Journal of Psycholinguistic Research, Vol. 28, No. 6, 1999

Processing of Grammatical Gender in a Three-Gender System: Experimental Evidence from Russian

Tatiana Akhutina,[1] Andrei Kurgansky,[1] Maria Polinsky,[2] and Elizabeth Bates[2,3]

*Four experiments investigated the effect of grammatical gender on lexical access in Russian. Adjective–noun pairs were presented auditorily, using a cued-shadowing technique in which subjects must repeat the second word (the target noun), following adjectives that are either concordant or discordant with the noun's gender. Experiment 1 demonstrates gender priming with unambiguous adjectives and phonologically transparent masculine or feminine nouns. Experiment 2 examines priming for transparent nouns against a neutral baseline (possible only for feminines and neuters), revealing that priming is due primarily to inhibition from discordant gender. Experiment 3 demonstrates gender priming with phonologically opaque masculine and feminine nouns. Experiment 4 returns to transparent masculine and feminine nouns with a different kind of baseline, using three versions of a single word root (*prost—*simple, in the feminine adjectival form* prostaja, *masculine adjectival form* prostoj, *and the adverbial form* prosto*), and shows that gender can also facilitate lexical access, at least for feminine nouns. We conclude that Russian listeners can exploit gender agreement cues "on-line," helping them to predict the identity of an upcoming word.*

The Russian language offers interesting opportunities to investigate the functional contribution of gender information to lexical access, because of its three-gender system with complex interactions between gender (an inherent property of nouns) and case (a property of nouns determined by the structure of the sentence), as well as its substantial word order variation. At the same time, three-gender systems pose a substantial challenge to experimental

---

The research reported here was supported by NIH/NIDCD RO1-DC00216, and by a grant from the James McDonnell Foundation. Maria Polinsky's participation was also supported by the UCSD Academic Senate (#960940S).

[1] Moscow State University, Moscow, Russia.
[2] University of California, San Diego.
[3] Address all correspondence to Elizabeth Bates, Center for Research in Language, 0526, University of California, San Diego, La Jolla, California 92093-0526.

0090-6905/99/1100-0695$16.00/0

design. First, two-gender systems can treat gender assignment as a binary decision, with one gender serving as the default form for new words entering the language. Because three-gender systems are not binary, the assignment of "default gender" is a more complex problem. Second, most two-gender systems have a roughly equal balance in both type and token frequency across the two genders (although type frequencies can be greater for one gender). By contrast, there are often large disparities in type and token frequency within three-gender systems, with lower-type frequencies in one gender (e.g., the neuter gender in many Indo-European languages). Third, the existence of three or more genders can lead to great complications in formal marking, on the noun itself and on agreeing elements. Because so many different elements are involved, each with its own morphophonemic constraints, it is difficult for languages to exploit similarity and apply straightforward formal principles like "all masculines are zero-marked" or "all feminines end in -a."

Given these complications, it is difficult to design priming studies in a three-gender system. For example, Bates, Devescovi, Hernandez, and Pizzamiglio (1996) were able to exploit a factorial design to study the interaction between gender (masculine vs. feminine) and regularity (phonologically transparent vs. phonologically opaque) in their first study of gender priming in Italian. As we will explain below, the structure of the Russian language does not permit a full orthogonal comparison of gender and regularity/transparency. Despite these limitations, Russian provides many interesting challenges and opportunities for the study of gender effects and their interaction with other factors. The four experiments presented below are limited in nature, but they open the door for more extensive studies of this kind.

ON RUSSIAN GENDER

To understand the experiments below, a brief overview of the Russian gender system is required (for details, see Corbett, 1991, pp. 34–43; Comrie, Stone, & Polinsky, 1996, pp. 104–117, and further bibliography there). Russian has three genders: masculine, feminine, and neuter. Masculines constitute about 46% of the nominal lexicon, feminines 41%, and neuters 13%. Although the number of neuter nouns has been declining, neuter remains a robust category that accepts new words and borrowings. Grammatical gender is correlated with semantic gender in the usual way, but there are many exceptions as a function of both declensional type and phonological factors (especially noun ending). Russian has at least six noun cases (nominative, accusative, dative, genitive, instrumental, and locative—cf. Zaliznjak, 1967), organized into four main declensional types. Table I illustrates these four types using nouns in the singular (from the experiments below).

Table I. Main Declensional Types in Russian (Singular Only)

	Type			
Case	I: "city"	II: "water"	III: "mud, dirt"	IV: "milk"
Nominative	gorod	voda	grjaz'	moloko
Accusative	gorod	vodu	grjaz'	moloko
Dative	gorodu	vode	grjazi	moloku
Genitive	goroda	vody	grjazi	moloka
Instrumental	gorodom	vodoj	grjaz'ju	molokom
Locative	gorode	vode	grjazi	moloke

The endings of some nouns in the nominative singular act as clear indicators of their gender. For our purposes, the most important correlations are the following: (a) nouns ending in a hard consonant are masculine, (b) nouns ending in -a are feminine if inanimate, (c) nouns ending in -o are neuter, (d) nouns ending in a palatalized consonant are either masculine or feminine.

There are many exceptions to these regularities that reduce their predictive value (see Comrie *et al.* 1996, pp. 112–117; Corbett, 1982); we have excluded most of them (variable-gender nouns, indeclinables, acronyms, pluralia tantum, diminutives) from our experiments and will not discuss them further. We have also restricted ourselves to singular nouns in the nominative case, which is the citation form.

With the exclusions just described, nouns in categories a–c above can be assigned gender unambiguously on the basis of their nominative singular ending, which means that they are phonologically transparent (Bates *et al.*, 1996). Nouns of category d have ambiguous endings, and will be referred to as phonologically opaque. The distinction between transparent and opaque nouns will be examined in the experiments below, but cannot be treated as a factor in a single design because it is not possible to match transparency and opacity along the various stimulus dimensions that we will require (e.g., frequency, number of syllables, number of phonemes, length of waveform). Gender priming of transparent nouns will be investigated in experiments 1, 2, and 4. Priming of opaque nouns will be examined experiment 3.

Russian gender agreement is manifested on adjectives, participles, past tense verbs, and numerals, illustrated in Table II. Table II shows that the agreement trigger can either follow or precede its target and that agreement is signaled by modifier endings. The pronunciation of these endings is dependent on the stress of a given adjective, and since our experiments involved auditory stimuli, we need to introduce two subtypes of adjectives.

Russian lexical items have mobile stress, i.e., stress can occur on any syllable in the word. In the standard pronunciation of Russian, adjectives that are

Table II. Gender Agreement in Russian

	Noun–adjective agreement	Noun–participle agreement	Noun–verb agreement	Noun–numeral agreement
Masculine	*plox-ój dom* bad house (nominative)	*zabróçsenn-yj dom* abandoned house (nominative)	*dom propal-∅* The house disappeared.	*ob-a doma* both houses
Feminine	*plox-ája kvartira* bad apartment (nominative)	*zabróçsenn-aja kvartira* abandoned apartment (nominative)	*kvartira propal-a* The apartment disappeared.	*ob-e kvartiry* both apartments
Neuter	*plox-óje çziliçsçce* bad dwelling (nominative)	*zabróçsenn-óje çziliçsçce* abandoned dwelling (nominative)	*çziliçsçce propal-o* The dwelling disappeared.	*ob-a çziliçsçca* both dwellings

stressed on the stem rather than the ending do not distinguish between feminine and neuter, e.g.,

masculine	glavnyj [glávnyj]
feminine	glavnaja [glávn™j™]
neuter	glavnoje [glávn™j™]

Adjectives that do not distinguish between feminine and neuter will be referred to below as ambiguous. Adjectives that always distinguish feminine and neuter, by virtue of stress on the ending, will be referred to as unambiguous. (Note that both ambiguous and unambiguous adjectives distinguish the masculine.) We will capitalize on this distinction in experiment 2.

The four experiments reported below represent a first step, conducted entirely within the auditory modality using the "cued shadowing" technique (Bates & Liu, 1996), in which subjects are asked to repeat auditory target words signaled by a voice change. To maximize comparability with existing studies of gender priming, we will focus on noun recognition in modifier-noun pairs similar to the stimuli that have been used in Italian, French, and German (Bates *et al.,* 1996; Grosjean, Dommergues, Carnu, Guillelmon, & Besson, 1994; Jacobsen, this volume).

GENERAL METHOD

Participants

Eighty young adults participated in one of the four experiments (20 per experiment). All were native speakers of Russian, college students ranging in age from 18 to 23, recruited from universities in the Moscow area. All were paid for their participation; they were naive about the goals of the experiment and reported no previous participation in any psycholinguistic study.

Materials

Stimuli for all experiments were auditory noun phrases consisting of a prenominal adjective and a noun, audio-recorded by a female native speaker of Russian at a natural tempo; special care was taken to ensure that the spoken items were not hyperarticulated. Words were chosen from three different word frequency dictionaries of Russian, which all yielded comparable results (Abakumova, Basis, & Badrieva, 1968; Steinfeldt, 1963; Zasorina, 1977). Excluding the top 500 most frequent words as well as words of low frequency, the resulting stimuli belong to the top quarter of the Russian noun vocabulary in terms of written and spoken frequency. We also excluded long polysyllabic words, proper names, special technical terms, acronyms, recent loan words, and words with leading voiceless consonants (see also Bates *et al.*, 1996). The final set used in one or more of the four experiments described below includes 112 nouns and 60 adjectives, all expressed in the singular form in nominative case, which is also the citation form.

The adjectives were audio-recorded separately from the nouns, in the context of a carrier phrase (target adjective plus dummy noun) to ensure a naturalistic rising intonation contour for the modifying adjective. Similarly, the nouns were recorded separately from the adjectives, in the context of a carrier phrase (dummy adjective plus target noun), to ensure a natural falling intonation at the noun's end. Thus, when adjectives and nouns were taken randomly from a target word set and combined into an adjective–noun phrase, the normal rise–fall contour for Russian auditory noun phrases was achieved. Stimuli were digitized and then trimmed out from the original continuous recording using the MacIntosh Sound Edit 16 system. The actual waveform duration (in milliseconds) of resulting separate words were then estimated via the PsyScope Experimental Shell (Cohen, MacWhinney, Flatt, & Provost, 1993). In each of the experiments reported below, the adjectives and nouns were arranged across conditions to achieve the best possible balance across the following dimensions (all known to affect reaction times): (1) actual waveform duration in milliseconds, (2) length in syllables, (3) length in phonemes, and (4) frequency of use. As described within each experiment below, adjective–noun pairs were compiled on-line for each participant by the PsyScope Experimental Shell, drawn quasirandomly from the relevant adjective and noun lists within the design constraints for that experiment [similar to the procedure described for Italian in Bates *et al.* (1996)]. Within each experiment, nouns were repeated within subjects, across conditions, but the individual randomization assured that order of presentation was variable and that experimental effects were not confounded with specific adjective–noun combinations (and hence were not influenced by the semantic or pragmatic peculiarities of any single adjective-noun phrase, e.g., *bol'çsaja sobaka* "big dog" vs. *bol'çsoe jabloko* "big apple"). Thus, subject A might see noun 1 first in the concordant condition, with adjective X, and then in the discordant condition, with adjective Y; subject B would see noun 1 first in the discordant condition, with adjective Q, then in the concordant, with adjective Z; and so forth. When collapsed across all 20 subjects within a given experiment, a global balance in stimulus dimensions such as frequency and length was maintained across experimental conditions.

Task and Procedure

In all four experiments, we used a variant of the cued-shadowing procedure (Bates & Liu, 1996), in which the participant is instructed to repeat the second word (always the noun) in each word pair. Adjective–noun pairs were presented auditorily via loudspeakers, at a comfortable amplitude. Within each pair, noun targets were presented immediately after the offset of the adjective prime (stimulus onset asynchrony = 0 ms). Subjects were allowed to respond until a response was registered. There were no situations in which subjects failed to respond; overly long response times were trimmed, as outlined below. Each trial was followed by a fixed intertrial interval of 1000 ms.

Participants were tested individually in a quiet room. They were told that they would be presented with a series of pairs of words, one pair at a time, and asked to repeat the second word in each pair as soon as they recognized that word, as fast as they were able without making a mistake. They were instructed to avoid early vocalizations (e.g., "Uhhhh . . . table"), beginning their vocalization only when they were sure that they had identified the word. Distance to the microphone (placed on a table in front of the participant) was set individually prior to the experiment, adjusted to assure reliable triggering of the voice key in the Carnegie Mellon Button Box, which was used to collect word onset latencies. This was established in a short practice (10 trials), which also served to adapt participants to the task. The stimuli used for the practice trials did not occur in the experimental trials.

Data Analysis

Reaction times (RTs) were measured from the onset of the target word to the onset of the subject's response, registered by the Carnegie Mellon Button Box and collected into the PsyScope data file. Trials on which errors occurred were excluded from reaction time analyses, and mean RTs were based on the average for that cell with the errors removed. All reaction times greater than 1500 ms were deemed outliers (based on previous studies) and were also excluded from further analyses. In fact, outlying RTs, errors, and failures to respond proved to be quite rare (see below). The individual randomization of adjectives and nouns eliminates the need for analyses of variance over items, so that all statistical analyses were conducted only over participants. In all of the experiments below, a multivariate analysis of variance approach was applied (McCall & Appelbaum, 1973).

EXPERIMENT 1

Experiment 1 was designed to parallel as closely as possible the method and materials used by Bates *et al.* (1996) for Italian. Because an orthogonal gender × regularity design is impossible in Russian (see above), experiment 1 was restricted to transparent masculine and feminine nouns, with gender-concordant or gender-discordant adjectives. No baseline was used, because there is no single adjective form that is equally compatible with both masculine and feminine nouns. Hence experiment 1 can determine whether adjective gender can prime recognition of nouns in Russian, but it cannot distinguish between facilitative and inhibitory priming.

Method

Materials and Design. Each of the 20 participants in experiment 1 was presented with 120 individually randomized adjective–noun pairs, constructed from a list of 52 transparent disyllabic nouns (26 masculine and 26 feminine nouns; no neuters were employed in this experiment) and 15 adjectives with stressed endings. Each adjective was recorded twice, once in the masculine and once in the feminine form. All the feminine nouns ended in -a (declensional type II, see Table I), and masculine nouns had nonpalatalized consonantal endings (declensional type I, see Table I). The respective masculine and feminine nouns were balanced for waveform duration [502.9 ms for feminine vs. 507.8 for masculine, $F(1, 50) < 1.0$] and frequency of use [26.9 for feminine vs. 30.1 for masculine, $F(1, 50) < 1.0$]. However, it proved impossible to achieve a satisfactory balance for word length in phonemes if/when length of the total waveform is controlled, due to the nature of transparent masculine vs. feminine markings in Russian. The mean length in phonemes for feminines was 4.54 ($SD = 0.58$), vs. 4.88 for masculines ($SD = 0.58$), a significant difference [$F(1, 50) = 4.55$, $p < .04$].

The overall design for experiment 1 was 2 (masculine vs. feminine noun) × 2 (gender-concordant vs. gender-discordant adjective), with 30 items in each of the four cells of the design. Adjectives and nouns were chosen at random by the PsyScope shell to produce unique pairings for each subject. Note that these 26 feminine and 26 masculine nouns do not divide evenly into 120 trials. It was not possible to find 30 feminine and 30 masculine nouns without disturbing the desired balance in frequency and length. Hence, through the individual randomization process (see General Method), some nouns were repeated more than others across conditions. However, the randomization process ensured that stimulus dimensions were balanced at a global level when data were collapsed over subjects.

Results and Discussion

Errors. Subjects made a minor number of errors on the task. The average error rate (including failures to respond, wrong word repeated, or RTs exceeding 1500 ms) was 0.75 per 120 trials (ranging from 0 to 2 per session for any individual participant). Hence the error data were not subjected to further analyses.

Reaction Times. Collapsed across all cells of the design, the mean RT was 693 ms (SD = 118), over items with a mean word duration of 505 ms (SD = 91 ms). Hence, when measured from the end of the word, RTs averaged 188 ms, which is compatible with values reported by Bates *et al.* (1996) for Italian. It thus seems fair to conclude that the cued-shadowing method (applied for the first time to Russian) yields reaction times well within the range that we would need in order to draw conclusions about "on-line" priming effects.

Results of the 2 (masculine vs. feminine target) × 2 (concordant vs. discordant adjective) MANOVA yielded significant main effects of noun gender [$F(1, 19) = 11.28$, $p < .003$] and concordance [$F(1, 29) = 36.8$, $p < .0005$], as well as a significant interaction [$F(1, 19) = 7.02$, $p < .016$]. Cell means are summarized in Table III.

The strong effect of concordance reflects an overall average of 33 ms priming, collapsed across feminine and masculine nouns. *Post hoc* analyses showed significant priming for both feminine nouns [$t(19) = 5.79$ $p < .0005$] and masculine nouns [$t(19) = 2.97$, $p < .008$], although the size of the priming effect was greater for feminines (47 ms) than masculines (20 ms).

Table III. Cell Means for Experiment 1 (Standard Deviations in Parentheses)

	Gender class of nouns		
Adjectives	Feminine	Masculine	Total
Concordant	678 ms (127)	674 ms (105)	676 ms (116)
Discordant	725 ms (123)	694 ms (117)	709 ms (120)
Priming	47 ms[a]	20 ms[a]	33 ms[a]

[a] $p < .05$.

The main effect of target gender indicates that, despite careful counterbalancing between masculine and feminine lists, the masculine nouns were shadowed (on average) 17 ms faster than feminine nouns (684 ms vs. 701 ms). Recall that masculine nouns were actually longer in mean number of phonemes, although the two genders were balanced for length of the actual waveform. Hence the fact that RTs were faster to masculine nouns cannot be an artifact of length (which would have predicted results in the opposite direction).

As Table III shows, the significant interaction between gender and concordance comes primarily from the discordant feminine condition. Additional *post hoc* analyses (Student's t test, two-tailed) showed no significant difference (4 ms) between feminine and masculine nouns in the gender-concordant condition [$F(19) = 0.48$, $p = 0.634$], but a highly significant difference of 31 ms in the discordant condition [$t(19) = 5.00$, $p < .0005$]. Hence the processing advantage observed for masculine nouns does not mean that masculines are processed more effectively overall (in the concordant condition, masculines are only 4 ms faster than feminines); rather, the main effect of gender is a by-product of the fact that discordant feminines are exceptionally difficult to process.

EXPERIMENT 2

Results of experiment 1 confirm that gender-marked adjectives can prime reaction times to Russian target nouns. However, in the absence of a neutral baseline (impossible for masculine nouns), we cannot determine whether these priming effects are facilitative or inhibitory. Experiment 2 addresses the baseline issue by examining gender priming for feminine vs. neuter nouns, exploiting the fact that gender-ambiguous adjectives are possible when masculine nouns are excluded.

Method

Materials and Design. A total of 30 adjectives were chosen for this experiment from the list described in the General Method, all different from those used in experiment 1. Each adjective was recorded twice: once in the feminine form, once in the neuter form. Among these 30 adjectives, 15 (half) were unambiguous with stressed and clearly audible endings. The other 15 adjectives are phonologically ambiguous—since they have unstressed endings, the difference between feminine and neuter forms cannot be detected when the words are presented auditorily.

Thirty nouns were selected; none had been used in experiment 1. With one exception (the four-syllable neuter noun *odejalo*—blanket), all words were three syllables in length. One half (15) were feminine, of declensional type II, and the other half (15) were neuter nouns of declensional type IV (see Table I). Recording in carrier phrases, digitization and measurement followed the procedures outlined in the General Method. Because the set of feminine and neuter nouns that meets the selection criteria for this experiment is relatively small, it was not possible to balance target nouns across all of the desired stimulus dimensions. The mean waveform duration for feminines was 752 (SD = 124 ms), which is approximately 130 ms greater than the mean waveform duration for neuter nouns (620 ms, SD = 94). The difference was statistically significant [$F(1, 28) = 10.47$ $p < .004$]. On the other hand, the feminine nouns were more frequent than the neuters (35.2, SD = 28, vs. 8.87, SD = 6). This difference in frequency was also statistically significant [$F(1, 28) = 12.89$, $p < .0002$]. A satisfactory balance was achieved in both word length in syllables and word length in phonemes ($F < 1.0$). These differences in length and frequency must be kept in mind in the interpretation of any differences that we find between feminine and neuter gender.

Stimuli were randomly assigned to conditions for each of the 20 participants by the PsyScope Experimental shell, assuring that any priming effects that we might obtain were not contaminated by specific lexical combinations.

Results and Discussion

Errors. The error rate in experiment 2 averaged 0.65 per 120 trials (with a range of 0–3), including all wrong responses, failures to respond, or false starts. Because these error rates are so low, they were not subjected to further analysis.

Reaction Times. The overall mean RT was 708 ms (SD = 132), measured from the onset of each target word. Because the mean waveform duration for individual words was 686 ms (SD = 128), this means that the average RT was only 22 ms when measured from the end of each word, further evidence that the cued-shadowing task yields truly "on-line" information about lexical access in Russian.

Experiment 2 lends itself to two different statistical designs: (1) a 2 (feminine vs. neuter nouns) × 2 (ambiguous vs. unambiguous adjectives) × 2 (concordant vs. discordant adjectives) design, or (2) a 2 (feminine vs. neuter nouns) × 3 (concordant, discordant, ambiguous adjectives) design. The physical structure of the stimulus set corresponds to the first design, but the psychological structure (from the point of view of the listener) corresponds to the second design, because ambiguous adjectives cannot be experienced as either concordant or discordant when they precede feminine or neuter nouns (such adjectives would be clearly discordant if masculine nouns had been used). We decided to conduct the analyses both ways: Table IV summarizes results from the 2 × 2 × 2 physical structure analysis, and Table V summarizes results from the 2 × 3 psychological structure analysis (treating all ambiguous adjectives as a neutral baseline).

In the 2 × 2 × 2 MANOVA, there was a main effect of concordance [$F(1, 29) = 9.91$, $p < .005$] and a main effect of unambiguous adjectives [$F(1, 19) = 41.53$, $p < .0001$]. The main effect of noun gender approached significance [$F(1, 19) = 3.46$, $p < .08$]. There was also a significant interaction between concordance and adjective transparency (unambiguous adjectives) [$F(1, 19) = 18.15$, $p < .0005$]. None of the other interactions reached significance ($F < 1.0$). As Table IV shows, there was significant priming (i.e., significant differences between concordant and discordant adjectives) in the unambiguous conditions, averaging 24 ms for feminine nouns preceded by unambiguous adjectives [$t(19) = 3.47$, $p < .003$], and 34 ms for neuter nouns preceded by unambiguous adjectives [$t(19) = 4.77$, $p < .0001$]. There was no significant difference between concordant and discordant conditions when ambiguous adjectives were used (−4 for feminine nouns, and −3 for neuter nouns). Of course this is not surprising, because (as noted above) listeners cannot hear the difference between concordant and discordant conditions when ambiguous adjectives with unstressed endings are used. Hence it seems more appropriate to analyze experiment 2

Table IV. Cell Means for Experiment 2 "Physical" 2 × 2 × 2 Design (Standard Deviations in Parentheses)

	Gender class of nouns		
Adjectives	Feminine	Neuter	Total
Concordant			
Unambiguous	707 ms	694 ms	700 ms
	(127)	(129)	(128)
Ambiguous[a]	710 ms	698 ms	704 ms
	(138)	(125)	(131)
Discordant			
Unambiguous	731 ms	728 ms	729 ms
	(137)	(141)	(139)
Ambiguous[a]	706 ms	695 ms	700 ms
	(149)	(120)	(134)
Priming			
Unambiguous	24 ms[b]	34 ms[b]	29 ms[b]
Ambiguous[a]	−4 ms	−3 ms	−3.5 ms

[a] Difference between concordant and discordant is not audible on ambiguous adjectives.
[b] Statistically reliable priming effect ($p < .05$).

Table V. Cell Means for Experiment 2 "Psychological" 2 × 3 Design (Standard Deviations in Parentheses)

	Gender class of nouns		
Adjectives	Feminine	Neuter	Total
Concordant	707 ms	694 ms	700 ms
	(127)	(129)	(128)
Neuter	708 ms	696 ms	702 ms
	(143)	(121)	(132)
Discordant	731 ms	728 ms	729 ms
	(137)	(141)	(139)
Total priming	24 ms[a]	34 ms[a]	29 ms[a]
Facilitation	1 ms	2 ms	2 ms
Inhibition	23 ms[a]	32 ms[a]	27 ms[a]

[a] Statistically significant priming score ($p < .05$).

using the psychologically valid 2 × 3 design (feminine/neuter nouns vs. concordant/discordant/ambiguous adjectives).

The second MANOVA yielded a significant main effect of adjective priming [$F(2, 18) = 26.77$, $p < .0005$]. The main effect of noun gender approached significance [$F(1, 19) = 3.30$, $p < .09$], reflecting slightly faster responses overall for neuter words (which, as we noted earlier, also had significantly shorter waveforms). The interaction between noun gender and adjective concordance was not significant [$F(2, 18) < 1.0$]. As can be clearly seen from Table V, reaction times pattern in similar ways within both genders. In particular, the mean RTs are virtually identical within the respective concordant and neutral conditions for each gender, while RTs in the discordant condition are significantly longer. Pairwise comparisons confirmed that there were no differences between concordant and neutral conditions ($t < 1.0$) and therefore no evidence for gender facilitation, but there was evidence for robust inhibitory priming (discordant vs. neutral), averaging 23 ms for feminine nouns [$t(19) = 4.69$, $p < .0005$] and 31 ms for neuter nouns [$t(19) = 4.48$, $p < .0005$]. Hence, using ambiguous adjectives as a neutral baseline, we obtain significant priming but no evidence for gender facilitation in this experiment.

EXPERIMENT 3

Experiment 3 was designed to determine whether gender priming would also occur for opaque nouns, using masculine and feminine nouns of declensional types I and III (see Table 1).

Method

Materials and Design. Thirty adjectives were selected from the list described in the General Method. Each adjective was recorded twice: once in feminine form, once in masculine, resulting in a total of 60 adjectives. A few of the adjectives used in experiment 1 were also used here. The reader should remember that adjectives always unambiguously distinguish masculine from feminine.

We selected 30 opaque nouns (all ending in a palatalized consonant) from the list described in the General Method. One half (15) were feminine (belonging to declensional type III, Table I) and the other half (15) were masculine (declensional type I). All were either mono- or disyllabic. Among the feminine nouns, 5 were monosyllabic and 10 were disyllabic. Approximately the same ratio was preserved among the masculine nouns: 4 monosyllabic and 11 disyllabic. The two pools of nouns did not differ significantly in word duration (558 ms, $SD = 96$ for feminine; 562, $SD = 93$ for masculine), frequency of use (7.70 for feminine vs. 3.66 for masculine), word length in phonemes (4.67, $SD = 0.82$ for feminine, 4.73, $SD = 0.88$ for masculine), or length in syllables (1.67, $SD = 0.49$ for feminine, 1.73, $SD = 0.46$ for masculine). $F(1,28) < 1.0$ in all analyses.

Results and Discussion

Errors. Again, as in the two previous experiments, errors were very rare (including failures to respond, wrong responses, or false starts), averaging 0.8 per 120 trials (ranging from 0–5 per session for any individual subject). No further analyses of errors were performed.

Reaction Times. Overall RTs averaged 724 ms ($SD = 114$), which is somewhat longer than the mean RTs observed in the first two experiments (691 ms and 709 ms, respectively). Because noun waveform durations averaged 560 ms ($SD = 93$), this means that the average RT measured from the end of each target word was 164 ms, still well within the range that we should expect for an "on-line" measure of lexical access.

Reaction times were subjected to a 2 (masculine vs. feminine noun) × 2 (concordant vs. discordant adjective) MANOVA. There were significant main effects of concordance [$F(1, 19) = 13.86$, $p < .001$] and noun gender [$F(1, 19) = 4.53$, $p < .05$], but the interaction was not significant [$F(1, 19) < 1.0$]. Cell means are summarized in Table VI, which shows that RTs to masculine nouns (719 ms) were faster than RTs to feminine nouns (730 ms). This may reflect the "default" status of masculine nouns (which are also higher in type frequency), but it might also reflect the fact that masculine nouns were slightly (although not significantly) shorter when measured in

Table VI. Cell Means for Experiment 3 (Standard Deviations in Parentheses)

	Gender class of nouns		
Adjectives	Feminine	Masculine	Total
Concordant	722 ms	711 ms	716 ms
	(114)	(112)	(113)
Discordant	738 ms	727 ms	732 ms
	(116)	(117)	(116)
Priming	16 ms[a]	16 ms[a]	16 ms[a]

[a] Statistically significant priming effect ($p < .05$).

mean number of phonemes (feminine = 4.93, masculine = 4.73). However, the priming effect was similar in magnitude (16 ms) within both genders and reached significance for each gender in *post hoc* tests [$t(19) = 3.69$, $p < .002$ for feminines; $t(19) = 2.20$, $p < .04$ for masculines]. These priming effects are smaller than those obtained with transparent nouns in experiments 1 and 2 (33 ms in experiment 1, 29 ms in experiment 2, collapsed across genders), a difference that may reflect the additional information provided by gender-transparent noun endings. Nevertheless, we may conclude that significant gender priming occurs whether or not formal gender markers are available on the target noun.

EXPERIMENT 4

The three experiments described so far testify to the robust nature of gender priming in Russian, for all three genders, for nouns with or without gender-transparent endings. However, the nature and direction of this gender effect is much less clear. With the transparent masculine and feminine nouns used in experiment 1, we found a significant interaction between noun gender and concordance, reflecting substantially slower times for feminine nouns in the discordant condition. There was no such interaction with the transparent feminine and neuter nouns used in experiment 2, but comparisons with the gender-ambiguous baseline condition led to the conclusion that the gender-priming effect is purely inhibitory in nature. The absence of evidence for a facilitative effect is troublesome: If gender agreement cues do play a role in word recognition, as a number of investigators have proposed (Bates *et al.,* 1996), then we should be able to find evidence for facilitation relative to some kind of gender-ambiguous baseline. If the priming effects that we have observed here are due entirely to the disruptive effects of gender mismatch (a condition rarely encountered in the real world), then the ecological validity of gender priming is not at all obvious. Can we uncover gender facilitation in Russian?

The key issue here is the choice of a neutral baseline. The adjective baseline used in experiment 2 could only be applied to the distinction between feminine and neuter nouns, and therefore does not generalize to the masculine gender, the largest and most important gender class in Russian, nor to the contrast between masculine and feminine (which, together, comprise about 85% of all Russian nouns). Experiment 4 was designed to approach the baseline issue from a different point of view, permitting an assessment of the facilitative and inhibitory contributions of gender-marked modifiers to recognition of masculine and feminine nouns. In contrast with experiments 1–3, which drew from a relatively large list of gender-marked adjectives, only one adjective was used in experiment 4: the word for "simple," which is *prostoj* in the masculine form and *prostaja* in the feminine form. There is no gender-ambiguous version of this adjective; indeed, there are no gender-ambiguous adjectives of any kind when masculine nouns are used. However, we were able to use a phonologically close and highly frequent adverb/particle, *prosto* "simply," similar in frequency and meaning to the English word *just.* We predicted that RTs to nouns following *prosto* would be significantly slower than RTs following a gender-concordant version of the adjective (providing evidence for facilitation), while RTs following this adverb would be significantly faster than responses following a gender-discordant version of the adjective (providing evidence for inhibition).

Method

Materials and Design. As noted above, only three prime words were used in this experiment: the masculine adjective *prostoj* "simple," the feminine adjective *prostaja* "simple," and the gender-neutral adverb/particle *prosto* "simply." These were combined with the same 26 feminine and 26 masculine nouns adopted in experiment 1, all transparent disyllabic forms. The design was 2 (masculine vs. feminine nouns) × 3 (concordant adjective, discordant adjective, neutral adverb). There were 156 trials, and each of the 52 nouns appeared once in each adjective condition. Within these design constraints, nouns were randomly assigned to conditions by the PsyScope shell, for each individual subject (see General Method and experiment 1).

Results and Discussion

Errors. As in experiments 1–3, errors were very rare (including wrong responses, failures to respond, and false starts), averaging 1.3 per 156 words (with a range of 0–9 for individual subjects). No further statistical analyses were performed.

Reaction Times. The overall mean RT obtained in this experiment was 611 ms (SD = 105). Because mean noun waveform duration was 505 ms (SD = 91), this means that RTs measured from the end of the word averaged 106 ms, well within the limits required for an "on-line" study of lexical access. Note that the word onset RTs in this experiment are more than 80 ms faster than those obtained with the same nouns in experiment 1 (611 ms for experiment 4 vs. 693 ms for experiment 1). To some extent, this may reflect the fact that experiment 4 contains more trials and more repetitions of individual words. It may also reflect random variations in baseline response time that often occur when results are compared across different groups of subjects. However, a more interesting explanation may lie in the fact that experiment 4 employed only three variants of the same root word as concordant, discordant, or neutral primes, compared with the large and variable set of adjective primes used in the other experiments. Although psycholinguists usually control carefully for the stimulus characteristics of the target words employed in priming experiments, less attention is paid to the stimulus characteristics of the prime. However, Liu, Bates, Powell, and Wulleck (1997) have shown that prime frequency is a significant contributor to reaction times in two-word priming studies. In particular, RTs are faster following a prime that is higher in frequency, after many other characteristics of both the prime and target are controlled. Liu *et al.* (1997) suggest that high-frequency primes can be processed more quickly, which means that the processing of a subsequent target word can begin more quickly as well (particularly important in designs like theirs and ours, where the interval between prime and target is very short). Along the same lines, the 80-ms RT advantage that we observed in experiment 4 may reflect the fact that the same adjective primes were repeated over and over, requiring minimal processing of the prime itself and thus permitting earlier processing (and faster recognition) of the target word.

A 2 (masculine vs. feminine noun) × 3 (concordant, discordant, baseline) MANOVA was conducted on the reaction time data, measured from target word onset. There were significant main effects of concordance [$F(2, 18) = 11.46, p < .001$] and noun gender [$F(1, 19) = 46.08, p < .0005$], but the interaction was not significant [$F(2, 18) = 1.36, p > .28$]. Cell means are summarized in Table VII, to facilitate comparison with results for the previous experiments.

To explore the overall effect of concordance, planned comparisons (pairwise Student's t tests, two-tailed) were conducted. The overall priming effect was 25 ms (discordant–concordant). The facilitative component (neutral–concordant) was 10.5 ms, and the inhibitory component (discordant–neutral) was 14.5 ms. So it seems that the gender-irrelevant adverb/particle *prosto* did operate much as expected, as a neutral baseline midway between concordant vs. discordant adjectives. *Post hoc* t tests (two-tailed) showed that

Table VII. Cell Means for Experiment 4 (Standard Deviations in Parentheses)

	Gender class of nouns		
Adjectives	Feminine	Masculine	Total
Concordant	607 ms (110)	592 ms (103)	599 ms (107)
Neutral	625 ms (108)	596 ms (105)	610 ms (106)
Discordant	636 ms (108)	613 ms (103)	624 ms (105)
Priming	29 ms[a]	21 ms[a]	25 ms[a]
Facilitation	18 ms[a]	4 ms	11 ms
Inhibition	11 ms	17 ms[a]	14 ms[a]

[a] *Statistically significant priming effect ($p < .05$).*

both the facilitative effect [concordant vs. neutral, $t(19) = 3.09$, $p = .006$] and the inhibitory effect [discordant vs. neutral, $t(19) = 2.66$, $p = .015$] were significant.

Although the interaction of gender by concordance was not significant, we conducted planned comparisons separately for masculine and feminine nouns, to facilitate comparison across experiments. The facilitative component of gender priming was significant for feminine nouns [17 ms, $t(19) = 3.39$, $p < .003$] but not for masculine nouns [4 ms, $t(19) = 0.69$, $p > .50$]. The inhibitory component was significant for masculine nouns [17 ms, $t(19) = 2.4$, $p < .025$] but not for feminine nouns [12 ms, $t(19) = 1.87$, $p < .076$], at least not by a two-tailed test (a less conservative one-tailed test might be warranted in this case, because we did indeed predict that discordant adjectives would be slower than the neutral baseline, for both genders, in which case the inhibitory component would also be significant for feminine nouns, at $p < .038$). Although all of these priming effects are relatively small, experiment 4 shows that gender-marked adjectives can facilitate lexical access in Russian.

GENERAL DISCUSSION

These four experiments provide further evidence of gender priming in a three-gender language with a complicated formal system of gender assignment and gender agreement. Although it is not possible in Russian to assess priming and morphophonemic transparency or regularity together in a single design, separate experiments showed that priming effects are obtained for both transparent nouns (with endings that are clearly and regularly marked for gender) and opaque nouns, which do not "wear their gender on their sleeve" (Bates *et al.,* 1996). The magnitude of these priming effects is small (ranging from 16 ms for opaque masculines and feminines in experiment 3, to 33 ms for transparent masculines and feminines in experiment 1), but no smaller than the effects that have been reported for gender priming in other languages using modifier-noun pairs in the auditory modality (e.g., Bates *et al.,* 1996, for Italian; Grosjean *et al.,* 1994, for French). Hence we may conclude that gender priming is a robust phenomenon in Russian, similar to the effects that have been observed with other two- and three-gender systems (e.g., for German, Hillert, & Bates, 1996; Jacobsen, this volume).

The existence and magnitude of these effects are clear, but their nature and direction are less obvious. Russian does not have a set of ambiguous adjectives that can be used as a neutral baseline for all the three genders, so experiment 2 used gender-ambiguous adjectives with feminine and neuter nouns only. Results suggested that the priming effect was due entirely to the inhibitory effects of a gender-discordant adjective (treating the ambiguous adjectives as a neutral baseline). However, we cannot generalize this finding to masculine nouns (the largest noun class), nor does it tell us anything about the disambiguation of masculines from feminines (with masculines and feminines together comprising roughly 90% of the nouns in Russian). Experiment 4 took up the same issue with a very different baseline, using three versions of a single word root (*prost*—simple, in the feminine adjectival form *prostaja,* masculine adjectival form *prostoj* and in the adverbial form *prosto*). These three primes were used with phonologically transparent masculine and feminine nouns (the same ones adopted in experiment 1, which had no baseline). Significant overall priming effects were obtained, and the neutral adverb *prosto* fell (as expected) right in between the concordant and discordant adjective conditions. Collapsed across genders, both the inhibitory and the facilitative component of gender priming reached significance. When the two genders were analyzed separately (a planned comparison that we conducted despite the absence of a gender × concordance interaction), the facilitative component only reached significance for feminine nouns. Although these facilitative effects are small, they are comparable in size to the significant effects of gender facilitation that have been observed in other languages, especially in the auditory modality.

The neutral baseline issue is not a simple one in any language, and it is even more complex in Russian, with three genders and a number of interacting formal cues. However, there is enough evidence here and in other languages to support the idea that gender agreement cues can be exploited "on-line," helping the listener to predict the identity of an upcoming word.

We would expect our current results for gender priming to vary with case and word order. Experiments designed to investigate these factors are currently underway.

REFERENCES

Abakumova, A., Basis, N., & Badrieva, L. (1968). *2380 naibolee upotrebitel'nyx slov russkoj razgovornoj reçci* [2380 most frequent words of spoken Russian]. Moscow: Universitet druzhby narodov im. P. Lumumby.

Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Perception & Psychophysics, 58(7),* 992–1004.

Bates, E., & Liu, H. (1996). Cued shadowing. In F. Grosjean & U. Frauenfelder (Eds.), *A guide to spoken word recognition paradigms* (pp. 577–581). Hove, England: Psychology Press.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments & Computers, 25,* 257–271.

Comrie, B., Stone, G., & Polinsky, M. (1996). *The Russian language in the twentieth century.* Oxford: Oxford University Press.

Corbett, G. G. (1982). Gender in Russian: An account of gender specification and its relationship to declension. *Russian Linguistics, 6,* 197–232.

Corbett, G. G. (1991). *Gender.* Cambridge: Cambridge University Press.

Grosjean, F., Dommergues, J.-Y., Cornu, E., Guillelmon, D., & Besson, C. (1994). The gender-marking effect in spoken word recognition. *Perception & Psychophysics, 56 (5),* 590–598.

Hillert, D., & Bates, E. (1996). *Morphological constraints on lexical access: Gender priming in German* (Tech. Rep. No. 9601). La Jolla: University of California, San Diego, Center for Research in Language.

Liu, H., Bates, E., Powell, T., & Wulfeck, B. (1997). Single-word shadowing and the study of lexical access. *Applied Psycholinguistics, 18(2),* 157–180.

McCall, R. B., & Appelbaum, M. I. (1973). Bias in the analysis of repeated measure designs: Some alternative approaches. *Child Development, 44 (3),* 401–415.

Steinfeldt, E. (1963). *çcastotnvj slovar' sovremennogo russkogo literaturnogo jazyka. 2500 naibolee upotrebitel'nyx slov* [Frequency dictionary of contemporary standard Russian. 2500 most commonly used words]. Tallin: n.p.

Zaliznjak, A. A. (1967). *Russkoe imennoe slovoizmenenie.* [Russian nominal morphology]. Moscow: Nauka.

Zasorina, L. N. (Ed.). (1977). *çcastotnyj slovar' russkogo jazyka: Okolo 40,000 slov* [Frequency dictionary of Russian: About 40,000 words]. Moscow: Russkij jazyk.