J Psycholinguist Res (2008) 37:69–85
DOI 10.1007/s10936-007-9060-0

ORIGINAL ARTICLE

Grammatical Gender in Speech Production: Evidence from Czech

Denisa Bordag · Thomas Pechmann

Published online: 6 November 2007

Abstract Three experiments demonstrate gender congruency effects (i.e., naming times of a picture are faster when the name of the target picture and a distractor noun are gender congruent) in Czech. In the first experiment, subjects named the pictures by producing gender-marked demonstrative pronouns and a noun. In the second and third experiments, subjects produced a gender-marked numeral (marked with a suffix) plus a noun. Two types of such suffixes exist in Czech. Some numerals vary in nominative singular with gender, others do not. The results show significant gender congruency effects in all experiments. They suggest that gender congruency effects can be obtained not only with free, but also with bound morphemes. In the second and third experiment the effect only emerged when the suffix was gender-marked (as opposed to gender-invariant), supporting the view that the gender congruency effect is due to competition at the level of phonological forms rather than at the grammatical level.

Keywords Grammatical gender · Czech · Gender congruency effect · Speech production

For several years there has been an ongoing debate on the nature of the most prominent effect observed in gender processing, the so-called gender congruency effect. The discussion started in 1993, when Schriefers reported this effect in Dutch for the first time (Schriefers 1993). In a picture-word interference experiment, he obtained faster reaction times when the target picture and the distractor word had the same gender than when they had a different gender. According to Schriefers, the effect reflects the competition for selection of a word's syntactic features: The activation of the gender node of the distractor word interferes with the selection of the target noun gender in the gender incongruent condition thus causing the delay in the picture

D. Bordag (✉) · T. Pechmann
Department of Linguistics, University of Leipzig, Beethovenstr. 15, 04107 Leipzig, Germany
e-mail: denisav@rz.uni-leipzig.de

naming compared with the condition, where both the target word and the distractor word have congruent genders.

Using the picture-word interference paradigm, several authors have addressed the question of the gender congruency effect in various languages since Schriefer's seminal paper, sometimes replicating his results (Van Berkum 1997; La Heij et al. 1998; Schriefers and Teruel 2000; Schiller and Caramazza 2002; for a review see Schriefers and Jescheniak 1999), and sometimes not (Costa et al. 1999; Alario and Caramazza 2002; Pechmann and Zerbst 2004). Gradually it became obvious that several aspects constrain the congruency effect and that its interpretation might be more complicated than suggested by Schriefers (1993). First, the gender congruency effect has been proven to be language specific, thus raising the question about cross-linguistic differences in a gender, or broader, in an NP production mechanism. Second, the question has also been raised, whether the effect is indeed a gender congruency effect or rather a determiner congruency effect (Schiller and Caramazza 2002)—in other words, whether it reflects competition at the grammatical (abstract gender nodes) or at the phonological level (phonological forms). Third, researchers have presented conflicting evidence concerning the scope of the congruency effect. Whereas some authors observed the effect only when an NP in the form of a free morpheme (article or other determiner) + noun are produced (Schiller and Caramazza 2003), others (Schriefers 1993) also obtained the effect with NPs consisting of an adjective ending with a gender-marked inflection (a bound morpheme) + a noun. Moreover, the effect only occurred when subjects named the pictures with gender-marked NPs, but not when bare nouns were produced, suggesting that the gender feature may not be always selected. Clearly, more empirical evidence and theoretical considerations are necessary to solve the apparent contradictions between existing results and their interpretations. The clarification of these issues is of special importance for present models of speech production. First, they broaden our knowledge about language specific versus general production mechanisms. Second, they address some fundamental topics in speech production modeling like modularity versus interactivity. Finally, they shed new light on the processing and storage of inflected forms. The experimental evidence reported in this paper contributes to all these areas.

Early and Late Selection Hypothesis

The selection of determiners or other modifiers is driven by different kinds of information. The selection mechanism must be supplied with the necessary semantic information: Is the object blue or red (if a color adjective is required), is it possessed by the speaker or by the speaker and somebody else (if a personal pronoun like *my* or *our* is required) etc. Lexical-grammatical information might also be required, for example the grammatical gender of the head noun (*klein-er Hund, klein-e Katze, klein-es Kind* in German) or the declensional class of the modifier itself (in Czech, see later). Phonological information may also play a role—the form of a determiner may depend on the form of the following word (another modifier or the head noun itself), e.g., whether it begins with a vowel or a consonant (e.g., *a giraffe* versus *an elephant*, *ma table* versus *mon ampoule* in French). The information needed for modifier selection is obviously language specific, which can partly account for differences in findings concerning the congruency effect in various languages.

An important cross-linguistic difference was pointed out by Miozzo and Caramazza (1999) and by Caramazza et al. (2001). In several experiments with Italian speakers, they systematically failed to observe the gender congruency effect reported for German and Dutch (Schriefers 1993; Van Berkum 1997; La Heij et al. 1998; Schriefers and Teruel 2000; Schiller and Caramazza 2002; for a review see Schriefers and Jescheniak 1999). Similarly, Costa et al. (1999) failed to observe the congruency effect in Catalan and Spanish and Alario and Caramazza (2002) in French (cf., however, Schriefers and Teruel (1999) for contrasting evidence from French). Caramazza and his colleagues argued that the effect is present only in languages in which the determiner form depends solely on the grammatical gender of the head noun (as it is the case in German and Dutch) and where the determiner selection can thus occur very early in the NP production process. In the Romance languages, where the specification of grammatical gender is not sufficient to retrieve the determiner form because the information about the phonological form of the head noun is also needed, the determiner selection is delayed until the phonological properties of the context are available. In Italian, for example, the singular masculine definite article can be either *il* or *lo*, depending on the word that follows: If the following word begins with a vowel, a consonant cluster of the form *s* + consonant, or *gn*, or an affricate, then the proper masculine determiner is *lo* (*lo sgabello* [the stool], *lo gnomo* [the gnome]; in other cases the correct determiner is *il*. Consequently, the determiner selection in languages like Italian must be a relatively late process. It can take place only after the phonological form of the word that follows has been specified. Based on this distinction, Caramazza differentiates between so called *early* and *late selection language*s. The congruency effect is observable only in the early selection languages because here the selection can be made as soon as the gender of the head noun becomes available. In the late selection languages, the activation of competing gender information would long have dissipated, rendering any competing activation ineffective.

Locus of the Congruency Effect

The proposal that the so called gender congruency effect in fact reflects the processes at the level where phonological forms and not grammatical features are selected, was first formulated by Miozzo and Caramazza (1999) in connection with their failure to find the effect in Italian and their late selection hypothesis. Their original claim was that if the congruency effect reflects competition between grammatical features and not between the forms of the determiner, then the effect would also have to be observable in the bare noun condition. However, this claim could not hold against the argument that grammatical gender is selected only when needed for production, i.e., when an NP with a gender-marked element and not a bare noun is produced (Levelt et al. 1999). Better evidence in support of the hypothesis that the congruency effect reflects processes at the level of phonological form was provided by Schiller and Caramazza (2002), who compared the production of German and Dutch NPs in singular and plural. In these two languages, there is only one definite article form for plural (*die* in German, *de* in Dutch), but three different forms in singular in German (*der, die, das*) and two (*het, de*) in Dutch. Whereas the singular forms are overtly marked for the three genders (masculine, feminine, neuter) in German and the two genders (common, neutral) in Dutch, the plural article form is invariant. Schiller and Caramazza conducted several experiments with consistent results: They observed

the congruency effect for singular NPs only and not for plural NPs. The authors reasoned that if the congruency effect was located at the level of grammatical feature selection (the so-called gender selection interference hypothesis, GSIH), longer naming latencies in the incongruent condition should have been observed both in singular and plural, i.e., irrespective of whether or not the gender feature surfaces in an overtly gender marked determiner. The fact that the effect appeared only in the singular supports, according to Schiller and Caramazza, the so-called determiner selection interference hypothesis (DSIH) claiming that the congruency effect originates at the level, where the phonological forms are selected. It is the phonological forms themselves (here determiners), not their abstract gender features, which compete for selection in singular. In plural there is no competition (only one determiner form), therefore no congruency effect can be observed.

Scope of the Congruency Effect: Bound versus Free Morpheme Hypothesis

Conflicting evidence has been collected concerning the scope of the congruency effect. Does the competition for selection apply only to free-standing morphemes like determiners, or is the retrieval of gender-marked inflections governed by the same principles? Studies exploring this issue are scarce and present contradictory results. Schriefers (1993) obtained a congruency effect when subjects produced NPs in the form of a definite article + adjective + noun, as well as in the form adjective + noun, suggesting that either the gender features compete for selection (GSIH), or else there is a competition for selection of the bound morphemes associated with the gender inflection of the adjective. The effect in the adjective + noun utterances was however substantially reduced (31 ms) compared to the effect in the definite article + adjective + noun utterances (56 ms). Schiller and Caramazza (2003), however, failed to replicate this experiment in both German and Dutch. Costa et al. (2003) could not observe the effect in adjective + noun phrases in Croatian, but they found the effect with personal pronouns in accusative case, that is with free morphemes. They conclude that the gender congruency effect only emerges when grammatically determined free-standing morphemes are produced. Based on this conclusion, Costa et al. (2003) further hypothesized that inflected forms are rather produced by transforming the base form into another form than by concatenative attaching of affixes to bare stems. If the latter case is true, inflectional suffixes would have independent phonological status, similarly to determiners. Consequently, it would be expected that the selection of the inflection markers is also a competitive process, for which the authors failed to find evidence [see Costa et al. (2003) for details].

The hypothesis that determiner forms (and not grammatical features) compete for selection is crucial for one more reason: It implies cascaded processing during speech production. Discrete serial models like that of Levelt (Levelt 1989; Levelt et al. 1999) assume that processes at the level of grammatical encoding (including e.g., the selection of a gender feature) are completed before the encoding at the phonological level starts, and that only the phonological form of the selected lexical node is encoded. This is incompatible with the claim that more than one determiner form is activated in the incongruent condition and that the interference arises from their competition for selection. On the other hand, cascaded or parallel activation models like those of Dell (1986) or Caramazza (1997) can account for the competition between two or more determiners at the level of form selection: All phonological forms (e.g., determiners)

that come into question are activated before the selection processes are resolved at the higher level of grammatical encoding. The specification of the locus of the congruency effect can thus contribute to a more general aspect of speech production theories, namely the issue of serial versus cascaded processing (Schiller and Caramazza 2003).

In this study we want to investigate the gender/determiner congruency effect in Czech. Czech is a highly inflected Slavic language with three genders (masculine, feminine, and neuter) and no articles. According to Caramazza's distinction, it belongs to the so-called early selection languages and therefore, it is expected that the congruency effect should be observed here. In the first experiment we addressed this issue using the picture-word interference paradigm. Subjects had to name pictures with NPs consisting of a demonstrative pronoun *ten* (m), *ta* (f), *to* (n) + a noun. The expected congruency effect was obtained. In the second experiment we took advantage of the presence of the so-called hard and soft adjectival declension in the Czech language to explore the locus and the scope of the congruency effect in more detail.

Experiment 1

In this experiment we attempted to replicate the gender/determiner congruency effect in Czech. It has been shown in previous research that the effect is present only in some languages and it is not certain whether all constraints determining the emergence of the congruency effect have already been discovered. According to the late/early selection distinction between languages, Czech should belong to the latter because the choice of a modifier form is not affected by its phonological context. Since there are no articles in Czech, an overtly gender marked demonstrative pronoun *ten* (m.), *ta* (f.), *to* (n.) [this] had to be produced in the naming task together with the target noun. The morphological status of this pronoun is not unambiguous: It could be interpreted as a complex word form consisting of a stem *t-* and a gender marked inflection *–en, -a, -o* (for nominative singular), but the interpretation that it is a free-standing morpheme seems to be better warranted. However, if we do not observe the congruency effect in this experiment, the status of the pronoun must be reconsidered, because in the previous research an effect was yielded mostly when gender marked freestanding morphemes were produced.

Method

Subjects

Thirty-two Czech native subjects, consisting mostly of students visiting the University of Leipzig for one or more semesters, participated in this experiment. They were paid for their participation.

Materials

A total of 36 pictures corresponding to monomorphemic (with two exceptions) Czech nouns were selected for naming. Twelve of the nouns were masculine, 12 feminine, and 12 neuter. All nouns ended with an inflection typical for their gender, i.e., a consonant for masculines, *-a* for feminines, and *−o* for neuters. Within each gender group, there was the same number of two and three syllable words. The same nouns were used also

as distractors. In addition to the 36 experimental items, there were also 12 practice items, four of each gender. The list of items used in the experiment can be found in the appendix.

The pictures were white line drawings presented on a black background. The distractors were also displayed in white characters (font type ibm, size 20 pts.).

The stimuli were presented in four blocks of 36 items (plus two warm-up items at the beginning of each block), 12 of each gender. In each block each picture appeared only once, either with a congruent, incongruent, identical or a neutral (xxxxx) distractor. The identical and neutral conditions allow us to control whether subjects process the distractors. Previous research also revealed that the presence of an identical condition enhances the distractor effects, especially if no semantic condition is included (Pechmann and Schriefers unpublished). Throughout the experiment, subjects thus saw each picture four times, once in each condition.

Similarly, each distractor appeared three times in the whole experiment, once in each condition, except for the neutral one. The repetition of pictures over blocks and participants was counterbalanced.

The items in each block were randomized, so that not more than two pictures of the same gender could follow each other, two pictures in the same condition were not allowed to follow each other and a distractor could appear only twice in direct succession in the same position (above or below the picture).

The following restrictions applied on the picture–word distractor combinations:

- the onsets were not phonologically similar,
- the target and the distractor word were not semantically related,
- the pictures that appeared in the incongruent condition with a distractor of one of the two remaining genders in one version was paired with a distractor of the other gender in the second version, so that, for example, a picture of a masculine appeared with a feminine for half of the subjects and with a neuter distractor for the other half of subjects in the incongruent condition.

Procedure

Subjects were tested individually in an experimental room in the presence of an experimenter, who scored for potential errors. They were given written instructions on the task they would perform and on the functioning of the voice key. The experiment was run using the ERTS software (Beringer 1999). The stimuli were presented on a NEC MultisyncXV 17 in. monitor and the responses were registered with an ECM-MS957 microphone connected to an ERTS voice-key.

The experiment consisted of three parts: a familiarization, a practice, and an experimental phase, and lasted approximately 30 min. During the familiarization session subjects were presented with all pictures that were used in the experiment. The name of each picture was printed below it. Participants were asked to pronounce the word and to memorize the picture with which it appeared. As soon as the experimenter pressed the space bar, the next picture appeared on the screen.

In the practice session pictures alone were randomly presented on the screen. Subjects named them with the demonstrative pronoun *ten, ta, to* and the appropriate noun. At their response, the picture and the distractor disappeared from the screen. If they did not use the name that was designated for a given picture, they were corrected

by the experimenter, who controlled the speed of the picture presentation by the space bar.

The last phase was the experiment itself: Subjects named the pictures in the presence of distractors, again using the demonstrative pronouns *ten, ta, to*, e.g., *ten ananas, ta kytara, to auto*. The SOA was 0. The stimuli were presented in the four blocks with a break between them. Each block began with two warm-up items, which were later excluded from the analyses.

Each trial started with a fixation point (*), which appeared in the middle of the screen for 1,000 ms, then a blank screen appeared for 500 ms, followed by both the picture and the distractor. After the voice key had been triggered, both the picture and the distractor disappeared from the screen, the experimenter scored the response for errors (by pressing either the right or left CTR key) and after a pause of 500 ms a new trial began. When the voice key was triggered, a small white square appeared in the upper left hand corner of the screen, signalling that the reaction time had been measured. If the voice key was not triggered within 1,500 ms (time out) after picture presentation, a sign "no voice key" appeared in the upper left hand corner of the screen and the trial was scored as incorrect.

Results and Discussion

Observations were discarded from the analyses whenever any of the following conditions held: a non-expected picture name had been used; the demonstrative pronoun was omitted; a non-speech sound preceded the utterance of the picture name or the pronoun, triggering the voice key; a disfluency occurred or an utterance was repaired; a speech onset latency was shorter than 200 ms or exceeded 1,500 ms, or deviated from the participant's mean by more than three standard deviations. On the basis of these criteria, a total of 292 observations (6.3%) were marked as incorrect.

As Table 1 shows, overall naming latencies in the gender-incongruent condition were 26 ms slower than naming latencies in the congruent condition. Analyses of variance (ANOVAs) were performed on the mean naming latencies per subject (F1) and per item (F2) with the distractor type (four levels) as a variable. They signaled a significant difference between the four conditions: $F1\ (3, 93) = 52.64$, $p < .01$; $F2\ (3, 105) = 72.63$, $p < .01$. Paired-samples t-tests comparing only the gender congruent and incongruent conditions revealed that this effect was significant: $t1(31) = 4.59$, $p < .01$; $t2(35) = 3.91$, $p < .01$.

The analyses clearly show a robust congruency effect in Czech, when pictures are named with a gender marked NP consisting of a demonstrative pronoun and a noun. They also provide further evidence that there is a competition for selection either between the gender features or between the demonstrative pronoun forms in early selection languages. In the next experiment we investigated the scope and the locus of the congruency effect in more detail.

Experiment 2

In this experiment we addressed two questions: First, is the congruency effect in Czech only present when freestanding gender marked morphemes are produced, or does it also appear in the production of NPs, in which the gender is marked on an inflectional

Table 1 Results obtained in the gender-congruent, gender-incongruent, identical, and neutral conditions in Experiment 1 (reaction times [RT] in ms and errors in %)

	Incongruent	Congruent	Identical	Neutral
RT	703.47 (11.1%)	677.45 (8.1%)	626.38 (2.4%)	639.95(3.7 %)

suffix? Second, does the inflection have to be formally marked for gender, or will the congruency effect also be observable with a formally invariant inflection?

To answer these questions we took advantage of the fact that there are two types of adjectival declension in Czech, the so-called hard and soft declension. Both these classes do inflect for gender, but in some cases, they take an invariant inflection for all three genders. For example, in nominative singular, they differ in that in the hard declension there are thee different endings for the three genders (*-ý, -á, -é* for masculines, feminines, and neuters respectively), whereas there is only one invariant ending *–í* for all three genders in the soft declension. On the other hand, in e.g., genitive singular both hard and soft adjectives take gender marked endings, which is not the case for e.g., genitive plural, where both hard and soft adjectives have one and the same gender invariant ending. Not only adjectives, but also some ordinal numerals belong to the adjectival declension. The ordinals *druh-ý, -á, -é* (second), *čtvrt-ý, -á, -é* (fourth), *pát-ý, -á, -é* (fifth) are declined according to the hard declension, the ordinals *prvn-í* (first) and *třet-í* (third) according to the soft declension. Color adjectives, which have been repeatedly used in other experiments on similar topics, all follow the hard declension (or are indeclinable) in Czech and are therefore unsuitable for the planned comparisons. In the experiment subjects therefore named pictures with an NP comprising either an ordinal from a hard adjectival declension (*druhý* [second] or *pátý* [fifth]) + a noun in one block, or an ordinal from a soft adjectival declension (*první* [first] or *třetí* [third]) + a noun in the other block.

Method

Subjects

Sixteen Czech native subjects participated in this experiment. They were also mostly exchange students studying temporarily at the University of Leipzig. None of them had participated in the previous experiment. They were paid for their participation.

Materials

The same 36 pictures as in Experiment 1 were used. The structure of the experiment was similar to the previous experiment, apart from the fact that subjects saw each picture eight times: four times in the soft (declension) condition, and four times in the hard (declension) condition. Pictures were again presented with gender congruent, incongruent, identical, and neutral distractors. The picture—distractor pairing was the same as in Experiment 1. In the hard condition, pictures were named with the ordinal *pátý* (fifth) in the congruent and incongruent conditions and with the numeral *druhý* (second) in the identical and neutral conditions (fillers). In the soft condition, pictures were named with the ordinal *první* (first) in the congruent and incongruent conditions and with the numeral *třetí* (third) in the identical and neutral conditions.

Table 2 Results from Experiment 2 (reaction times [RT] in ms and errors in %)

	Incongruent	Congruent	Identical	Neuter
Hard	786.78 (15.6%)	749.76 (11.5%)	697.81 (3.0%)	712.90 (5.6%)
Soft	712.90 (6.4%)	704.09 (8.3%)	690.44 (4.0%)	703.86 (4.7%)

Thus, the NPs from the two critical conditions (congruent versus incongruent) began with the same onset /*p*/ in both the soft and the hard block. Half of the subjects saw the hard block first, and the other half the soft block first.

To initiate the naming with the particular ordinals, the pictures (and the distractors) were presented in frames consisting of one or three lines in the soft block (corresponding to the ordinals *first* and *third*), and with two or five lines in the hard block (corresponding to the ordinals *second* and *fifth*).

Procedure

The procedure was similar to that of the previous experiment. The main difference was that two more conditions, hard and soft declension blocks, were present. The experiment lasted approximately 40 min and again had three phases: a familiarization, a practice, and an experimental phase. The familiarization phase was the same as in Experiment 1. In the practice phase, pictures appeared in the frames without any distractors. Subjects who saw the hard block first had to name the pictures with the ordinals from the hard declension in the practice phase, subjects who saw the soft block first with the ordinals from the soft declension. Participants were instructed to name the pictures in the more complex frames (three lines in the soft block, five lines in the hard block) with the higher ordinals and pictures in the less complex frames with the lower ordinals. The experimental phase started with a practice block of 12 practice items either from the hard or from the soft condition, depending on whether a hard or a soft block followed. Before the other block (hard versus soft) started, there was again a practice block of 12 items from the appropriate condition. Subjects always listened to the corresponding instructions before each experimental phase. The trial structure was identical to that in Experiment 1.

Results and Discussion

Observations were excluded on the basis of the same criteria as in Experiment 1 and additionally, in those situations where subjects named a wrong ordinal. Thus, 340 observations (7.4%) were discarded.

As Table 2 shows, subjects were 37 ms slower in the gender incongruent than in the gender congruent condition when naming the pictures with an ordinal from a hard declensional class and only less than 9 ms when the ordinal was from a soft declensional class. Analyses of variance (ANOVAs) were performed on the remaining mean naming latencies per subject (F1) and per item (F2) with Congruency and Type of Declension as independent variables (the control conditions [identical and neutral] were excluded from the analyses). They yielded significant main effects of Congruency: $F1(1, 15) = 11.04$, $p < .01$; $F2\ (1, 35) = 12.35$, $p < .01$, Type of Declension (soft

versus hard): $F1(1, 15) = 14.50$, $p < .01$; $F2\ (1, 35) = 103.11$, $p < .01$, and a significant interaction between these two factors: $F1\ (1, 15) = 5.94$, $p < .05$; $F2\ (1, 35) = 6.58$, $p < .05$. A Scheffé post hoc test revealed that, whereas the two conditions (congruent versus incongruent) in the soft condition did not differ from each other, the gender congruent and gender incongruent conditions in the hard condition differed both from the soft condition and from each other. In other words, the gender congruency effect was only observed in the hard condition.

The results obtained in Experiment 2 are important for several reasons. First, they demonstrate that the congruency effect from the first experiment can be replicated, showing that it is a robust effect in Czech. Second, the effect emerged only in the hard condition, where the ending is gender marked. In the soft condition, where the adjective has an invariant ending for all three genders, no congruency effect was observed. This finding indicates that it is rather the phonological forms, not the abstract gender features, which compete for selection. Finally, the fact that the effect emerged at all in the experiment, where no determiners, but rather ordinals with gender marked inflections were produced, sheds doubts on the claims that only free-standing morphemes can compete for selection. The consequences of these findings are discussed in detail in the General Discussion.

However, it could be argued that the absence of the gender congruency effect in the soft condition rather reflects a strategy applied to the particular experimental task than automatic processes employed in speech production. In Experiment 2, subjects named pictures in two blocks. In one block only soft ordinals were used, in the other block only hard ordinals. Consequently, subjects produced one and the same ending *–í* throughout the whole soft block. It could be the case that after several soft ordinals were produced, the production system registered that always one and the same ending is needed and accommodated to this situation by bypassing the gender selection (and maybe even the complete grammatical encoding) and simply retrieved the ending *–í* all the time. This could be also an alternative explanation, why the reaction times were generally faster in the soft condition.

To distinguish between the two possible interpretations of the effect, we modified Experiment 2 by reorganisation of the ordinals in the two blocks, so that in each block both soft and hard ordinals were produced. If the congruency effect in the soft condition is absent in this design as well, then it cannot be attributed to a task strategy.

Experiment 3

Method

Subjects

Fourteen Czech native subjects from the same population as the subjects in Experiment 1 and 2 participated in this experiment. None of them had participated in the previous experiments. They were paid for their participation.

Materials

Materials were identical with those in Experiment 1 and 2. There were again two conditions, soft and hard (crossed with gender congruent and incongruent). This time

Table 3 Results from Experiment 3 (reaction times [RT] in ms and errors in %)

	Incongruent	Congruent	Identical	Neuter
Hard	751.16 (14.3%)	713.70 (13.0%)	594.82 (5.4%)	610.09 (7.1%)
Soft	654.25 (8.9%)	647.91 (7.7%)	630.29 (3.8%)	645.04 (5.6%)

their division into blocks was different. In one block, pictures in the congruent and incongruent conditions were named with the hard ordinal *pátý* (fifth) and in the identical and neutral conditions (fillers) with the soft ordinal *třetí* (third). In the other block, pictures were named with the soft ordinal *první* (first) in the congruent and incongruent conditions and with the hard ordinal *druhý* (second) in the identical and neutral conditions. Thus, in both blocks ordinals from a hard and a soft declension had to be named and no ending could be selected in advance, because subjects could not predict, whether a soft or a hard ordinal would be produced.

Procedure

The procedure was the same as in Experiment 2.

Results and Discussion

Observations were excluded on the basis of the same criteria as in Experiment 2. Thus, 330 observations (8.9%) were discarded from further analyses.

As Table 3 shows, the effects in Experiment 3 even numerically almost exactly mirror the results of Experiment 2. Subjects were again 37 ms slower in the gender incongruent than in the gender congruent condition when they named the pictures with a hard ordinal and only 6 ms when the ordinal was from a soft declensional class. Analyses of variance (ANOVAs) yielded significant main effects of Congruency: $F1\ (1, 13) = 18.45$, $p < .01$; $F2\ (1, 35) = 11.08$, $p < .01$, Type of Declension (soft versus hard): $F1\ (1, 13) = 10.92$, $p < .01$; $F2(1, 35) = 206.88$, $p < .01$, and a significant interaction between these two factors: $F1\ (1, 13) = 21.38$, $p < .05$; $F2\ (1, 35) = 5.22$, $p < .05$. A Scheffé post-hoc test again revealed that the congruent and incongruent conditions did not differ from each other in the soft condition. In the hard condition they differed both from the soft condition and from each other: The gender congruency effect was only present in the hard condition.

In this experiment we thus replicated the results of Experiment 2. We again observed a robust congruency effect, however only when subjects produced NPs with ordinals from the hard declension, which have different inflectional endings for masculine, feminine, and neuter nouns in nominative singular. The effect was not observed when subjects produced soft ordinals with an invariant ending for all three genders in nominative singular, suggesting that the overt gender marking is the crucial variable. The ordinals from the soft and hard declension were named in random order. Thus subjects could not develop a strategy according to which gender selection in the soft condition would be disregarded on the basis of the nature of the task, as it might have been argued in the previous experiment, where in one block only items from one condition (hard or soft) were named.

General Discussion

The experiments reported in this study provide new evidence with regard to all three areas of research mentioned in the introduction concerning the so-called gender congruency effect. First, the early selection hypothesis predicts that the congruency effect should emerge in languages, where the form of the gender marked element does not depend on its phonological context, but can be accomplished as soon as the gender feature of the noun is retrieved. According to this hypothesis, Czech belongs to the so-called early selection languages, and thus the congruency effect should be present. Indeed, this is what the data in all three experiments demonstrated.

Second, there is conflicting evidence regarding the locus of the congruency effect. Schiller and Caramazza (2002) argue that if the competition for selection takes place at the level where grammatical features are selected, the congruency effect should emerge irrespective of whether it surfaces in a gender marked element or in an invariant element for all genders. However, their experiments with singular and plural noun phrases in Dutch and German consisting of a definite article + a noun demonstrate that the congruency effect can only be observed in the singular, where the definite article has a different form in all genders. In the plural, where the definite article *die* or *de* is identical for all genders in German and Dutch respectively, the effect was absent.

The data from our Experiments 2 and 3 are in line with the data reported by Schiller and Caramazza (2002). The congruency effect appeared only in the hard condition, i.e., when subjects produced utterances where gender overtly surfaced as an inflection, which is suffixed to the lexical stem of a ordinal. In the soft condition, where the inflection has the same form *–í* for all three genders in nominative singular, the congruency effect was not observed. According to Schiller and Caramazza (2002), this indicates that it is the phonological forms of the inflections (in Schiller and Caramazza's terms of the determiners) that compete for selection, not the abstract gender features of the head nouns as argued by Schriefers (1993). Furthermore, if the locus of the congruency effect is at the level of phonological encoding, as the data suggest, then cascaded processing must be involved. Discrete serial processing, as proposed by Levelt (1989) and Levelt et al. (1999), does not allow other than the *selected* phonological forms that are activated. Clearly, if the two gender marked forms (inflections or determiners of the picture and of the distractor) can compete for selection, they both must be activated at the same time.

Is there an alternative hypothesis to explain why the congruency effect only emerges when overtly gender marked forms are produced, maintaining the idea of gender feature selection? One possibility is that the gender feature in the soft condition is not selected at all, because it has no consequences for the selection of the appropriate inflection in nominative singular. Under this assumption, it would also be easy to explain the pattern of results observed in Experiments 2 and 3, namely that the naming with soft ordinals is faster than the naming with hard ordinals in both the gender incongruent *and* congruent condition. In the soft condition, gender does not affect further encoding, thus it is bypassed, which results in short naming latencies. In the hard condition, gender selection is obligatory because it is relevant for the selection of the proper adjectival inflection. When the gender of the distractor and the target noun collide (congruent condition), the selection proceeds fast, though still more slowly than in the soft condition. The hard incongruent condition yields

the longest naming latencies, indicating that competition for selection between the gender feature of the distractor and the target noun takes place.

This alternative hypothesis assumes that the locus of the congruency effect is at the level of grammatical encoding, where the selection of the abstract gender features takes place. It further assumes that the production system "knows" when the selection of the gender feature can be bypassed, because it has no consequences for further encoding. But how could the system "know"? One possibility is a hierarchical feature selection hypothesis. In our case it would mean that the feature "declensional class of the modifier" and the feature "case" (because in e.g., genitive or dative singular the inflection for the three genders in the soft condition is no longer invariant) would have to be selected prior to the gender feature. If the declensional class of the modifier was soft and if the case was nominative, then the selection of the gender feature and probably even number (the inflection *–í* appears in all three genders also in nominative plural) would be bypassed. Such a mechanism, which allows selection of some features under one condition and not under another, is not only problematic for anaphoric agreement and elliptical constructions, but, as also noted by Schiller and Caramazza (2003), who consider a similar system, it seems ad hoc and rather unlikely. So far it is, however, the only hypothesis compatible with discrete serial models of processing like the one of Levelt and probably the only one that can currently explain the observed data.[1]

Another hypothesis that could explain the present results while maintaining the claim that they reflect the competition for selection between abstract gender features, assumes cascaded processing and feedback between the levels of phonological and grammatical encoding: The activated (but not yet selected) gender feature(s) send(s) activation to the corresponding phonological form(s). If the activation converges on just one inflection, information is sent back through feedback to the level of grammatical encoding, signalling that the selection of e.g., grammatical gender can be bypassed, because all activation coming from that source converges on just one

[1] An intermediate position concerning the selection of abstract gender features at the level of grammatical encoding is taken e.g., by Schriefers and his colleagues (Schriefers et al. 2002, 2005). According to their account, the abstract gender feature is always selected (also when actually not needed for production), but it is a non-competitive process with no implications for production latencies. In four experiments they compared gender processing in German singular and plural NPs. According to their hypothesis, abstract gender features are selected in a non-competitive way both in singular and plural, but in plural they activate not only the plural, but also the singular determiner forms/endings (singular-as-default hypothesis). They observed facilitation (faster RTs) for feminine noun phrases in plural compared to singular and argued that it is because the feminine singular determiner form *die* is identical with the plural gender invariant form *die*. The convergence of the activation on just one determiner form in the case of feminines explains why the plural determiner *die* of feminine nouns is selected faster than the same plural determiner of masculine and neuter nouns, which have different determiner forms in singular (*der* and *das*, respectively). The gender effect thus originates at the phonological level. The problem with this interpretation is that in such a case, not only facilitatory effects with feminine nouns (reflecting convergence of activation on one phonological form), but also inhibitory effects with masculine and neuter nouns (reflecting the competition between the different phonological forms) should be observed in plural production—as claimed also by the authors. A detailed look at their data however reveals that the evidence for this competition is very inconsistent and putative and in some of their experiments absent altogether. To conclude, so far there is no strong empirical evidence which would unequivocally help to decide whether the absence of the gender congruency effects in some cases (bare noun naming, naming with gender-invariant modifiers) is due to the bypassing of the gender feature selection, or due to automatic gender node selection with (in this case absent) competition between the phonological forms.

phonological form anyway. Such a mechanism would be compatible e.g., with the Interactive Activation Model of Dell (1986).

Obviously, more empirical evidence is necessary to decide whether the observed congruency effect is located at the level of grammatical feature selection or at the level where phonological forms are selected. However, so far it seems very likely that no matter how this question will be finally answered, the fact that the occurrence of a congruency effect is contingent on whether or not there are multiple forms to choose from, supports cascaded processing. The evidence reported in this study contributes to the issue in a unique way, because it yields evidence that overt gender-markedness versus invariance affects gender processing, without involving manipulation with other noun features. All other studies concerned with this issue (Schiller and Caramazza 2002; Schriefers et al. 2002, 2005) compare naming of NPs in singular and plural in German and Dutch and thus have to formulate additional hypotheses about the interaction of the two noun features (i.e., gender and number). As the interaction of different grammatical features (e.g., number, case, gender, declensional class of nouns) during grammatical and phonological encoding is so far largely unresolved, it is important to search for evidence which can disentangle possibly confounded factors and expand the research also to other languages.

Another crucial question concerns the scope of the congruency effect: Does the competition for selection appear only when free-standing morphemes are involved, or does the selection of bound morphemes like suffixes follow the same principle? Our data from Czech clearly show that the competition also occurs between bound morphemes (or the corresponding gender feature), suggesting that the processes of gender production are identical in both cases. The presence of the congruency effect for ordinal + noun NPs is in accordance with the results reported by Schriefers (1993) and contrasts with those of Schiller and Caramazza (2003) and Costa et al. (2003). It is, however, notable that the congruency effect observed in our Experiments 2 and 3 (37 ms) with gender-marked hard ordinals is not only unreduced compared to Experiment 1 (26 ms), but numerically it is even larger. In this respect, the evidence for the competition between the gender-marked inflection reported in our paper is unique, because in the previous studies the congruency effect with bound morphemes was either not observed at all, or it was radically reduced (Schriefers 1993; Schriefers et al. 2005). The results of our study shed doubts on the claim that the evidence concerning the competition between bound morphemes is equivocal just because the gender marked inflections appear at the end of a word, which makes the measurement of the competition effects in their selection more difficult (Schiller and Caramazza 2003; Schriefers et al. 2005). It seems more likely that language specific properties (e.g., the degree of inflectionality, properties of the modifiers carrying the gender-marked inflections, the role of gender in the particular nominal system and its interaction with other features like number, declensional class etc.) also play a role, which makes the investigation of the congruency effect in other languages even more desirable. Our finding that the congruency effect can also be observed when gender surfaces as a gender marked inflection furthermore supports the assumption that free-standing morphemes like determiners or stems are stored in the same way as bound morphemes like inflectional suffixes (at least if we assume that the congruency effect reflects competition at the level of form selection). On this view, both inflectional suffixes and stems have an independent phonological status. From the absence of the congruency effect in the production of inflectional suffixes reported by Costa et al. (2003), the

authors conclude that inflected forms are produced by transforming a base form into other forms rather than by attaching affixes to a bare stem. In contrast, our results support the view that at least regular inflected forms are produced concatenatively, by simple addition of pieces of phonological information.

The last issue to be addressed here is the morphological status of the demonstrative pronoun *ten, ta, to* used in our Experiment 1: Is it a free-standing morpheme or is it a complex word consisting of a stem *t-* and an inflectional affix? If no congruency effect had been observed in Experiments 2 and 3, where the ordinals unambiguously consisted of a stem and an inflectional suffix, it would have been possible to argue that the effect in Experiment 1 occurred because a free-standing gender marked morpheme was involved in the production. Experiments 2 and 3, however, revealed that the congruency effect is present also in the absence of any gender marked free morpheme in Czech. Here, we want to argue that if the locus of the congruency effect is at the level of phonological forms, then the basic distinction concerning its scope (if necessary at all) should not be between bound and freestanding morphemes, but rather between stored and computed forms. If we accept this view, then it is irrelevant how the Czech demonstrative pronouns *ten, ta, to* are to be linguistically analyzed. The question then rather is whether they are stored as wholes or whether they are computed (by addition of an inflectional suffix to the stem). Our knowledge about the way regular and irregular forms are produced and about the role of frequency allows us to decide with great certainty that the forms of the pronoun are stored as whole units. First, its declension is highly irregular and experiments with regular and irregular forms clearly show that such forms are stored (Clahsen 1999). Second, the pronoun (and its forms) is very frequent and there is again ample evidence that very frequent forms are not computed each time they are needed anew, but are stored as whole units (Stemberger and MacWhinney 1986). If a distinction between congruency effects reflecting competition between stored units and competition between gender marked suffixes contained in computed forms is made, then the effect observed in our Experiment 1 would very likely fall into the first category, whereas the effect observed in Experiments 2 and 3 would fall into the second category. Whatever the interpretation of the congruency effect in Experiment 1 is, the results of Experiments 2 and 3 clearly show that the competition that leads to the congruency effect in Czech is not restricted to freestanding morphemes, or better, to stored units.

Acknowledgements The research reported in this paper was supported by a grant from the Deutsche Forschungsgemeinschaft (German Research Council, DFG Pe/357–9). We thank Jana Hrachovcová, Nataša Ukropcová and Michael Welt for their help with preparing and running the experiments as well as Merrill F. Garrett for his comments on an earlier draft of this manuscript.

Appendix:

Materials Used in Experiments 1–3

potok	(m)	brook	kotva	(f)	anchor
obraz	(m)	picture	ruka	(f)	hand
telefon	(m)	phone	trumpeta	(f)	trumpet
dopis	(m)	letter	ryba	(f)	fish
mravenec	(m)	ant	hlava	(f)	head
ananas	(m)	pineapple	kravata	(f)	long tie

prsten	(m)	ring	oko	(n)	eye
orel	(m)	eagle	auto	(n)	car
autobus	(m)	bus	letadlo	(n)	plane
pytel	(m)	sack	laso	(n)	lasso
bagr	(m)	digger (excavator)	pero	(n)	pen
koberec	(m)	carpet	koleno	(n)	knee
houba	(f)	mushroom	kolo	(n)	bike
kytara	(f)	guitar	okno	(n)	window
paprika	(f)	pepper	zrcadlo	(n)	mirror
vana	(f)	bathtub	sedlo	(n)	saddle
kniha	(f)	book	kladivo	(n)	hammer
cibule	(f)	onion	rameno	(n)	shoulder

References

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

Beringer, J. (1999). *Experimental run time system*. Frankfurt: BeriSoft Cooperation.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14*, 177–208.

Caramazza, A., Miozzo, M., Costa, A., Schiller, N., & Alario, F.-X. (2001). Lexical selection: A cross-language 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.

Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences, 22*, 991–1060.

Costa, A., Sebastian-Galles, N., Miozzo, M., & Caramazza, A. (1999). The gender congruity effect: Evidence from Spanish and Catalan. *Language and Cognitive Processes, 14*, 381–391.

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

Dell, G. S. (1986). A spreading-activation model of retrieval in sentence production. *Psychological Review, 93*, 231–241.

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. (1989). *Speaking: From intention to articulation*. Cambridge: MIT Press.

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.

Pechmann, T., & Zerbst, D. (2004). Syntactic constraints on lexical selection in language production. In T. Pechmann & C. Habel (Eds.),*Multidisciplinary approaches to language production* (pp. 279–303). Berlin: Mouton.

Schiller, N. O., & Caramazza, A. (2002). The selection of grammatical features in word production: The case of plural nouns in German. *Brain and Language, 81*, 342–357.

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.

Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 841–850.

Schriefers, H., & Jescheniak, J. D. (1999). Representation and processing of grammatical gender in language production: A review. *Journal of Psycholinguistic Research, 28*, 575–600.

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, and Cognition, 31*, 159–168.

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.

Stemberger, J. P., & MacWhinney, B. (1986). Frequency and the storage of regularly inflected forms. *Memory & Cognition, 14*, 17–26.

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