Running Head: DETERMINER AND INFLECTION SELECTION

The selection of determiners and inflections during language production: A task comparison

Niels Janssen[1], Niels O. Schiller[2], and F.-Xavier Alario[3]

[1]University of La Laguna, Spain

[2]Leiden Institute for Brain and Cognition, Leiden University, The Netherlands

[3]Aix-Marseille Université & CNRS, France

Address for correspondence:
Niels Janssen
Grupo Neurociencia Cognitiva & Psicolingüística
Dpto. de Psicología Cognitiva, Social y Organizacional
Facultad de Psicología - Universidad de La Laguna
Campus de Guajara s/n - La Laguna, 38205 - Tenerife, España
Tel. (+34) 922 31 65 02 (+6104)
fax (+34) 922 31 91 38
njanssen@ull.es

Abstract

The research presented here addressed the question whether determiner and inflection retrieval rely on shared or distinct selection mechanisms. In three experiments, determiner and inflection retrieval was assessed by having Dutch participants produce gender-marked determiner NPs, gender-marked inflection NPs, as well as bare nouns. Performance was tested in two tasks: picture-word interference (Experiment 1) and simple picture naming (Experiments 2 and 3). In Experiment 1, a gender congruency effect was found for determiners, but not for inflections, nor for bare nouns. Likewise, in Experiments 2 and 3, a gender by number interaction was found for determiners, but not for inflections, or bare nouns. These results suggest that different mechanisms underlie the production of determiners and inflections. A further discussion of the determiner results clarifies the constraints that the observed pattern places on theories of lexical selection.

Keywords: noun phrase production; grammatical gender; determiners; inflections; picture-word interference; simple picture naming

The production of noun phrases (NPs) requires the retrieval of a wide variety of linguistic forms. For example, producing the Dutch NP 'de grote auto' [the big car] requires retrieval of morphologically simple word forms like the noun 'auto' and the determiner 'de', and morphologically complex word forms like the inflected adjective 'grot.e'. A central issue in language production is the nature of the mechanism by which these linguistic forms are selected. Much of the research on this topic has considered the production of linguistic forms such as nouns and, to a lesser extent, verbs (e.g., Levelt, 1989). However, a complete account of the lexical selection mechanism must describe the production of other types of linguistic elements as well. Here we focused on the retrieval of determiners and inflections. Based on previous research, we focused on the question whether determiners and inflections are retrieved through the same or distinct selection mechanisms (e.g., Lemhöfer, Schriefers, & Jescheniak, 2006; Schiller & Costa, 2006).

As we will argue below, there is an empirical discrepancy in the data that has been used to support various theories of determiner and inflection retrieval. Our aim in this paper is to clarify these data in two ways: First, to establish a clear empirical pattern upon which theories of determiner and inflection retrieval can be based, and second, clarify how the available data can be brought to bear on a specific mechanism of determiner selection.

General aspects of determiner and inflection retrieval

The question whether determiners and inflections share the same selection mechanism is rooted in two fundamentally different views on the role of morphology in language production. These two views differ on the assumption whether determiners and inflections have representations in the mental lexicon that are of the

same, or different types. A common assumption is that determiners and inflections have representations that are of the same type (e.g., Dell, 1986; Levelt, 1989). According to this view, the representation for a Dutch determiner like '*de*' does not differ from that of a gender-marked adjectival inflection '.*e*' (i.e., they are both lexemes). If this view were correct, one may expect determiners and inflections to share the same selection mechanism. An alternative proposal is that there are fundamental differences in the representations for determiners and inflections (e.g., Anderson, 1992; Aronoff, 1976). According to this view, there are representations in the mental lexicon for determiners, but not for inflections. Inflections are represented by a transformation that changes grammatical features into phonology. If this view were correct, and without further hypotheses made, one would not expect determiners and inflections to share the same selection mechanism[1].

Empirical observations regarding the selection of determiners and inflections have come from a variety of domains. For example, analyses of speech error corpora often suggest that determiners and inflections do not participate in speech errors (e.g., Bock & Levelt, 1994; Garrett, 1982; Stemberger, 1985). This has been taken to suggest that the speech production system treats determiners and inflections in the same way. By contrast, in studies with aphasic patients, both associations and dissociations have been observed between errors involving determiners and those that involve inflections (e.g., Miceli, Silveri, Romani, & Caramazza, 1989). This could suggest that the speech production system treats determiners and inflections in different ways.

More recently, the hypotheses that determiners and inflections are selected through the same *vs.* distinct mechanisms have been investigated with chronometric

[1] These two views are commonly referred to as the "item-and-arrangement" and the "item-and-process" theories of morphology.

tasks eliciting non-erroneous speech. This research has made extensive use of a property present in many languages, whereby the selection of determiners and inflections is constrained by grammatical information (e.g., the syntactic features *number*, or *grammatical gender*). Consider, for example, a series of representative examples from Dutch, the language used in the current study (Table 1). The form of the determiner and the adjectival inflection both depend on the gender feature of the corresponding noun (*common* or *neuter*) in singular, but they do not depend on gender in plural NPs. Note also that the adjective form does not depend on gender in definite NPs (examples 1 and 3). These properties have been used to build contrastive experimental conditions in two different tasks that are of interest in the current study: picture-word interference (PWI) and simple picture naming (SPN).

TABLE 1 HERE

The Gender Congruency effect in the PWI task

In the PWI task, participants name pictures with superimposed distractor words. In the experiments that are relevant to our purposes, the grammatical gender of the distractor word and that of the target picture name were manipulated independently. When the distractor word and the target picture have the same gender, the pair is gender congruent; when they are different, it is gender incongruent. With this manipulation, a *gender congruency effect* has been repeatedly observed in Germanic and Slavic languages: picture naming latencies are faster in the gender congruent condition compared to the gender incongruent condition (note that the

effect is not observed in Romance languages; for discussion of the cross-linguistic contrast, see Caramazza, Miozzo, Costa, Schiller, & Alario, 2001).

The gender-congruency effect was originally observed when the picture naming task required the production of gender-marked determiners (example 1 in Table 1), as well as when it required the production of gender-marked inflections (example 2 in Table 1; Schriefers, 1993). However, subsequent empirical investigations have made clear that only NPs with determiners (and not inflections) show the effect (Schiller & Caramazza, 2003; Costa, Kovacic, Fedorenko, & Caramazza, 2003; Schiller & Costa, 2006; but see the recent study by Bordag & Pechmann, 2008 discussed below) The contrasting effects for these two utterance formats have been interpreted in terms of different mechanisms for the selection of determiners and inflections (e.g., Schiller & Caramazza, 2003). We postpone the discussion of the details of these mechanisms (e.g., competitive or facilitative) to the General Discussion.

The Gender by Number interaction in the SPN task

In the SPN task, the experimental conditions are defined by combinations of the gender and number features that drive the selection of determiners and inflections, without the need of including distractor words. The rationale hinges on the fact that in Dutch and German, the plural form of the determiner does not vary as a function of gender, and is identical to one of the singular forms (e.g., in Dutch, the common gender form; Table 1 and Figure 1; in German, the feminine gender form). In this context, consider, for example, the processing hypothesis that both gender and number features contribute independently to the activation of determiner (or inflection) forms. Three different scenarios could arise. First, a possible scenario is

that there is a *cost* in the production of plural forms of neuter gender NPs, due to activation of a competing determiner form, compared to plural forms of common gender NPs (Figure 1, left panel). Accordingly, all other things being equal, the production of neuter plural NPs can be expected to be slower than the production of common plural NPs. By contrast, such a difference should not occur when singular NPs are produced (Figure 1, right panel). In other words, under this processing hypothesis, a *gender by number interaction* is predicted during NP production.

An alternative scenario is that there is a *benefit* in the production of plural forms of common gender NPs, due to the joint activation of the determiner form by its gender and number features, compared to plural neuter gender NPs (Figure 1, left panel). Such a difference should again not be observed in singular NPs (Figure 1, right panel). Therefore this scenario also predicts a gender by number interaction – albeit of a different shape than the one outlined above. The particular shape of the interaction can be used to distinguish between a competitive and facilitative mechanism of lexical selection. Again, we postpone the discussion of the details of the mechanism of lexical selection to the General Discussion.

Finally, the gender by number interaction involves a comparison between words from different genders and different numbers. Consequently, it is possible that an obtained gender by number interaction arises due to aspects related to the processing of the noun rather than the determiner (or inflection). Under such circumstances the gender by number interaction would be theoretically irrelevant for theories of determiner (or inflection) selection. To verify this point, a bare noun production experiment can be run with the same materials. The gender by number interaction should then be compared across the utterance formats that require production of determiners, inflections, and bare nouns, and will only be of theoretical

interest if it varies as a function of utterance format. Critically, the interaction pattern of gender by number in bare noun naming should be different from the one found in determiner and inflection NPs.

FIGURE 1 HERE

Using this logic, Lemhöfer et al. (2006; see Schriefers, Jescheniak, & Hantsch, 2005 for German) asked Dutch participants to name pictures with common and neuter gender names in singular and plural NPs. In these studies, gender-marked determiners were elicited by determiner+adjective+noun NPs (example 1 in Table 1), and gender-marked inflections were elicited by adjective+noun NPs (example 2 in Table 1). Also included was a bare noun control condition. The results focused on a comparison of the gender by number interaction across the three utterance formats. This was examined in two different analyses. In the first analysis, utterance format included the determiner, inflection, and bare noun formats; in this analysis, the gender by number interaction was modulated by the utterance format variable. In the second analysis, utterance format only included the determiner and inflection formats; in this analysis, the gender by number interaction was not modulated by the utterance format variable (we will consider the numerical details of these results in the discussion of Experiment 2).

These results were taken as an indication that the same gender by number interaction was present in the production of determiners and inflections, while the gender by number interaction was different in bare noun utterances. The comparable

results between the determiners and inflection utterance formats led Lemhöfer et al. (2006, see also Schriefers et al. 2005) to argue that the same mechanism underlies the production of determiners and inflections. In particular, Lemhöfer et al. and Schriefers et al. have argued for a mechanism in which determiner and inflection forms are selected in a competitive manner (e.g., Roelofs, 2003).

In short, then, the picture naming evidence published so far points to an empirical discrepancy. Although initially the findings from the PWI task suggested similar patterns for determiners and inflections, a thorough investigation revealed contrastive results in the production of determiners and inflections. This could mean that they are retrieved through different selection mechanisms. By contrast, in the SPN task, comparable results have been found for the two utterance formats. This in turn has prompted the hypothesis that the same mechanism is involved in both types of forms. In order to account for this apparent conflict, Lemhöfer et al. (2006) proposed that the PWI and SPN tasks differ in terms of their sensitivity to detect the effects of inflection retrieval. More specifically, this post-hoc assumption states that the effects of inflection retrieval are unreliably detected by the PWI task, but are reliably detected by the SPN task. Under this assumption the contrastive patterns of results between the PWI and SPN tasks are no longer theoretically relevant, and the available data favor the hypothesis that the same mechanism underlies determiner and inflection production.

Before this conclusion can be accepted, however, some important considerations have to be taken into account. First, the differential sensitivity hypothesis has been proposed post-hoc and still awaits an explicit empirical test. Secondly, the different patterns of results in the PWI and SPN tasks were obtained with different sets of materials and different participant groups. Hence, there is a

possibility that the differences between the two patterns are not due to differences in task sensitivity, but simply reflect differences between the experimental materials and participant groups involved. Finally, as reviewed above, research using the PWI task initially suggested similar patterns for determiners and inflections, an observation that was not replicated in subsequent studies. It is therefore important that the reliability of the patterns observed in the SPN task is assessed just as well.

We report three experiments aimed at clarifying the similarity and differences between the selection processes involved in determiner and inflection retrieval. In Experiment 1, participants performed a PWI task with a gender manipulation along the lines described above. In Experiment 2, the same participants were confronted with the same pictures in a SPN task. The materials and design were similar to those used by Lemhöfer et al. (2006). Finally, Experiment 3 tested the same task and conditions as Experiment 2, intended to ascertain the reliability of the results we report.

Experiment 1 – Picture-Word Interference

In this experiment, three naming formats were compared: determiner+adjective+noun, adjective+noun, and bare noun naming conditions. The distractor words were either gender congruent or incongruent with the picture names. On the basis of previous findings, we expected a congruency effect in determiner NPs, but no effect in adjective NPs or bare noun naming.

Methods

Participants

Sixty native Dutch speakers, all students at the Radboud University Nijmegen in The Netherlands, took part in Experiment 1. These sixty participants were divided into three groups of 20 participants and were assigned to the determiner+adjective+noun, adjective+noun, and bare noun naming conditions. The same group of participants was assigned to the same utterance format condition in Experiment 2. Each participant was paid three euro upon completion of the experiment.

Materials and Design

The 48 target pictures were identical to the set of pictures with common and neuter gender names used in Lemhöfer et al. (2006; see Appendix B). Because in the SPN task employed by Lemhöfer et al. there were no distractor words, two sets of distractor words were selected. All distractor words were concrete words referring to everyday objects. Half had a neuter gender name, and half had a common gender name. Neuter and common gender words were paired with neuter and common gender picture names to form the Gender Congruent condition. These distractor words were then repaired with pictures of the other gender class to form the Gender Incongruent condition. Thus, the same distractor words appeared in the Gender Congruent and Incongruent conditions. Care was taken that distractor word and picture name pairs did not have any semantic or phonological overlap. The adjectives in the determiner+adjective+noun and in the adjective+noun NPs were 'klein/kleine' (small)

and 'groot/grote' (big) and were elicited by varying the size of the pictures (see Lemhöfer et al. for further details).

The counterbalancing of the factors Gender, Gender Congruency, and Size was done as described in the paper by Lemhöfer et al. (2006). This led to the creation of sixteen individual participant lists, with 96 experimental items each. Finally, six additional pictures were selected for practicing the experimental tasks and for use as warm-up trials. From these six pictures 12 practice trials were created.

Procedure

The experimental software was NESU (Nijmegen Experimental Set Up). Participants sat in front of a PC in a sound-proof booth. They wore headphones and spoke into a microphone. The microphone was connected to a button-box and provided voice-key measurements with 1 ms accuracy.

The experiment was preceded by a picture familiarization phase, in which participants named the 48 experimental and 6 practice pictures. On each trial, a fixation point appeared for 700 ms. Next, a picture appeared and after 1 second the name of the picture appeared beneath the picture. The appearance of the picture's name was a cue for the participant to produce the name of the picture aloud. The picture and its name remained on the screen for 1 second. After a blank screen of 1,500 ms, the next trial started. Next, the participants practiced the experimental task. They were instructed to name the pictures while ignoring the presented distractor words, using singular NPs with determiner+adjective+noun, adjective+noun or bare nouns, depending on the naming condition. On each trial, a fixation cross appeared for 700 ms. Next, the picture display appeared for 1,500 ms or until participants made a

vocal response. Finally, there was a pause of 2,000 ms before the next trial started. After the practice phase, participants performed the experiment proper which was identical in trial structure to the practice phase. The experiment lasted about 20 minutes.

Results

For this analysis, and all other analyses reported herein, trials on which a participant's response deviated from the response intended by the experimenter were discarded from the analysis. In addition, trials on which reaction times (RT) were more than 2.5 standard deviations above the overall participant's and item's mean, as well as those RTs smaller than 300 and larger than 2,000 ms were discarded from the analyses.

The variables Utterance Format (between participants, within-items), and Gender Congruency (within-participants, within-items) were analyzed with participants (F1) and items (F2) as random factors.

TABLE 2 & 3 HERE

The RT analysis revealed a marginally significant interaction between Utterance Format and Gender Congruency ($F1(2, 57) = 5.67$, $p < .006$; $F2(2, 94) = 3.00$, $p < .06$), suggesting that the Gender Congruency effect depended on the Utterance Format. This was confirmed in a further analysis which revealed an effect

of Gender Congruency in the determiner+adjective+noun, but not in the adjective+noun or bare noun utterances (see Tables 2 and 3). The results from error analyses were consistent the RT results (see Appendix for details of the error analyses).

Discussion

The results of Experiment 1 revealed a gender congruency effect in the production of NPs that require retrieval of a gender-marked determiner, but not in the production of NPs that require retrieval of a gender-marked inflection.

While these results are consistent with those obtained in previous studies (Costa et al., 2003; Schiller & Caramazza 2003; Schiller & Costa, 2006), there are also studies that have reported results that stand in contrast with those reported here (Bordag & Pechmann, 2008; Schriefers, 1993). For example, in a recent study, Bordag and Pechmann asked Czech participants to name pictures in the context of distractor words that were gender congruent or gender incongruent. Pictures were named using NPs with gender marked demonstrative determiners (e.g., ta_{fem} auto [this car]), or with an inflected gender marked numeral (pat-a_{fem} auto [fifth car]). The results revealed gender congruency effects for both determiners and inflections.

The results presented here do not allow us to determine why some studies have reported the effect and others have not. Nevertheless, one may speculate about the causes of this discrepancy. First, a survey of the literature concerning inflection production reveals that the gender congruency effect is not observed in 6 experiments (Exp 1 reported here; Exp 2 and 3 in Costa, Kovacic, Fedorenko, & Caramazza, 2003; Exp 1b and 4a in Schiller & Caramazza, 2003; Exp 1a in Schiller & Costa, 2006), and

is observed in 3 (Exp 2 and 3 in Bordag & Pechmann, 2008; Exp 2 in Schriefers, 1993). Consequently, it cannot be ruled out that the positive gender congruency effect simply reflects a type 1 error. Furthermore, a hitherto unaddressed issue concerns the possibility that the gender congruency effect in inflection production depends on particular properties of the carrier word, where inflection production only yields a gender congruency effect if the carrier has certain properties. This is a plausible hypothesis, since distinct types of carrier words have been used across studies (e.g., definite and indefinite articles, demonstratives, numerals, etc). If such a hypothesis were verified, then the gender congruency effect should no longer be considered a reflection of inflection retrieval per se; Inflection production should be studied with that broader perspective in mind. Future research should be geared towards resolving this issue.

Experiment 2 – Simple Picture Naming

In Experiment 2, the participants of Experiment 1 performed the SPN task. The order of these two tasks was counterbalanced across participants. We expected to find a gender by number interaction in the production of determiner+adjective+noun and adjective+noun NPs, but not in bare noun naming.

Methods

Participants

Participants in Experiment 2 were the same as those in Experiment 1. They remained assigned to the same Utterance Format condition as in Experiment 1.

Materials and Design

The target pictures were identical to those used in Experiment 1 and in Lemhöfer et al. (2006; see Appendix B). The design was identical to that used by Lemhöfer et al. (2006), except that singular NPs were elicited by the presentation of a single picture, and plural NPs were elicited by the presentation of two identical pictures (as in Janssen & Caramazza, 2003).

Procedure

Given that the same target pictures were used in Experiments 1 and 2, and that the same participants did both Experiments 1 and 2, participants performed a familiarization stage with the target pictures only if they had not already done so. The trial structures for the practice and experiment proper were also identical to that used in Experiment 1.

Results

The variables Utterance Format (between participants, within-items), Gender (within-participants, between-items), and Number (within-participants, within-items) were analyzed with participants (F1) and items (F2) as random factors. Main effects of Utterance Format are not theoretically relevant and thus not reported.

TABLE 4 & 5 HERE

The RT analysis revealed a marginally significant three-way interaction between Utterance Format, Gender, and Number ($F1(2, 57) = 3.10$, $p < .06$; $F2(2, 92) = 3.43$, $p < .04$). The nature of this interaction was evaluated in two ways. First, the interactions of Gender and Number were computed separately for the three utterance formats. In this analysis, the Gender by Number interaction was significant in the determiner+adjective+noun, but not in the adjective+noun, nor in bare noun utterances (see Table 4). Further t-tests revealed that the Gender by Number interaction in determiner+adjective+noun utterances was due to comparable RTs for singular and plural NPs for common gender words, and faster RTs for singular compared to plural NPs for neuter gender words (see Table 5).

Further information about the three-way interaction can be gained by conducting a supplementary analysis in which the Gender by Number interaction is compared across two rather than three different levels of the Utterance Format variable. In this analysis, the Gender by Number interaction differed between determiner+adjective+noun and bare noun utterances ($F1(1, 38) = 5.51$, $p < .03$; $F2(1, 46) = 6.00$, $p < .02$), but not between adjective+noun and bare noun utterances (both $Fs < 1$), and was marginally significant between determiner+adjective+noun and adjective+noun utterances ($F1(1, 38) = 2.35$, $p = .13$; $F2(1, 46) = 3.00$, $p < .09$). The results from error analyses were consistent with the RT results (see Appendix for details of the error analyses).

Discussion

The results of Experiment 2 revealed that the gender by number interaction depended on the utterance format. Whereas there was a gender by number interaction in determiner+adjective+noun utterances, there was, to our surprise, no such interaction in the adjective+noun or bare noun utterances.

At face value, the results of Experiments 2 seem at odds with those reported by Schriefers et al. (2005) and Lemhöfer et al. (2006). Both studies claimed to have observed evidence indicating that determiner and inflections share the same selection mechanism. However, upon closer inspection of those results, it is not obvious that there is strong support for this claim. In Schriefers et al., the critical gender by number interaction was significant in the utterance format that required production of determiners, but was only marginally significant in the utterance format that required production of inflections (by-participants $p = .05$, by-items $p = .28$). Moreover, an analysis that compared the gender by number interaction between these two utterance formats was significant (by-participants $p = .04$[2], by-items $p = .06$). Likewise, in Lemhöfer et al., the results of that same interaction was p = ns by participants and $p = .10$ by items[3], and in the analysis reported above, the results were $p = .13$ by participants and $p = .09$ by items. Accepting the null hypothesis of no differences between the determiner and inflection conditions on the basis of these data may be premature. In short, it is not obvious that results of Lemhöfer et al., Schriefers et al, and those observed here provide strong support for the claim that determiners and inflections are produced through a shared selection mechanism.

[2] It is possible that this p-value is incorrect. In Schriefers et al. (2005, footnote 8) it is reported as $F1(2, 124) = 14.49$, $p = .04$. However, with this F value and degrees of freedom, the correct p-value is: $p < .001$.

[3] Lemhöfer et al (2006) do not report the interaction between gender and number separately for each utterance format.

In light of these considerations, we conducted Experiment 3.

Experiment 3 - Simple Picture Naming

In Experiment 3, we attempted to further establish the reliability of the pattern of results observed in Experiment 2. A new group of participants performed the SPN task in the three utterance formats with new materials. If the results obtained in Experiment 2 are reliable, we expected to find a gender by number interaction in the production of determiner+noun NPs, but not in the production of adjective+noun NPs or in bare noun naming.

Methods

Participants

Fifty participants took part in Experiment 3. There were 15 participants in the determiner+noun, twenty in the adjective+noun naming, and 15 in the bare noun condition. All participants were native speakers of Dutch and had not participated in Experiments 1 or 2. They were paid three euro upon completion of the experiment.

Materials and Design

For the determiner+noun condition, the materials and design were identical to those used in Janssen and Caramazza (2003; see Appendix C). There were 60

pictures, half with common and half with neuter gender names. These 60 pictures appeared in a singular and in a plural condition, leading to a total of 120 experimental items. Note that in these 120 experimental items there are 90 items that require determiner form ‘de’, and 30 items that require ‘het’. Following Janssen and Caramazza's design, another 30 filler pictures with neuter gender names that were named in singular were added. In total, there were 150 items in the experiment. The stimulus displays that were used to elicit singular and plural NPs were created as in Experiment 1. Finally, eight pictures were selected as practice items. Half of these pictures had common and half had neuter gender names.

The determiner+noun condition was tested earlier in time than the adjective+noun and bare noun conditions. In the analyses of the determiner+noun condition we identified ten pictures in the common gender and ten pictures in the neuter gender conditions with relatively low name agreement and high error rates. These pictures were not included in the adjective+noun and bare noun conditions (see Appendix C for details). The analyses reported here were conducted on the materials that were shared between the three naming conditions (for further considerations see footnote 5 in the results section). In the adjective+noun and bare noun conditions there were forty pictures with half common and half neuter gender names. As in the determiner+noun condition, twenty additional filler pictures with neuter gender names were added to counterbalance the number of trials on which a neuter or common gender inflection was required. Stimulus displays to elicit singular and plural NPs were created as in Experiment 1.

In contrast to the size adjectives in Experiments 1 and 2, adjectives in Experiment 3 were colors, elicited by presenting the outline of the pictured objects in one of four colors. Color adjectives follow identical rules for gender inflection than

size adjectives. The colors all had one-syllable names (i.e., ‘rood’ [red], ‘groen’ [green], ‘blauw’ [blue], and ‘paars’ [purple]. RGB values were: red [247,0,0]; green [0,206,0]; blue [0,0,247]; purple [222,16,214]. These colors were paired with the pictures on the basis of the following three constraints. Pictures did not appear in all colors, but if a particular picture would appear in a certain color in the singular condition, that picture would also appear in that same color in the plural condition. Second, phonological overlap between the color name and the picture name should be minimal. Finally, each color appeared equally often in each condition of the experiment. Practice pictures were the same as in the determiner+noun condition.

The materials of the bare noun condition were identical to those used in the adjective+noun condition.

The items were pseudo-randomized into a stimulus list with the following four constraints: (1) there were never more than three consecutive trials requiring the same determiner; (2) there were never more than three consecutive trials with an identical number; (3) there were no consecutive trials that shared phonological onset; (4) there were no consecutive trials on which there was a semantic relationship between picture names.

Procedure

The procedure was similar to the one used in Experiment 2. The experimental software was Psyscope (Cohen, MacWhinney, & Flatt, 1993). In the determiner+noun condition, participants were told to use the gender-marked distal demonstrative determiner (e.g., die auto [that car]/ dat boek [that book]; die autos [those cars]/ die boeken [those books]). In the adjective+noun condition, participants were told to

produce the appropriate color name and the depicted object in singular and plural NPs. Each naming condition lasted about twenty minutes.

Results

The analyses were conducted as in Experiment 2.

TABLE 6 & 7 HERE

The three-way interaction between Utterance Format, Gender, and Number was significant ($F1(2, 47) = 12.50$, $p < .001$; $F2(2, 76) = 11.14$, $p < .001$). As in Experiment 2, this interaction was evaluated in two ways. First, separate analyses for the three utterance formats revealed a Gender by Number interaction in the determiner+noun, but not in the adjective+noun, or bare noun utterances (see Table 6). Further t-tests revealed that for determiner+noun utterances, singular NPs differed from plural NPs for common gender words, and that singular NPs also differed from plural NPs for neuter gender words, albeit in the opposite direction (see Table 7).

Second, we computed three-way interactions between Gender, Number, and utterance format restricted to two levels. The Gender by Number interaction differed between determiner+noun and bare noun utterances ($F1(1, 28) = 20.33$, $p < .001$; $F2(1, 38) = 17.00$, $p < .001$), and between determiner+noun and adjective+noun utterances ($F1(1, 33) = 15.19$, $p < .001$; $F2(1, 38) = 12.50$, $p < .002$), but not between

adjective+noun and bare noun utterances (both Fs < 1)[4]. The results from error analyses were consistent with the RT results (see Appendix for details of the error analyses).

Discussion

The results of Experiment 3 establish the reliability of the findings reported in Experiment 2. Notably, the difference between determiner and inflection utterances, which was not clearly established in previous data (Experiment 2; Lemhöfer et al., 2006; Schriefers et al., 2005), was clearly significant. Together, the available evidence allows rejecting the null hypothesis that determiner and inflection utterances show the same pattern in these experiments. In other words, in this experiment as in Experiment 2, the gender by number interaction depended on the utterance format. Whereas there was a gender by number interaction in determiner+adjective+noun NPs, there was no such interaction in the adjective+noun NPs, or in bare noun naming.

One possible reason for the difference in results between the two NP utterance formats is that determiners appear at the onset of NPs, whereas inflections appear only at the end of the first word of the NP. If articulation could start on the basis of partial encoding of the response, not including the inflection, then response times need not be sensitive to manipulations of inflection retrieval processes. However, there is ample evidence that even the last word of an NP (here the noun) is processed at the lexical and phonological levels before articulation starts (Alario & Caramazza, 2002; Costa & Caramazza, 2002; Jescheniak, Schriefers, & Hantsch, 2003; Damian & Dumay, 2007; see also discussion in Schiller & Costa, 2006). Thus, it is unlikely that

[4] When these analyses were computed for the full set of items in the determiner+noun condition the same results were found.

differences in the location of the gender-marked element in the NP can account for the different results between the two utterance formats.

General Discussion

The goal of the current research was to clarify the differences and similarities between the selection of determiners and inflections during noun phrase production. This was done by comparing performance in the PWI and the SPN tasks. In the PWI task (Experiment 1), a gender congruency effect was found in NPs that required the production of gender-marked determiners, but not when they required production of gender-marked inflections, or when bare nouns were produced. Similar results were found in the SPN task (Experiments 2 and 3). In this task, a gender by number interaction was found in NPs that required the production of determiners, but not when they required production of gender-marked inflections, or when bare nouns were produced.

These results do not support the claim that the PWI and SPN tasks differ in their sensitivity to detect effects of inflection retrieval (Lemhöfer et al., 2006). If such had been the case, we would have expected a gender congruency effect in determiner but not inflection retrieval in the PWI task of Experiment 1, and a gender by number interaction in both determiner and inflection retrieval in the SPN task of Experiments 2 and 3. However, instead of this we observed the same pattern of results in the PWI and SPN tasks for determiner and inflections.

Furthermore, the results of Experiments 2 and 3 are not inconsistent with those of Lemhöfer et al. (2006) and Schriefers et al. (2005). As we argued in the discussion section of Experiment 2, careful consideration of the numerical details of their

statistical analyses suggests that their interpretation of these data might have been incorrect. Overall then, a pattern of results emerges that is consistent across tasks and across studies: Whereas determiner retrieval is sensitive to gender and number manipulations in picture naming tasks, inflection retrieval is not. This pattern suggests that different selection mechanisms may underlie the production of these two morphological forms.

Such a conclusion may be premature, however. This is because the available evidence provides no positive information about the selection of inflections, other than the absence of effects in these tasks. Of course, if determiner and inflection retrieval were identical in every respect, such a difference would not be expected. Accordingly, we can speculate that there are at least some differences in the retrieval of these two form-types, where inflection retrieval involves a transformation of grammatical features into a phonological form (e.g., Anderson, 1992; Aronoff, 1976; see also discussion in Costa et al., 2003). Pending further evidence, however, we focus the rest of this discussion on the retrieval of determiners, for which positive evidence has been found in the picture naming tasks.

Mechanisms of determiner selection

The two systematic findings regarding determiner NP production are a gender congruency effect in the PWI task, and an interaction between gender and number in the SPN task. Here we consider whether these two effects place constraints on theories of lexical selection, for which two basic mechanisms have been contrasted. First, a competitive lexical selection mechanism assumes that target selection time is an increasing function of the activation of other activated lexical representations (e.g., Roelofs, 2003). Alternatively, in a non-competitive selection hypothesis, lexical

selection time is independent from the activation of other lexical representations, depending only on the activation process of the actual target (e.g., Mahon, Costa, Peterson, Vargas, & Caramazza, 2007).

Schriefers and colleagues (e.g., Schriefers, 1993; Lemhöfer et al., 2006) have claimed that the gender congruency effect in the PWI task reflects a competitive determiner selection mechanism. This explanation rests on the interpretation of the gender congruency effect (i.e., slower naming latencies in the context of incongruent distractors compared to congruent distractors) as an inhibitory effect. On this assumption, the effect would arise as a consequence of a delay in the selection of the target determiner form due to the activation of an incorrect, competing determiner form by the incongruent distractor word. However, the assumption that the gender congruency effect reflects an inhibitory effect is problematic, given that the gender congruency effect is not evaluated against a neutral baseline. In other words, it is equally possible that the gender congruency effect reflects a facilitatory effect, where naming latencies are faster with congruent than with incongruent distractors. Thus, the assumption that the gender congruency effect is an inhibitory effect is not tenable, and hence, it cannot be taken as supportive of a theory of competitive determiner selection.

In the SPN task, the competitive and non-competitive theories of lexical selection predict different types of gender by number interactions in the production of determiner NPs. A competitive theory of lexical selection assumes that, in plural neuter NPs, competition will arise during determiner selection because of the activation of alternative (i.e. common and neuter) determiner forms (Figure 1, left panel). Such competition will not be present in plural common NPs, nor in singular NPs irrespective of gender (Figure 1, right panel). This hypothesis therefore predicts a

gender by number interaction where there is an extra-cost in the production of plural NPs with neuter gender nouns. By contrast, according to a non-competitive theory of lexical selection, plural common determiners will benefit from activation from both gender and number features. (Figure 1, left panel). Such will not be the case for plural neuter determiners, and hence there should be faster selection of plural common determiners. More generally, relative to a singular baseline, this hypothesis predicts a gender by number interaction where there is a benefit in the production of plural NPs with common gender nouns.

In other words, a competitive theory of lexical selection predicts a gender by number interaction referred to here as *cost-type* interaction, and the non-competitive theory predicts a gender by number interaction referred to here as *benefit-type* interaction. In the following we discuss how these two different types of gender by number interactions must be evaluated in light of the effects in the bare noun baseline condition.

As we discussed in the Introduction, a gender by number interaction in determiner NPs is only meaningful if this interaction is absent in bare noun naming. In addition, a further constraint applies if a distinction is to be made between a cost-type and a benefit-type interaction. This constraint is that the interaction can only be interpreted straightforwardly in the absence of *a main effect of number* in the bare noun naming condition. If there is a baseline number effect in bare noun naming, this would require an adjustment of the differences between plural and singular conditions in determiner NPs. To illustrate the important consequences of such adjustments for number effects, compare the hypothetical *cost*- and *benefit*-type gender by number interactions depicted in the top row of Figure 2, to the gender by number interactions in the bottom row of Figure 2. In the top row, a main effect of number is absent, while

in the bottom row a main effect of number is present. As can be seen under 'difference', a positive main effect of number could reverse the cost-type interaction into a benefit-type interaction (Figure 2, bottom row, left panel), and negative main effect of Number could reverse a benefit-type interaction into a cost-type interaction (Figure 2, bottom row, right panel). The observation of a cost-type or benefit-type gender by number interaction in determiner NPs can therefore only be interpreted as such if there is no main effect of number in bare noun naming.

FIGURE 2 HERE

Careful consideration of all the studies in the literature that have used the SPN task with a bare noun baseline condition reveals that only two studies meet the requirements laid out in the previous paragraph. In Janssen and Caramazza (2003), and in Schriefers et al. (2005), a gender by number interaction is reported in determiner NPs, but not in bare nouns, and in both studies there is no main effect of number in bare noun naming. The other studies reported a main effect of Number in bare noun naming (i.e., Experiments 2 and 3 reported here; Schriefers et al., 2002), or did not report a test of this effect (i.e., Lemhöfer et al., 2006).

FIGURE 3 HERE

A graphical representation of the results of the studies of Janssen and Caramazza (2003) and Schriefers et al. (2005) is presented in the top row of Figure 3. There are two important points to be made. First, it should be clear that in these two studies, the gender by number interactions are neither exclusively of the cost, nor of the benefit-type, but rather show both a plural cost and a plural benefit. Second, the results are not consistent across the two studies. In Janssen and Caramazza, there is a plural cost for neuter gender nouns, and a plural benefit for common gender nouns. In Schriefers et al., there is a plural cost for masculine and neuter gender nouns that is only marginally significant, and a plural benefit for feminine nouns. A possible interpretation of these results is that the gender by number interaction of the cost-type is not observed consistently across studies, while the gender by number interaction of the benefit-type is reliably observed across studies. If this interpretation were correct then the results obtained in the SPN task would favor a non-competitive mechanism of determiner form selection. However, it should be noted that this interpretation is based on only two studies.

One final aspect of the results reported here that deserves further discussion is the experiment-to-experiment variability in the gender by number interaction in determiner NPs. Specifically, the gender by number interaction observed in the determiner NPs of Experiment 2 differs from the one observed by Lemhöfer et al. (2006; see also Figure 3) even though the same sets of materials were used. However, as we argued above, a gender by number interaction cannot be interpreted directly without taking into account the main effect of number in the bare noun naming condition. When the main effects in bare noun naming are taken into account (see Figure 3), the differences between various experiments become less pronounced. One

reason why this could be case is that different stimulus displays were used between the two studies and this influenced the main effect of number in the experiments.

Summarizing, in contrast to the claims made in the literature (e.g., Schriefers, 1993; Lemhöfer et al., 2006), neither the gender congruency effect in the PWI task, nor the gender by number interaction in the SPN task provide compelling support for a competitive lexical selection mechanism. If anything, the current results may suggest a non-competitive mechanism of determiner form selection, a proposal that is supported by recent studies that have compared the competitive and non-competitive mechanisms of lexical selection in the context of determiner and noun production. For example, Alario, Ayora, Costa and Melinger (2008) showed that, in a PWI task, the production of determiner NPs was sped up rather than inhibited by the presentation of distractor words that were strong competitors to the target determiner. The authors argued on the basis of these results that the mechanism of lexical selection is not competitive, a conclusion that fit well studies investigating the selection of nouns (e.g., Finkbeiner & Caramazza, 2006; Janssen, Schirm, Mahon, & Caramazza, 2008; Mahon, et al., 2007).

To conclude, the aim of the current research was 1) to clarify the empirical data with respect to determiner and inflection retrieval, and 2) to interpret these data in terms of a competitive or facilitative mechanism of determiner form selection. The picture that emerged from the research presented here is that there are effects in the production of determiners, whereas there are no such effects in the production of inflections. Importantly, this pattern of results generalizes across tasks (PWI and SPN) and across studies (Janssen & Caramazza, 2003; Lemhöfer et al., 2006; Schriefers et al., 2002; 2005). Taken together, the evidence we have reviewed and reported here suggests a system of lexical selection by which linguistic forms such as

determiners and nouns are selected by a non-competitive lexical mechanism (e.g., Mahon et al., 2007), and that the selection of inflections may well involve a phonological transformation (e.g., Anderson, 1992).

References

Alario, F.-X., Ayora, P., Costa, A., & Melinger, A. (2008). Grammatical and nongrammatical contributions to closed-class word selection. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 34*, 960-981.

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

Anderson, S. R. (1992). *A-Morphous Morphology*. Cambridge, UK: Cambridge University Press.

Aronoff, M. (1976). *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

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

Caramazza, A., Miozzo, A., Costa, A., Schiller, N. O., & Alario, F.-X. (2001). A crosslinguistic 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.

Cohen, J. D., Macwhinney, B., & Flatt, M. (1993). Psyscope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavioral Research Methods, Instruments and Computers, 25*, 257-271.

Costa, A., Alario, F. X., & Caramazza, A. (2005). On the categorical nature of the semantic interference effect in the picture-word interference paradigm. *Psychonomic Bulletin & Review, 12*, 125-131.

Costa, C., Kovacic, D., Fedorenko, E., & Caramazza, A. (2003). The gender congruency effect and the selection of freestanding and bound morphemes: Evidence from Croatian. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1270-1282.

Costa, A., & Caramazza, A. (2002). The production of noun phrases in English and Spanish: Implications for the scope of phonological encoding in speech production. *Journal of Memory and Language, 46*, 178-198.

Damian, M. F., & Dumay, N. (2007). Time pressure and phonological advance planning in spoken production. *Journal of Memory and Language, 57*, 195-209.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*, 283-321.

Garrett, M. F. (1982). Production of speech: Observations from norma and pathological language use. In A. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19-76). London: Academic Press.

Janssen, N., & Caramazza, A. (2003). The selection of closed-class words in noun phrase production: The case of Dutch determiners. *Journal of Memory and Language, 48*, 635-652.

Janssen, N., Schirm, W., Mahon, B. Z., & Caramazza, A. (2008). Semantic interference in a delayed naming task: Evidence for the response exclusion hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 249-256.

Jescheniak, J. D., Schriefers, H., & Hantsch, A. (2003). Utterance format affects phonologuical priming in the picture-word task: Implications for models of phonological encoding in speech production. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 441-454.

Lemhöfer, K., Schriefers, H., & Jescheniak, J. D. (2006). The processing of free and bound gender-marked morphemes in speech production: Evidence from Dutch. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 32*, 437-442.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA, USA: MIT Press.

Mahon, B.Z., Costa, A., Peterson, R., Vargas, & Caramazza, A. (2007). Lexical Selection is Not by Competition: A Reinterpretation of Semantic Interference and Facilitation Effects in the Picture-Word Interference Paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 503-535

Miceli, G., Silveri, M.C., Romani, C., & Caramazza, A. (1989). Variation in the pattern of omissions and substitutions of grammatical morphemes in the spontaneous speech of so-called agrammatic patients. *Brain & Language, 36*, 447-492.

Roelofs, A. (2003). Goal-Referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review, 110*, 88-125.

Schiller, N. O., & Caramazza, A. (2003). Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory and Language, 48*, 169-194.

Schiller, N. O., & Costa, A. (2006). Different selection principles of freestanding and bound morphemes in language production. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 32*, 1201-1207.

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., & Hantsch, A. (2002). Determiner selection in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 941-950.

Schriefers, H., Jescheniak, J. D., & Hantsch, A. (2005). Selection of gender-marked morphemes in speech production. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 31*, 159-168.

Stemberger, J. P. (1985). *An interactive activation model of language production.* In A. Ellis (Ed.), Progress in the Psychology of Language. London: Erlbaum.

Tables

Table 1. Examples of Dutch gender-marked determiners (examples 1 and 3) and inflections (examples 2 and 4) as a function of gender (common versus neuter) and number (singular NPs in examples 1 and 2; plural NPs in examples 3 and 4).

Example	Number	Common	Neuter
1	sg	**de** grote auto *[the big car]*	**het** grote boek *[the big book]*
2	sg	grot**e** auto *[big car]*	groot boek *[big book]*
3	pl	**de** grote autos *[the big cars]*	**de** grote boeken *[the big books]*
4	pl	grot**e** autos *[big cars]*	grot**e** boeken *[big books]*

Table 2. Overview of F and p statistics for the analyses of naming latencies in the picture-word interference task of Experiment 1.

Picture-Word Interference		Participants		Items	
		F(1,19)	*p*	F(1,46)	*p*
Determiner+adjective+noun	Congruency	12.83	< .002	5.31	< .03
Adjective+noun	Congruency	< 1	ns	< 1	ns
Bare noun	Congruency	< 1	ns	< 1	ns

Table 3. Mean naming latencies and error percentages (between brackets) as a function of Congruency (congruent versus incongruent) in the picture-word interference task in Experiment 1.

Picture-Word Interference			
	Congruent	Incongruent	*difference*
Determiner+adjective+noun	822 (13.0)	842 (14.3)	20
Adjective+noun	726 (9.1)	725 (9.2)	-1
Bare noun	745 (9.9)	747 (10.1)	2

Table 4. Overview of F and p statistics for the analyses of naming latencies in the simple picture naming task of Experiment 2.

Simple Picture Naming		Participants		Items	
		F(1,19)	*p*	F(1,46)	*p*
Determiner+adjective+noun	Gender	12.90	< .002	4.67	< .04
	Number	2.57	= .12	2.07	.15
	Gender*Number	5.00	< .04	4.81	< .04
Adjective+noun	Gender	11.50	< .004	3.60	< .07
	Number	13.00	< .002	4.63	< .04
	Gender*Number	< 1	ns	< 1	ns
Bare noun	Gender	20.52	< .001	5.21	< .03
	Number	4.53	< .05	6.51	< .01
	Gender*Number	1.22	= .28	2.28	= .14

Table 5. Mean naming latencies and error percentages (between brackets) as a function of Gender (common versus neuter) and Number (singular versus plural) in the simple picture naming task in Experiment 2.

Simple Picture Naming	Common			Neuter		
	Singular	Plural	diff	Singular	Plural	diff
Determiner+adjective+noun	739 (10)	734 (6)	-5[*]	749 (11)	774 (13)	25[¥]
Adjective+noun	705 (7)	720 (9)	15[a]	723 (9)	740 (13)	17[a]
Bare noun	665 (6)	686 (9)	21[a]	708 (7)	714 (8)	6[a]

[*] ns; [¥] at p1 < .02; p2 < .003; [a] In the absence of an interaction, see main effects computed in Table 3.

Table 6. Overview of F and p statistics for the analyses of naming latencies in the simple picture naming task of Experiment 3.

Simple Picture Naming		Participants*		Items	
		F(1,14)	*p*	F(1,38)	*p*
Determiner+noun	Gender	3.39	= .09	< 1	ns
	Number	2.27	= .15	2.69	= .11
	Gender*Number	23.67	< .001	25.05	< .001
Adjective+noun	Gender	1.58	= .22	< 1	ns
	Number	4.10	= .06	5.64	< .03
	Gender*Number	< 1	ns	< 1	ns
Bare noun	Gender	< 1	ns	< 1	ns
	Number	19.18	< .001	30.88	< .001
	Gender*Number	< 1	ns	< 1	ns

* Degrees of freedom for the by-participants F test in the adjective+noun condition are (1, 19).

Table 7. Mean naming latencies and error percentages (between brackets) as a function of Gender (common versus neuter) and Number (singular versus plural) in the simple picture naming task in Experiment 3.

Simple Picture Naming	Common			Neuter		
	Singular	Plural	diff	Singular	Plural	diff
Determiner+noun	843 (10)	783 (6)	-61[*]	778 (6)	810 (13)	33[¥]
Adjective+noun	751 (9)	765 (10)	14[a]	761 (12)	777 (16)	16[a]
Bare noun	602 (4)	635 (6)	32[a]	608 (5)	636 (11)	28[a]

[*] at p1 < .002, p2 < .001; [¥] at p1 < .006; p2 < .006; [a] In the absence of an interaction, see main effects computed in Table 3.

Figures

Figure 1. Graphical presentation of the relationship between the features grammatical gender, number and the form of the determiner in Dutch plural (left panel) and singular (right panel) NPs. Note the same type of relationship is true for bound inflectional morphemes.

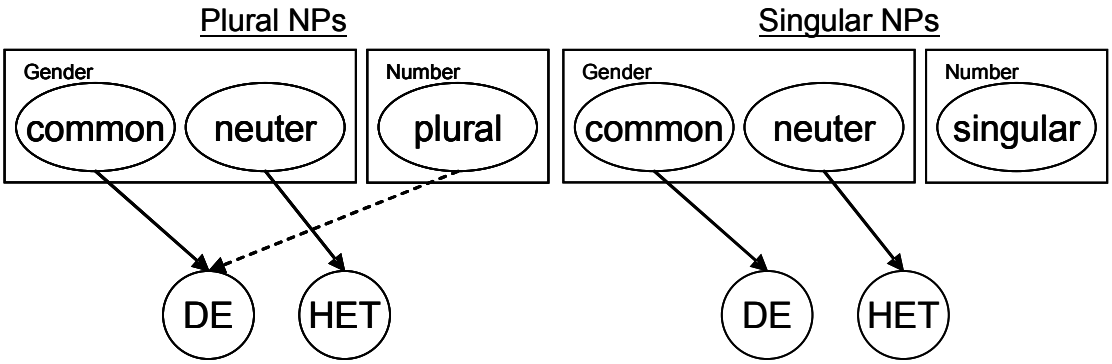


Figure 2. Hypothetical predictions of the competitive and the non competitive lexical selection hypotheses: top row without a main effect of Number in the bare noun naming condition, and bottom row with such an effect. The left column shows cost-type interactions under the assumption of competitive selection. The right column shows benefit-type interactions under the assumption of non-competitive selection (see text for details).

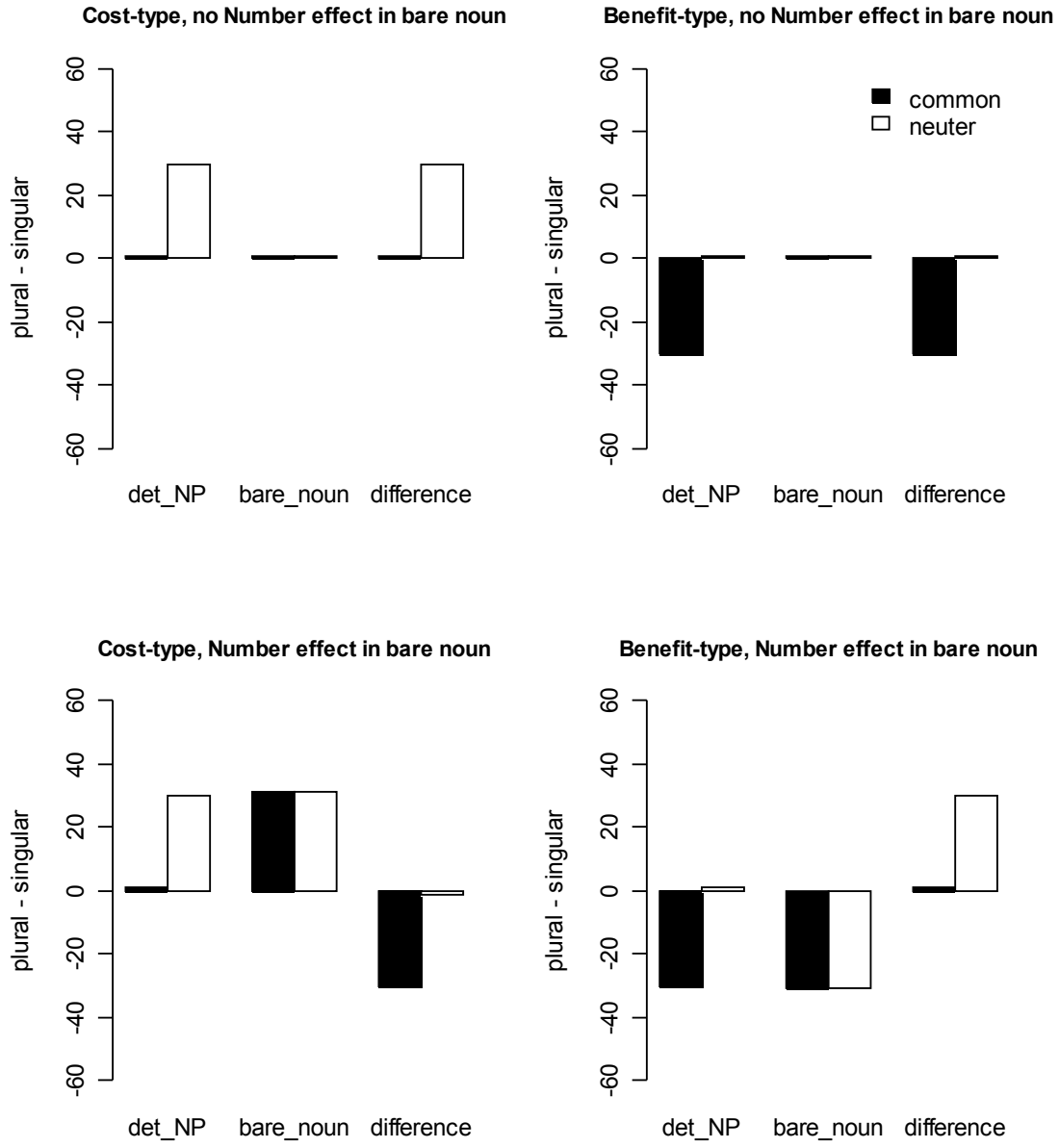
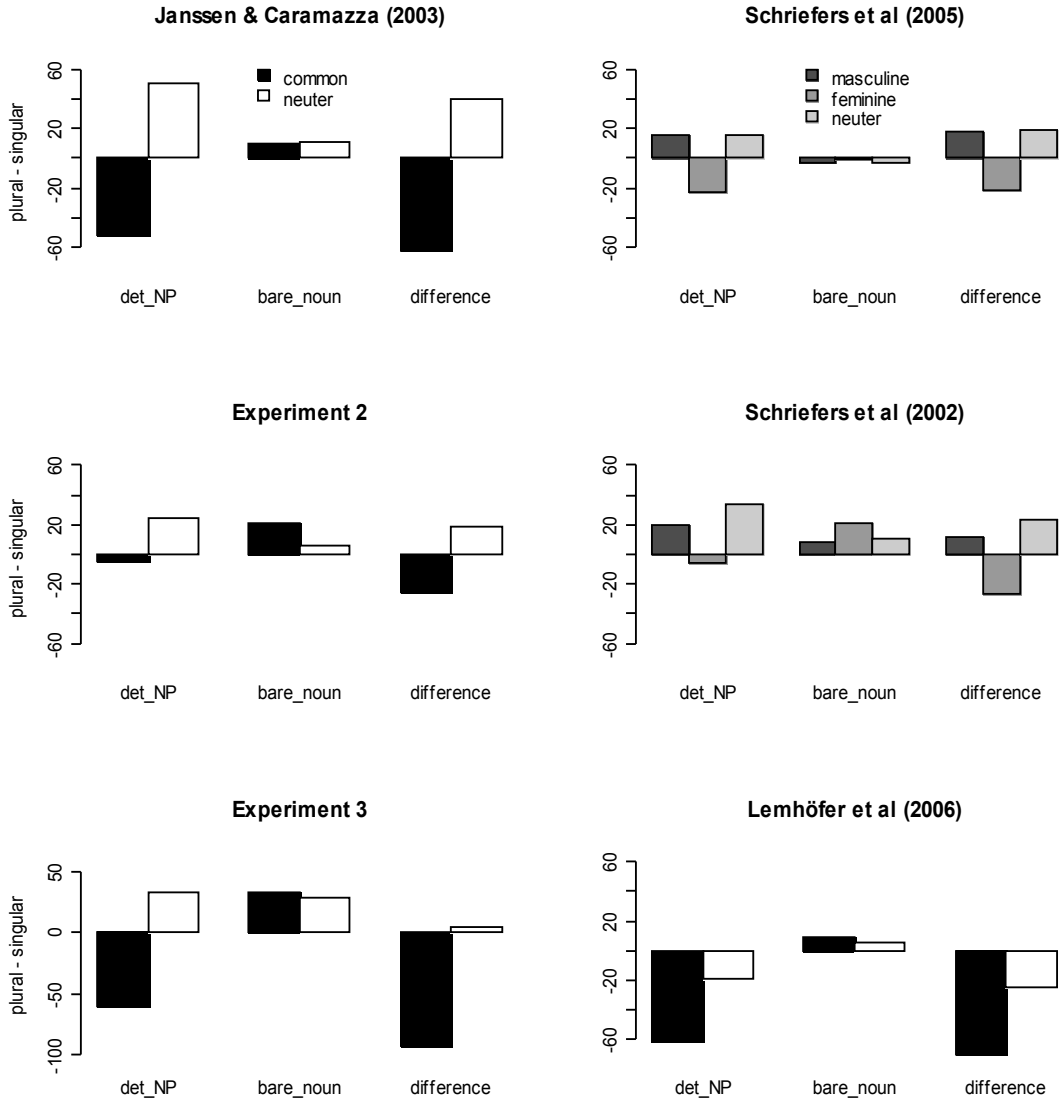


Figure 3. Graphical overview of the studies in the literature that have used the SPN task, following the presentation used in Figure 2.



Appendix A

Overview of F and p statistics for the *analyses of errors* in the determiner NPs, adjective NPs, and bare noun naming conditions in the picture-word interference task of Experiment 1, and the simple picture naming task of Experiments 2 and 3.

Picture-Word Interference - EXP 1		Participants		Items	
		F(1,19)	*p*	F(1,47)	*p*
Determiner+adjective+noun	Congruency	1.83	= .20	1.50	= .23
Adjective+noun	Congruency	< 1	ns	< 1	ns
Bare noun	Congruency	< 1	ns	< 1	ns

Simple Picture Naming - EXP 2		Participants		Items	
		F(1,19)	*p*	F(1,46)	*p*
Determiner+adjective+noun	Gender	7.54	< .02	3.50	= .07
	Number	< 1	ns	< 1	ns
	Gender*Number	3.72	< .07	3.06	= .09
Adjective+noun	Gender	4.06	= .06	3.70	= .07
	Number	6.75	< .02	3.78	= .06
	Gender*Number	< 1	ns	< 1	ns
Bare noun	Gender	< 1	ns	< 1	ns
	Number	3.97	= .06	5.10	< .05
	Gender*Number	< 1	ns	< 1	ns

Simple Picture Naming - EXP 3		Participants*		Items*	
		F(1,14)	*p*	F(1,38)	*p*
Determiner+noun	Gender	< 1	ns	< 1	ns
	Number	< 1	ns	< 1	ns
	Gender*Number	7.12	< .02	8.92	< .005
Adjective+noun	Gender	7.10	< .02	2.92	= .09
	Number	1.82	= .19	1.16	= .29
	Gender*Number	2.89	= .11	< 1	ns
Bare noun	Gender	3.70	= .07	1.86	= .18
	Number	7.74	< .02	5.74	< .03
	Gender*Number	4.41	= .06	1.83	= .18

* degrees of freedom values for the by-participants F test in the adjective+noun condition are (1, 19)

Appendix B

Materials used in Experiment 1 and 2.

Gender	Number		Congruency	
	Singular	Plural	Congruent	Incongruent
Common	borstel (brush)	borstels	zwaan (swan)	masker
	spons (sponge)	sponzen	ladder (ladder)	potlood
	schaar (scissors)	schaar	leeuw (lion)	wiel
	vos (fox)	vossen	harp (harp)	wapen
	bijl (axe)	bijlen	schoen (shoe)	vliegtuig
	lepel (spoon)	lepels	muur (wall)	been
	aap (monkey)	apen	bloem (flower)	ei
	riem (belt)	riemen	appel (apple)	spook
	douche (shower)	douches	vork (fork)	brein
	sigaar (cigar)	sigaren	dolfijn (dolphin)	vlot
	bel (bell)	bellmen	muts (hat)	fornuis
	tent (tent)	tenten	jas (jacket)	stuur
	lamp (lamp)	lampen	raket (rocket)	bot
	kroon (crown)	kronen	beer (bear)	varken
	bril (glasses)	brillen	auto (car)	hart
	jurk (dress)	jurken	fiets (bike)	schaap
	bus (bus)	bussen	sleutel (key)	horloge
	vogel (bird)	vogels	trechter (funnel)	harnas
	kast (cupboard)	kasten	pijp (pipe)	orgel
	Spiegel (mirror)	spiegels	boot (boat)	brood
	hond (dog)	honden	fles (bottle)	blad
	stoel (chair)	stoelen	berg (mountain)	zout
	tafel (table)	tafels	zeep (soap)	web
	deur (door)	deuren	molen (mill)	vergiet
Neuter	skelet (skeleton)	skeletten	masker (mask)	zwaan
	hert (deer)	herten	potlood (pencil)	ladder
	anker (anchor)	ankers	wiel (wheel)	leeuw
	konijn (rabbit)	konijnen	wapen (weapon)	harp
	tapijt (carpet)	tapijten	vliegtuig (airplane)	schoen
	zadel (saddle)	zadels	been (leg)	muur
	zwaard (sword)	zwaarden	ei (egg)	bloem
	net (net)	netten	spook (ghost)	appel
	nest (nest)	nesten	brein (brain)	vork
	hek (fence)	hekken	vlot (raft)	dolfijn
	touw (rope)	touwen	fornuis (oven)	muts
	pistol (gun)	pistolen	stuur (wheel)	jas
	kasteel (castle)	kastelen	bot (bone)	raket
	mes (knife)	messen	varken (pig)	beer
	gewicht (weight)	gewichten	hart (heart)	auto
	oor (ear)	oren	schaap (sheep)	fiets

dak (roof)	daken	horologe (watch)	sleutel
eiland (island)	eilanden	harnas (harnas)	trechter
slot (lock)	sloten	orgel (organ)	pijp
vuur (fire)	vuren	brood (bread)	boot
paard (horse)	paarden	blad (leaf)	fles
raam (window)	ramen	zout (salt)	berg
boek (book)	boeken	web (web)	zeep
bed (bed)	bedden	vergiet (colander)	molen

* Words under Number/Singular were the pictures in Experiment 1 and words under Congruency were the distractors in Experiment 1. Words under Number were the pictures in Experiment 2.

Appendix C

Materials in Experiment 3

Gender	Number	
	Singular	Plural
Common	arm (arm)	armen
	auto (car)	autos
	beer (bear)	beren
	berg (mountain)	bergen
	bijl (axe) *	bijlen
	borstel (brush)	borstels
	bus (bus)	bussen
	deur (door)	deuren
	fiets (bike)	fietsen
	gieter (watering can) *	gieters
	hond (dog)	honden
	jurk (dress) *	jurken
	kaars (candle)	kaarsen
	kikker (frog)	kikkers
	krant (newspaper) *	kranten
	lamp (lamp)	lampen
	leeuw (lion) *	leeuwen
	molen (windmill)	molens
	pijl (arrow)	pijlen
	raket (rocket) *	raketten
	riem (belt)	riemen
	sigaar (cigar) *	sigaren
	stekker (plug)	stekkers
	stoel (chair)	stoelen
	trompet (trumpet)	trompetten
	vlieg (fly) *	vliegen
	voet (foot)	voeten
	vogel (bird) *	vogels
	vos (fox)	vossen
	wortel (carrot) *	wortels
Neuter	anker (anchor)	ankers
	bed (bed)	bedden
	been (leg)	benen
	boek (book)	boeken
	bureau (desk)	bureaus
	fornuis (oven) *	fornuizen
	hart (heart)	harten
	hek (fence)	hekken
	hert (deer) *	herten
	kasteel (castle) *	kastelen

kompas (compass) *	kompassen
konijn (rabbit)	konijnen
mes (knife)	messen
nest (nest) *	nesten
net (net)	netten
oor (ear)	oren
paard (horse)	paarden
pistol (gun) *	pistolen
potlood (pencil)	potloden
raam (window) *	ramen
schaap (sheep) *	schapen
skelet (skeleton) *	skeletten
slot (lock)	sloten
touw (rope)	touwen
varken (pig)	varkens
vergiet (colander) *	vergieten
vest (vest)	vesten
wiel (wheel)	wielen
zadel (saddle)	zadels
zwaard (sword)	zwaarden

* Pictures excluded from the adjective+noun and bare noun conditions.