outliers.

Table 2 displays mean reaction times and error rates for the experimental trials, broken down by number and gender class. As in Experiment 1, plural noun phrases were produced significantly slower than singular noun phrases (700 vs. 687 ms), $F_1(1, 15) = 8.26$, $MSE = 484.79$; $F_2(1, 57) = 7.37$, $MSE = 954.48$, and elicited more errors (7.7% vs. 3.7%), $F_1(1, 15) = 32.54$, $MSE = 0.44$; $F_2(1, 57) = 5.80$, $MSE = 1.97$. Although gender was significant in the participant analysis, both in the analysis of reaction times, $F_1(2, 30) = 10.20$, $MSE = 459.78$; $F_2 < 1$, and in the analysis of error rates, $F_1(2, 30) = 5.62$, $MSE = 1.09$; $F_2(2, 57) = 2.07$, $MSE = 2.33$, $p = .14$, this effect was not confirmed in the item analyses. Most important, in contrast to Experiment 1, there was no Gender Class × Number interaction, neither in the analysis of reaction times, $F_1(2, 30) = 1.01$, $MSE = 349.85$, $p =$

---

[3] The fact that the 20-ms effect for nouns of masculine gender was not reliable might be explained by the fact that for this item group the number effect in part manifested itself in the error-rate pattern. In fact, the 6.2% increase from singular to plural noun phrases was reliable in both the participant and item analysis, $t_1(15) = 2.33$, $p < .05$; $t_2(19) = 2.36$, $p < .05$. There was no such effect for the other two gender classes.

.38; $F_2 < 1$, nor in the analysis of error rates, $F_1(2, 30) = 1.42$, $MSE = 1.33$, $p = .26$; $F_2 < 1$.

The absence of a Gender Class × Number interaction in Experiment 2 indicates that the corresponding interaction in Experiment 1 was due to determiner competition as captured in the singular-as-default hypothesis. For masculine and neuter nouns, the selection of the plural determiner is hindered by the activation of the different singular determiners, whereas for feminine nouns, the selection of the plural determiner profits from the fact that the *singular* default and the number feature *plural* activate the same determiner.

The singular-as-default hypothesis also implies a specific prediction concerning the speed with which singular noun phrases can be produced. According to this hypothesis, only the singular determiner should be activated on singular trials, whereas the plural determiner is not activated. If this is correct, singular noun phrases with a definite determiner should be equally fast for all three gender classes. By contrast, according to the gender-dominance hypothesis, the plural determiner also becomes activated in producing singular noun phrases. If this is the case, feminine singular noun phrases with a definite determiner should be produced faster than the corresponding masculine or neuter singular noun phrases because for feminine singular noun phrases the singular determiner and the plural determiner are identical, whereas for masculine and neuter singular noun phrases, two different determiners should become activated. The results from Experiment 1 show no advantage of singular noun phrases with feminine nouns over singular noun phrases with either masculine or neuter nouns (for reaction times, both Fs < 1; for error rates, $F_1[2, 30] = 1.53$, $MSE = 0.67$, $p = .23$; $F_2 < 1$). However, the interpretation of this finding is complicated because possible effects of determiner retrieval might have been obscured by differences in the ease of item recognition, lemma retrieval, or both of these factors across the item sets used in the three gender conditions. In fact, when participants produced singular bare nouns in Experiment 2, there was some effect of gender, although it was only reliable in the participant analysis of reaction times, $F_1(2, 30) = 8.71$, $MSE = 389.92$; $F_2(2, 57) = 1.21$, $MSE = 4,282.71$, $p = .31$. To take this variable into account, we performed a joint analysis of the singular conditions of Experiment 1 and 2, with experiment as the between-participant factor and gender class as the within-participant factor. This analysis revealed no significant main effect of gender (for reaction times, $F_1[2, 60] = 1.57$, $MSE = 578.83$, $p = .22$; $F_2[2, 57] = .29$, $MSE = 6,731.01$, $p = .75$; for error rates, $F_1[2, 60] = 3.01$, $MSE = 0.70$, $p = .06$; $F_2[2, 57] = 1.80$, $MSE = 0.93$, $p = .17$), but it did reveal a significant interaction of these two factors for reaction times, $F_1(2, 60) = 4.78$, $MSE = 578.83$; $F_2(2, 57) = 4.27$, $MSE = 771.56$. This pattern reflects a smaller difference between naming latencies for singular noun phrases and singular bare nouns for items of feminine gender than for items of either masculine or neuter gender, resulting from particularly fast bare-noun-naming latencies for the former item set. Subsequently computed t tests on reaction times revealed significant differences between singular bare nouns and singular noun phrases for all but the feminine nouns in the participant analysis. It is important to note that this pattern runs exactly opposite to what one would predict if plural determiners were activated during singular noun phrase production. In such a case, the difference should have been particularly large for items of feminine gender.

Together, the results of Experiments 1 and 2 suggest that in producing a plural noun phrase with definite determiners both the singular and the plural determiner are activated, whereas in producing a singular noun phrase only the singular determiner becomes activated. This is in line with the predictions of the singular-as-default hypothesis.

Before turning to the General Discussion, we report one more experiment. In Experiments 1 and 2, the proportion with which the three determiners (*der* [singular masculine), *das* [singular neuter], and *die* [singular feminine and plural for all three gender classes]) occurred was kept equal, to avoid the fact that participants would be biased to strategically prepare some particular determiner. That is, across critical and filler trials, each determiner occurred 80 times. However, controlling for the frequency of occurrence of the three determiners inevitably led to an imbalance of singular and plural noun phrases. Across critical trials and filler trials, there were 180 singular trials and only 60 plural trials. Obviously, given the German determiner system, it is logically impossible to balance the occurrence of the specific determiners and the occurrence of singular and plural trials simultaneously. However, it could be argued that the predominance of singular trials in the previous experiments artificially induced the use of the singular feature as a default. To rule out this possibility, we performed Experiment 3. It was a replication of Experiment 1, with the only difference being that this time the proportion of singular and plural trials rather than the proportion of the different determiners was controlled. This was accomplished in Experiment 3 by eliminating the filler trials of Experiments 1 and 2. As a consequence, there were 120 trials left, half singular trials and half plural trials. We hypothesized that if the use of *singular* as the default number feature was an artifact induced by the unequal proportion of singular and plural trials in Experiment 1, then the Gender Class × Number interaction as obtained in Experiment 1 should disappear in Experiment 3. If, however, the use of *singular* as the default number feature is a genuine characteristic of determiner retrieval, the interaction should be replicated in Experiment 3. If the interaction was obtained, it could take two different forms. First, it might take the same form as in Experiment 1, that is, as a plural cost for noun phrases with masculine and neuter nouns, and as a significantly reduced plural cost (or none at all) for noun phrases with feminine nouns. Second, the equal proportion of singular and plural trials (as opposed to a predominance of singular trials in the preceding experiments) might speed up the production of plural noun phrases relative to singular noun phrases for all gender classes. In this case, there should still be a Gender Class × Number interaction. However, the net plural costs for masculine and neuter should be somewhat reduced relative to Experiment 1, and for feminine trials there could even be some tendency toward a plural gain.

Experiment 3

Experiment 3 was a replication of Experiment 1. However, rather than controlling the relative proportion of the different determiners, we controlled for the relative proportion of singular and plural trials.

Method

Participants. Sixteen students from the University of Leipzig took part in the experiment.

Materials. The materials were identical to those used in the critical trials of Experiment 1. The 120 filler trials of Experiment 1 were not used in the present experiment.

Design. The design was identical to the one used in Experiment 1.

Procedure. The procedure was the same as the one used in Experiment 1.

Results and Discussion

The raw data were treated the same as in Experiment 1. According to these criteria, 120 observations (6.3%) were marked as erroneous and 47 observations (2.5%) as outliers.

Table 3 displays mean reaction times and error rates broken down by number and gender class. In contrast to Experiment 1, the effect of number on reaction times did not reach significance (both Fs < 1), suggesting that in the present experiment there was indeed a general reduction of overall plural costs. However, the plural condition still yielded a significantly higher error rate than did the singular condition, $F_1(1, 15) = 12.27$, $MSE = 0.49$; $F_2(1, 57) = 4.85$, $MSE = 0.99$. This reduced impact of the number effect is presumably due to the equal distribution of singular and plural trials, the (necessarily correlated) dominance of responses with the determiner *die,* or both these variables. The effect of gender class on reaction times was marginally significant in the participant analysis, $F_1(2, 30) = 3.07$, $MSE = 784.04$, $p = .06$, and not significant in the item analysis, $F_2(2, 57) = 1.09$, $MSE = 361,381$, $p = .34$. Gender did not have an effect on error rates. Most important, as in Experiment 1, the Gender Class × Number interaction was significant in the analyses of reaction times, $F_1(2, 30) = 5.25$, $MSE = 569.21$; $F_2(2, 57) = 3.75$, $MSE = 920.91$ (for error rates, both Fs < 1). T tests were computed to further analyze this interaction. For feminine nouns the 18-ms decrease in the plural condition was reliable, $t_1(15) = 2.22$; $t_2(19) = 2.87$. For neuter nouns the 20-ms increase in the plural condition was marginally significant in the participants analysis, $t_1(15) = 1.95$, $p = .07$; $t_2(19) = 1.41$, $p = .17$. None of the other effects were reliable.

Thus, keeping the proportion of singular and plural trials equal (and thus having a dominance of responses with the determiner *die*) led to a reduction of the overall plural costs. Nevertheless, the critical Gender Class × Number interaction again was significant and took a form in support of the singular-as-default hypothesis.

A comparison of Tables 1 and 3 confirms this picture. Reaction times and error rates for singular noun phrases were highly comparable in Experiments 1 and 3. By contrast, the reaction-time difference between the singular and the plural conditions in Experiment 3 was smaller than in Experiment 1, and this reduction was almost identical for all three gender classes (15-, 12-, and 14-ms reduction for nouns of masculine, feminine, and neuter gender, respectively).

General Discussion

The investigation of the production of noun phrases with gender-marked elements such as definite determiners provides a means for studying the way in which lexically specified syntactic features like grammatical gender are used in the course of language production. As discussed in the introduction, Schriefers (1993, see also Levelt et al., 1999) proposed that in producing a noun phrase with a gender-marked determiner, speakers have to retrieve an abstract feature specifying the noun's grammatical gender. This process is assumed to be a competitive process. Miozzo and Caramazza (1999, see also Caramazza et al., 2001), however, suggested that the competition occurs at the level of the actual determiners and not at the level of some abstract gender representation. In the present article, we put different versions of this determiner-selection hypothesis to an experimental test.

In our experiments, we used a simple picture-naming task. Participants named pictures of one or two target objects by either singular or plural noun phrases. Unlike in the picture–word task used in previous studies on gender and determiner processing, there should be no competition between different abstract gender features in this task. The target noun's lemma makes available its grammatical gender, and because no additional stimulus also activating gender-related information is present (such as a gender-congruent or gender-incongruent distractor word, as in Schriefers, 1993), there should be no competition in the process of retrieving the abstract gender representation.

However, if determiner selection is a competitive process, as proposed by Caramazza and colleagues (Caramazza et al., 2001; Miozzo & Caramazza, 1999), one should expect effects from competition between different determiners even in our pure production task. We specified three different hypotheses with respect to the question of how gender and number should affect naming latencies for the case that different determiners compete for selection. The results of Experiments 1 and 3 (noun phrases with definite determiners) showed a Gender Class × Number interaction, whereas this interaction was absent in Experiment 2 (bare nouns). This interaction has meanwhile been replicated in German and Dutch. For German, we replicated the Gender Class × Num-

Table 3

Mean Reaction Times (in ms) and Error Rates (in Percentages) by Number and Gender for Experiment 3

| | Gender | | | | | | | | | | | |
| | Masculine | | | | Feminine | | | | Neuter | | | |
Number	M	SE	%	SE	M	SE	%	SE	M	SE	%	SE
Singular	656	10.8	5.9	0.9	652	13.1	4.4	1.1	646	13.4	4.7	0.8
Plural	661	13.2	8.4	1.6	634	15.2	5.6	1.4	666	16.9	8.4	1.7
Difference[a]	5	8.0	2.5	1.4	−18	8.3	1.2	1.7	20	10.2	3.7	2.0

[a] Difference = plural minus singular.

ber interaction in an additional experiment in which the production of singular or plural noun phrases was cued with a delay of 100 ms relative to picture presentation. For Dutch, a Gender Class × Number interaction has been obtained by Janssen and Caramazza (2002, see also Caramazza et al., 2001). Dutch distinguishes nouns of common gender (requiring the determiner *de* in singular and plural noun phrases) from nouns of neuter gender (requiring the determiner *het* in singular noun phrases and *de* in plural noun phrases). Janssen and Caramazza found a plural cost for neuter gender nouns and a plural gain for common gender nouns, resulting in a reliable Gender Class × Number interaction. In their experiment, singular and plural utterances were induced by presenting a single object or two identical objects, respectively. These results show that our pattern of results also generalizes to a different language with a similar determiner system as German, and that this pattern does not depend on specific details of eliciting singular and plural utterances.

The overall pattern of results (see Figure 1 for a summary) is in line with the predictions of one of these hypotheses, namely the singular-as-default hypothesis.

According to this hypothesis, a noun will always automatically activate its gender-marked singular determiner. When the target utterance is a plural noun phrase, the conceptually specified feature plural will in addition activate the plural determiner. Because German nouns of masculine and neuter gender have different singular and plural determiners, these determiners compete if a plural noun phrase is to be produced. By contrast, because the singular and plural determiners for feminine nouns are identical, no determiner competition should be observable for plural noun phrases with feminine nouns. Figure 2 gives an outline of how the singular-as-default hypothesis can be captured in recent psycholinguistic models of the mental lexicon (e.g., Jescheniak & Levelt, 1994; Levelt et al., 1999).

One should note that the model representation in Figure 2 incorporates two assumptions. First, definite determiners are rep-

resented at the lemma level as well as at the level of phonological form. Second, at the lemma level but not at the word-form level, there are two distinct representations for the determiner *die*, one for its function as the singular feminine determiner, and one for its function as the plural determiner for all three gender classes. In this scenario, the observed interaction between gender and number is due to the fact that for feminine nouns the singular and plural determiner lemmas converge on one common phonological form, whereas this is not the case for masculine and neuter nouns. Thus, for plural noun phrases of all three gender classes there would be determiner competition at the lemma level. At the phonological level there should also be competition for masculine and neuter nouns but not for feminine nouns.

Of course, there are other scenarios conceivable. First, one could introduce the assumption that the determiners for singular feminine and for the plural of all three gender classes have one shared lemma representation and that this lemma is connected to one corresponding phonological form. In this case, the Gender Class × Number interaction would be due to competition at the lemma level and not at the level of phonological form. For plural noun phrases with masculine or neuter nouns, two determiner lemmas would compete for selection, whereas such competition would not obtain for plural noun phrases with feminine nouns. Although our data do not exclude this possibility, it would be in conflict with the definition and function of the lemma level. The lemma level is assumed to represent syntactic information independent of its eventual phonological realization (e.g., Levelt et al., 1999). Therefore, it appears to be odd to stipulate that two lexical elements with clearly different syntactic function are represented by a single lemma. This problem becomes even more evident for a language like German in which determiners carry information not only about number and gender but also about case. For example, the determiner *die* also functions as the accusative singular determiner for nouns of feminine gender. Furthermore, such a move would imply that the lemma level representation is informed about pho-
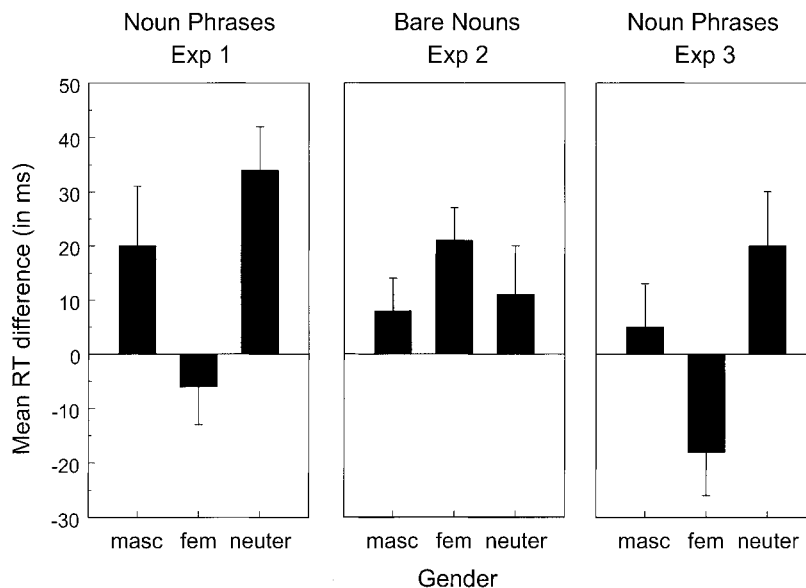

Figure 1. Mean reaction time (RT) differences (plural utterance minus singular utterance) and standard errors (in milliseconds) by gender for Experiments 1–3. Exp = experiment; masc = masculine; fem = feminine.
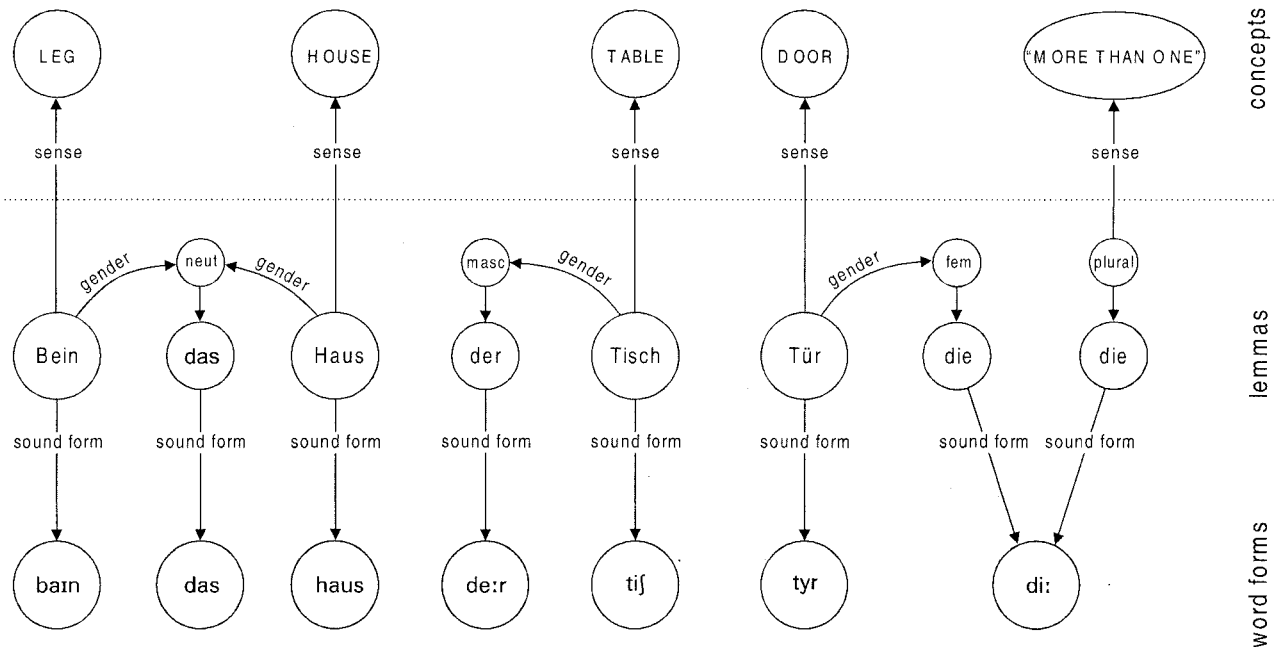
Figure 2. Simplified model representation of a fragment from the German mental lexicon. masc = masculine; fem = feminine; neut = neuter.

nological properties, an assumption that is in conflict with its original motivation.

Second, one could introduce the assumption that determiners are not represented at the lemma level at all. Rather, the diacritic plural feature or the gender feature could be linked directly to the determiner's phonological representation. However, each abstract gender node connects not only to definite determiners but also to a number of other gender-marked elements, including pronouns (e.g., Jescheniak et al., 2001; Schmitt, Meyer, & Levelt, 1999) as well, and this fact argues against linking syntactic features directly to the determiner's phonological representation without an intervening lemma representation. In line with this argument, Alario and Caramazza (2002) have recently provided evidence from French that also strongly suggests that direct activation of determiners from individual syntactic features does not provide a complete account of determiner selection.

In conclusion, the present data demonstrate that in noun phrase production singular and plural determiners compete for selection. This competition pattern was shown to be asymmetric such that singular determiners are coactivated during the production of plural noun phrases, whereas plural determiners are not substantially activated during the production of singular noun phrases. This supports the hypothesis that *singular* acts as the default value of the number feature of nouns, in line with a corresponding proposal made in the context of number agreement between subject and verb (see Bock, 1995). Furthermore, given the pattern of the competition effects, there must be some shared lexical representation for determiners with the same form (i.e., the singular definite determiner for the gender class feminine and the plural definite determiner for all gender classes). There are strong theoretical reasons to locate this overlap in representation at the level of word forms in line with previous empirical findings and sug-

gestions (Cutting & Ferreira, 1999; Dell, 1990; Jescheniak & Levelt, 1994). This, in turn, fits well with the assumption that determiner competition occurs at the level of word forms, an assumption that also allows one to preserve the view that the lemma level is an abstract level of syntactic representation that is independent of the phonological form of the eventual utterance. Finally, our observation opens up the possibility that at least some part of the gender-congruency effect obtained in the picture–word task is to be explained in terms of determiner competition as opposed to competition between abstract gender representations.

References

Alario, F.-X., & Caramazza, A. (2002). The production of determiners: Evidence from French. *Cognition, 82,* 179–223.

Bock, J. K. (1995). Producing agreement. *Current Directions in Psychological Science, 4,* 56–61.

Caramazza, A., Miozzo, M., Costa, A., Schiller, N. O., & Alario, F.-X. (2001). A cross-linguistic investigation of determiner production. In E. Dupoux (Ed.), *Language, brain, and cognitive development: Essays in honor of Jacques Mehler* (pp. 209–226). Cambridge, MA: MIT Press.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12,* 335–359.

Cutting, C. J., & Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 318–344.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes, 5,* 313–349.

Janssen, N., & Caramazza, A. (2002). *The selection of closed-class words in noun phrase production: The case of Dutch determiners.* Manuscript submitted for publication.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonolog-

ical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 824–843.

Jescheniak, J. D., Schriefers, H., & Hantsch, A. (2001). Semantic and phonological activation in noun and pronoun production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1058–1078.

La Heij, W., Mak, P., Sander, J., & Willeboordse, E. (1998). The gender-congruency effect in picture–word tasks. *Psychological Research, 61,* 209–219.

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

Miozzo, M., & Caramazza, A. (1999). The selection of determiners in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 907–922.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42,* 107–142.

Schiller, N. O., & Caramazza, A. (2002). *Gender or determiner selection interference? Evidence from noun phrase production in German and Dutch.* Manuscript submitted for publication.

Schmitt, B., Meyer, A. S., & Levelt, W. J. M. (1999). Lexical access in the production of pronouns. *Cognition, 69,* 313–335.

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). Grammatical gender in speech production: A review. *Journal of Psycholinguistic Research, 28,* 575–600.

Schriefers, H., & Teruel, E. (2000). Grammatical gender in noun phrase production: The gender interference effect in German. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1368–1377.

van Berkum, J. J. A. (1997). Syntactic processes in speech production. The retrieval of grammatical gender. *Cognition, 64,* 115–152.

Appendix

Names of Experimental Objects Used in Experiments 1–3 (With Approximate English Translations)

Masculine	Feminine	Neuter
Affe (monkey)	Birne (pear)	Auto (car)
Frosch (frog)	Boje (buoy)	Bein (leg)
Fuß (foot)	Bombe (bomb)	Bett (bed)
Hammer (hammer)	Brille (glasses)	Blatt (leaf)
Helm (helmet)	Eule (owl)	Buch (book)
Hund (dog)	Flasche (bottle)	Faß (barrel)
Kaktus (cactus)	Gabel (fork)	Glas (glass)
Kamm (comb)	Gans (goose)	Haus (house)
Knopf (button)	Hose (trousers)	Kamel (camel)
Koffer (suitcase)	Kerze (candle)	Kissen (pillow)
Korb (basket)	Leiter (ladder)	Klavier (piano)
Löffel (spoon)	Nase (nose)	Lasso (lasso)
Magnet (magnet)	Palme (palm tree)	Messer (knife)
Ofen (stove)	Pfeife (pipe)	Nest (nest)
Rock (skirt)	Schlange (snake)	Pferd (horse)
Schrank (closet)	Sichel (sickle)	Regal (shelf)
Schuh (shoe)	Sonne (sun)	Schaf (sheep)
Stuhl (chair)	Tür (door)	Schiff (ship)
Teller (plate)	Tulpe (tulip)	Zebra (zebra)
Tisch (table)	Vase (vase)	Zelt (tent)