Perception & Psychophysics
2004, ?? (?), ???-???

Gender and lexical access in Bulgarian

ELENA ANDONOVA
New Bulgarian University, City, Bulgaria

SIMONA D'AMICO and ANTONELLA DEVESCOVI
University of Rome "La Sapienza", Rome, Italy

and

ELIZABETH BATES
University of California at San Diego, La Jolla, California

Two procedures were used to explore the effects of semantic and grammatical gender on the recognition and processing of Bulgarian nouns, in relation to other factors that are known to affect lexical access. This study in a three-gender language was modeled on previous work in Italian, a two-gender language (Bates, Devescovi, Pizzamiglio, D'Amico, & Hernandez, 1995). Words were presented auditorily in randomized lists in two tasks: (1) repeat the word as quickly as possible and (2) determine the grammatical gender of the noun as soon as possible and indicate the decision by pressing a button. Reaction times in both tasks were influenced by phonological factors, word frequency, and irregularity of gender marking, but semantic and grammatical gender affected only gender monitoring. The significant contribution of semantic gender to processing in Bulgarian contrasts with previous findings for Italian. Also, we obtained an interaction between sex of the subject and noun gender, reflecting a bias toward one's own grammatical gender "counterpart" (especially for females). Reanalysis of the prior study in Italian showed a similar interaction but confirmed no effects of the semantic gender of the noun, suggesting that these two *natural gender* effects can dissociate. Possible reasons for cross-linguistic differences are discussed, with implications for comparative studies of gender and lexical access.

Recently, there has been a surge of interest in the nature and effects of grammatical gender in real-time language processing (Akhutina et al., 2001; Akhutina, Kurgansky, Polinsky, & Bates, 1999; Bates, Devescovi, Pizzamiglio, D'Amico, & Hernandez, 1995; Cacciari, Carreiras, & Cionini, 1997; Cacciari & Cubelli, 2003; Desrochers & Brabant, 1995; Desrochers & Paivio, 1990; Deutsch, Bentin, & Katz, 1999; De Vincenzi, 1999; Dominguez, Cuetos, & Segui, 1999; Friederici & Jacobsen, 1999; Heim, Opitz, & Friederici, 2002; Jacobsen, 1999; Jakubowicz & Faussart, 1998; Jescheniak, 1999; Jescheniak & Schriefers, 1999; Miceli et al., 2002; Plemmenou, Bard, & Branigan, 2002; Schiller & Caramazza, 2003; Schriefers & Teruel, 2000; van Berkum, 1997; Vigliocco & Franck, 1999; Vigliocco, Lauer, Damian, & Levelt, 2002). Grammatical gender is of interest for several reasons. First, gender processing plays a role in every component of language, including phonology, grammar, lexical semantics, and discourse (especially coreference). Hence, the study of gender contributes to our understanding of interactions among various components of the language processor. Second, the relation between grammatical and semantic gender poses a particularly interesting domain in which to study relationships between form and meaning. The correlation between grammatical and semantic gender is weak in many languages, involving arbitrary or indirect relations (e.g., the word for *bottle* is feminine in German, but the word for *girl* is neuter). In contrast, there are often strong correlations between grammatical gender and the phonological shape of the word. The relative strength or weakness of these correlations varies across languages, which may lead to different profiles of sound–meaning interaction in the study of lexical access. In this article, we will explore the interaction between semantic and grammatical gender in Bulgarian, including phonological cues, and will contrast these findings with those in an earlier study of lexical access in Italian (Bates et al., 1995). Although Italian and Bulgarian are both Indo-European languages, they are members of different branches (Romance and Slavic, respectively), with structurally distinct gender systems: only two genders in Italian (masculine and feminine) and three genders in Bul-

This study was conducted with research support from a McDonnell Foundation grant for "Crosslinguistic studies of aphasia." E.B.'s participation in data analysis and manuscript preparation was supported by Grant DC00216 from NIH/NIDCD. Partial support for E.A. during manuscript preparation was provided by a travel grant from NATO. We thank those who helped in the preparation and running of the experiment—above all, Daniela Angelova, Agop Erdeklian, and Ivailo Milenkov, as well as Encho Gerganov and Simona D'Amico for suggestions on the post-experimental stage. Address correspondence to E. Bates, Center for Research in Language, 0526, University of California at San Diego, La Jolla, CA 92093-0526 (e-mail: bates@crl.ucsd.edu).

Note—This article was accepted by the previous editorial team, headed by Neil Macmillan.

garian (masculine, feminine, and neuter). As we shall see, our results for Bulgarian include effects of semantic gender that have rarely been demonstrated before, a result that may be due, at least in part, to differences between two- and three-gender languages (see also Akhutina et al., 1999).

The two tasks used in this study and in the earlier study by Bates et al. (1995) vary in the degree to which they require conscious reflection on the property of gender. The first, repetition of spoken word targets (*cued shadowing*; Bates & Liu, 1996), requires very little metalinguistic reflection; the second task, classification of words according to gender (also known as gender monitoring or gender assignment), requires conscious attention to grammatical gender in order to reach a deliberate decision. With these two techniques, we will examine three questions concerning gender effects: (1) Does semantic gender have facilitative effects similar to those observed in other studies of lexical access with semantic variables such as imageability and concreteness? (2) Will faster responses be observed when there is a transparent relationship between grammatical gender and phonological properties of the word (in particular, the nature of the word ending)? (3) Will we observe interactions between sex of the subject (another form of *natural gender*) and grammatical and/or semantic gender of the noun?

The peculiarities of Bulgarian make it especially interesting for a cross-linguistic comparison of this kind. As is the case for all Slavic languages, gender is marked on nouns, adjectives, some verb forms, numerals, pronouns, and demonstratives; and in Bulgarian, gender is also marked on articles. Thus, gender plays an important role in subject–verb agreement, modifier–noun agreement, and pronoun–referent agreement. For example, finite verb forms agree with the subject in person and number, and in some past tense forms, they also agree with the subject in gender.

Noun gender in Bulgarian falls into three categories: masculine, feminine, and neuter. In general, masculine is considered the unmarked gender in Bulgarian, whereas feminine and neuter are both considered marked forms. The most common phonological markings for these three genders are the following: Most masculine nouns end in a consonant, most feminine nouns end in *a*/*ia*, and neuter nouns typically end in other vowels (*o*, *e*, *i*, *u*). These are the most productive patterns for each gender (Aleksieva, 1977; Vachkov, 1997), and any form of gender marking that deviates from these three patterns will be referred to here as a *mismatch*. In fact, there are many such mismatches—for example, nouns such as *prolet* (*spring*), which is feminine despite its consonantal ending; nouns such as *bashta* (*father*), which is masculine despite ending in *a*; generic gender nouns such as *lisitsa* (*fox*), which is feminine even though it is used to refer to animals of both sexes; and ambiguous nouns such as *drob* (*liver* and *mathematical fraction*), where one of the meanings of the word is associated with masculine and the other with feminine gender, respectively. In addition, approximately 2,500 abstract nouns in Bulgarian (Pashov, 1994) are derived from adjectives by means of a word-formative pattern that uses the suffixes *-ost* and *-est*; all of these nouns are in the feminine gender, even though the consonantal ending would lead to an expectation of masculine gender. Certain pairs of nouns in Bulgarian in masculine and feminine gender exhibit a regular word-formative pattern, by which feminine nouns are derived from masculine nouns by means of a suffix—for example, *-ka* as in *uchitel* (*male teacher*) versus *uchitelka* (*female teacher*)—which contributes to the greater average length of feminine nouns.

The availability of a neuter form in a three-gender language usually means that the correlation between biological and grammatical gender is somewhat stronger (although far from perfect) than in many two-gender languages, because a relatively low proportion of animate items end up in the neuter category. Hence, the *cue validity*, or information value, of semantic gender for grammatical gender assignment may be higher in Bulgarian than it is in Italian—perhaps leading to semantic gender effects of the sort that were not observed in Italian. Bates et al. (1995) reported absolutely no effects of semantic gender on either of the tasks that we employed here, including gender classification. They concluded that the arbitrary relation between grammatical and semantic gender for most nouns has resulted in a tendency to ignore any grammatical–semantic correlations that might exist, even in a task in which explicit and conscious decisions about gender assignment must be made. The results presented below will suggest some limits on the generality of this conclusion.

METHOD

Subjects

All the subjects were native Bulgarian speakers, university students, 18–28 years old. Fourteen males and 14 females participated in the word repetition task, and 9 males and 10 females participated in the gender-monitoring task. Most participated in both the word repetition and the gender-monitoring tasks, in that order. Following Bates et al. (1995) for Italian, all the analyses reported below were conducted over items, using mean scores over subjects as dependent variables (percentages correct in gender choice, reaction times [RTs] in gender choice and word repetition).

Materials

Stimulus selection was based on several criteria, including an approximately equal ratio of nouns in each of the three genders, selection of nouns in the singular only, and exclusion of acronyms, slang, proper names, and recently borrowed foreign words. Previous RT studies of Bulgarian (unpublished) and other languages (especially Bates et al., 1995) have shown that it is advisable to avoid words beginning with fricatives and affricates. Such words were excluded from the list (fricative-initial words were included in Italian but were controlled in regression analyses). Both words with and without semantic (biological) gender were included. In the Italian study, the criterion for semantic gender (+sexed) was "reference to an individual capable of sexual reproduction at some point in its life, whether or not sexual status of that individual is obvious to the speaker/listener" (Bates et al., 1995). In our study, however, semantic gender was treated in terms of the perception and awareness of the sexual status of the referent by the speaker/listener, which appeared more relevant

to our purposes than the more scientifically accurate but not so automatized classificatory principle employed for Italian. Some implications of this methodological decision will be considered later.

The main word sources used were dictionaries and two published grammars of Bulgarian (Pashov, 1994; Stojanov, 1980). A recurrent problem in psycholinguistic research in Bulgarian is the lack of reliable word frequency counts. Therefore, a subjective frequency norming on a 7-point scale (7, *most frequent*) was conducted with 17 subjects with characteristics similar to those of the subjects in the two main experimental tasks, yielding indirect data on the frequency distribution of items. The results from these ratings, as well as the description of other predictor variables, are included in Table 1. Another difference from the Italian study reflects the absence of age-of-acquisition data in Bulgarian. A final set of 13 predictor variables was derived for each of the 386 words selected for this study: three phonological factors (number of syllables, word length in milliseconds, and presence/absence of an initial vowel), four semantic factors (concreteness, humanness, animacy, and semantic gender), subjective word frequency (ratings from 1 to 7), grammatical gender, and a number of related characteristics that require further explanation, including marker–gender mismatch, generic nouns, diminutives, and abstract action nominals.

Marker–gender mismatch refers to the nature of the relationship between word ending and grammatical gender. This is similar to the phonological transparency/opacity variable used in Bates et al.'s (1995) study of Italian, but its definition in Bulgarian is slightly more elaborate. In Italian, the final vowel *-e* constitutes a single gender-opaque ending for both nouns and adjectives, contrasting with the high-probability *-a* ending for feminine nouns and adjectives and *-o* ending for masculine nouns and adjectives. Bulgarian does not have a single opaque, or baseline, pattern. Hence, marker–gender mismatch in Bulgarian refers to any exceptions (including some that are relatively productive) to the typical relation between the ending of a noun and its gender.

Semantic gender (as was noted above) was defined here as a characteristic of nouns referring to living beings (humans and animals) that are perceived as different on the basis of biological sex—for example, *balerina* (F) meaning *female ballet dancer*. It is not used here to refer to words representing animate creatures whose sexuality is entirely nonobvious to human language users (e.g., insect or fly). Hence, humanness and semantic gender are both subsets of animacy, which are, in turn, all subsets of concreteness. Generic gender nouns are those that are used to refer to both male and female representatives of the category although the word is in one specific gender only—for example, *majmuna* (*monkey/ape*, generic feminine) and *vrabche* (*sparrow*, generic neuter). Abstract action nominals are derived from verbs referring to an activity with the same semantics. In this sense, they do not evoke object-oriented but action-oriented imagery and are thus not perceived as concrete—for example, *minavane* (*passing*). All deverbal nouns are in the neuter gender with a clear unambiguous gender marking (*-ane* or *-ene*). Finally, a coding for the presence/absence of diminutive markers was made, to capture any effects that might be due to the correlation between gender and diminutive marking; indeed, diminutives in this data set are in the feminine and neuter genders only, which reflects the most productive pattern of their distribution in the Bulgarian language as a whole.

The 386 words were recorded by a female native speaker of Bulgarian in a neutral intonation with a falling tone, were digitized using the Macintosh SoundEdit system, and were placed in a sound file within the PsyScope experiment preparation package developed by MacWhinney et al. at Carnegie Mellon University. For each item, RT was measured from the onset of the stimulus word to the subject's response (as measured from the onset of the vocal response in word repetition and from the buttonpress in gender monitoring). A complete list of word stimuli is available in an appendix on our Web site at http://crl.ucsd.edu/~aszekely/ipnp/bugender.html.

Procedure

The subjects were tested individually in a quiet room and were told to repeat each word as quickly as possible without making a mistake for the word repetition task. In the gender-monitoring task, they were asked to press one of three buttons to indicate the gender of the noun they have just heard (from the same series of words just heard in word repetition, but in a different randomized order). In addition, since gender in Bulgarian is a three-member grammatical category, and in order to avoid bias, the subjects were assigned randomly to one of six spatial configurations of the three buttons in the gender-monitoring task—that is, M–F–N, M–N–F, F–N–M, F–M–N, N–M–F, or N–F–M. The subjects first performed the word repetition task and then the gender-monitoring task. This fixed order was chosen to ensure that word repetition performance was not influenced by instructions to focus on grammatical gender (a requirement for the gender-monitoring task). Accuracy and RT means were averaged over subjects and were entered into analyses over items.

RESULTS AND DISCUSSION

Performance on the Outcome Measures

In the gender-monitoring task, the average percentage of correct gender assignment on the 386 nouns was 97%

Table 1
Composition of the Word List ($N = 386$)

Grammatical gender	masculine	137	35.5%
	neuter	116	30.1%
	feminine	133	34.5%
Number of syllables	1	46	11.9%
	2	185	59.8%
	3	119	30.8%
	4	31	8.0%
	5	5	1.3%
Vowel initial	(0,1)	62	16.1%
Semantic gender	(0,1)	105	27.2%
Concrete	(0,1)	312	80.8%
Human	(0,1)	47	12.2%
Animate	(0,1)	119	30.8%
Generic	(0,1)	46	11.9%
Diminutive	(0,1)	29	7.5%
Abstract action nominals	(0,1)	18	4.7%
Marker–gender mismatch	(0,1)	63	16.3%
Word length (msec)	$M = 796.12$	$SD = 167.95$	range = 408–1,294

(range = 75%–100%, *SD* = 4%). These high accuracy levels indicate that the task was relatively easy for Bulgarian speakers. However, there was enough variance in accuracy to justify correlations, using gender-monitoring accuracy as an outcome variable.

RTs were measured from two reference points: the beginning and the end of the auditorily presented word. This was done because the length of words varies considerably in Bulgarian, both in milliseconds and in the number of syllables within each word. The average gender-monitoring RT (GMRT) from the end of words was 698 msec (range = 257–1,198 msec, *SD* = 158 msec), which reveals the relative difficulty of this task in comparison with the Italian study, where the corresponding average RT was 238 msec. This difference may reflect the complexity of gender assignment in Bulgarian, as well as the additional time required for the execution of a three-way choice among masculine, feminine, and neuter genders. The average GMRT from the beginning of words was 1,494 msec (range = 1,192–1,878 msec, *SD* = 125 msec), approximately twice as long as the GMRT from the end of words (see above).

Error rates were extremely low on the word repetition task (less than 1%), so low that they will be excluded from further consideration below in analyses of the effects of predictor variables on behavioral outcomes. The RT needed for the execution of the word repetition task (WRRT) was considerably shorter than that for the GMRT task. Mean WRRT from the beginning of words was 1,043 msec (range = 759–1,479 msec, *SD* = 122 msec), and from the end of the word it was 247 msec (range = 11–493 msec, *SD* = 86 msec). On average, these WRRTs were 453 msec shorter than the GMRTs described above. This difference is due to the more complex processing required for the GMRT task, including the selection of one of three possible gender labels, as well as the degree of conscious reflection required for the monitoring task. In contrast with Bates et al.'s (1995) study of Italian, Bulgarian RTs were not considered from the word uniqueness point, since in many cases it is unclear what that point would be in the highly complex morphology of Bulgarian.

Correlations among the outcome measures are summarized next. RTs for the two tasks (word repetition and gender monitoring) were highly correlated, both from the beginning of the word ($r = .49$, $p < .001$) and from the end of the word ($r = .61$, $p < .001$). On the gender-monitoring task, RTs from the beginning versus the end of the word were also positively correlated ($r = .31$, $p < .001$). By contrast, the two different RT points were correlated negatively in the word repetition task ($r = -.25$, $p < .001$). This interesting contrast may reflect some important differences between these two tasks. Gender monitoring takes substantially longer overall, and it requires explicit judgments of gender that are (as we shall see shortly) influenced by the word's ending. On both these grounds, it seems that gender classification decisions are usually not made until the word is complete; hence, results are similar whether RTs are measured from the beginning or the end of the word. By contrast, word repetition is substantially faster and does not require any explicit judgments of the word or its ending. Hence, repetition can be initiated as soon as the word is recognized, which may occur before the entire word has been heard. As a result, measurements taken from the beginning versus the end of the word are sensitive to different characteristics of word structure on the word repetition task. In the correlation and regression analyses reported below, we will use *word-final RTs* as our dependent variable on both tasks, to maximize sensitivity to possible task-based differences in the factors that influence RTs. However, we will also use word length as a predictor, so that the large variations in length in our lexical materials are taken into account.

Finally, the outcome variable of percentage of correct choice of gender was negatively correlated with the RTs in gender monitoring ($r = -.33$, $p < .001$, from the end of the word, and $r = -.30$, $p < .001$, from the beginning of the word), reflecting a single dimension of word difficulty (resulting in lower accuracy and longer RTs). This also shows that there is no speed–accuracy tradeoff in this task.

Relationships Among the Predictor Variables

Tables 2A–2C list correlations among all our predictor variables, computed over items.

Three of the semantic variables (concreteness, humanness, and animacy) were significantly correlated in the expected directions. All three also were correlated significantly with semantic gender. In particular, animacy was positively and strongly correlated with semantic gender ($r = .92$, $p < .001$), for obvious reasons. However, unlike the Italian study, with its broader definition of semantic gender, animacy and semantic gender are not identical here (e.g., we do not assign semantic gender to *insect* or *butterfly*), which allows finer distinctions of effects to be drawn.

The variables associated with word form were also significantly correlated, including a strong positive relation between length in milliseconds and number of syllables ($r = .73$, $p < .001$) and a weaker positive relation of both length measures to presence of an initial vowel (i.e., vowel-initial nouns tended to be longer in this data set). As in other languages, word length was negatively correlated with frequency ratings, although the relationship reached significance only when length was mea-

Table 2A
Correlations Among Predictor Variables

	Frequency	Concrete	Human	Animate
Frequency	–			
Concrete (0,1)	−.12*	–		
Human (0,1)	+.03	+.18‡	–	
Animate (0,1)	−.17‡	+.30‡	+.56‡	–
Semantic gender (0,1)	−.16†	+.30‡	+.59‡	+.92‡

*$p < .05$. †$p < .01$. ‡$p < .001$.

Table 2B
Correlations Among the Predictor Variables

	Vowel Initial	No. of Syllables	Mismatch	Generic	Diminutives	Abstract Action Nominals	Word Length
Vowel initial	–						
No. of syllables	+.22‡	–					
Mismatch	+.06	−.23‡	–				
Generic gender	+.04	+.01	−.16†	–			
Diminutives	−.04	+.17‡	−.13*	−.07	–		
Abstract action nominals	−.06	+.20‡	−.10	−.08	−.06	–	
Word length	+.18‡	+.73‡	−.05	+.07	+.23‡	+.02	–
Masculine	+.01	−.26‡	−.18‡	+.14†	−.19‡	−.16†	−.11*
Feminine	+.00	+.02	+.47‡	−.05	−.01	−.16†	+.13*
Neuter	−.06	+.26‡	−.29‡	−.10	+.21‡	+.34‡	−.02

*$p < .05$. †$p < .01$. ‡$p < .001$.

sured in milliseconds. Diminutives were also correlated with word length, reflecting the fact that Bulgarian and other Slavic languages create diminutives through addition of a suffix (predominantly in the neuter and feminine genders in Bulgarian).

Particularly important for our purposes here, Tables 2A–2C show that grammatical gender is confounded with a host of variables that could influence accuracy and RTs. For example, the masculine nouns in this data set tend to be higher in concreteness, animacy, and humanness, and they are correlated more often with semantic gender. Masculine words also tend to be shorter, and they are more likely to be classified as generic (could refer to masculine or feminine referents). Masculine nouns rarely occur as diminutives (an interesting contrast with other Slavic languages, such as Russian), never as abstract action nominals, and very rarely as words with irregular (mismatching) gender marking (although a small number of masculine nouns, including some used in this study, do end with *-a,* mimicking feminine nouns). By contrast, feminine nouns in this data set tend to be lower in concreteness, but they are not significantly higher in animacy, humanness, or semantic gender. Feminine gender is not found in abstract action nominals (*deverbal* nouns), but feminine nouns do appear much more often with an irregular (mismatching) ending. Finally, neuter nouns tend to be low in humanness, animacy, and semantic gender, but not in concreteness. Diminutives occur more often in neuter gender, as do abstract action nominals, but neuters in this data set do not carry an irregular (mismatching) ending. All of these results point to the importance of conducting stepwise regression analyses evaluating the contribution of each factor after all the others are controlled.

Relations Between Predictor and Outcome Variables

Pearson product–moment correlations among all the predictor and outcome variables are reported in Table 3. Table 4 summarizes the joint and unique variance contributed by these predictor variables to performance on the gender-monitoring and word repetition tasks, controlling for some of the confounds among predictors that are evident in Table 3.

Frequency ratings. Frequency and/or familiarity have been shown to play an important role in language processing across a wide range of tasks and languages. That is why the absence of frequency effects in Bates et al.'s (1995) study was quite surprising, leading the authors to conclude that "spoken-word frequency contributes very little to these measures of auditory lexical access in Italian." In the present study, subjective ratings of frequency

Table 2C
Correlations Among the Predictor Variables

	Frequency	Concrete	Human	Animate	Semantic Gender
Vowel initial	−.09	−.04	+.03	+.01	+.03
No. of syllables	−.07	−.01	+.21‡	+.10	+.08
Mismatch	+.05	−.39‡	+.01	−.16†	−.14†
Generic gender	−.14†	+.18‡	+.33‡	+.55‡	+.53‡
Diminutive	−.25‡	+.14†	−.11*	−.06	−.06
Abstract action nominals	+.08	−.45‡	−.08	−.15†	−.14†
Word length	−.12*	−.03	+.24‡	+.10*	+.09
Masculine	+.05	+.17†	+.20‡	+.15†	+.15†
Feminine	+.00	−.15†	−.04	−.04	−.05
Neuter	−.05	−.02	−.17‡	−.11*	−.11*

*$p < .05$. †$p < .01$. ‡$p < .001$.

Table 3
Correlations Between Predictor and Outcome Variables

	% Correct	GMRT	WRRT
Frequency	+.14†	−.13*	−.16†
Concrete	n.s.	−.19‡	n.s.
Human	n.s.	n.s.	+.25‡
Animate	+.17‡	−.20‡	+.16†
Vowel	n.s.	+.18‡	+.12†
Syllable	n.s.	+.30‡	+.64‡
Gender–marker mismatch	−.14†	n.s.	−.15†
Semantic gender	+.19‡	−.21‡	+.10*
Generic	n.s.	n.s.	+.10*
Diminutives	n.s.	n.s.	−.19‡
Abstract action nominals	n.s.	n.s.	n.s.
Word length	n.s.	+.45‡	+.86‡
Masculine	n.s.	n.s.	n.s.
Feminine	n.s.	n.s.	n.s.
Neuter	n.s.	n.s.	n.s.

Note—GMRT, gender-monitoring reaction time; WRRT, word repetition reaction time. *$p < .05$. †$p < .01$. ‡$p < .001$. n.s., nonsignificant.

were significantly correlated with all three outcome variables: More frequent nouns took less time to process in both tasks and elicited fewer errors in gender classification (Table 3).

We established earlier that these frequency ratings are confounded with a number of semantic and phonological predictors. However, stepwise regression analyses (Table 4) yielded small but significant effects of frequency, after other factors were controlled, for gender-monitoring accuracy and for RTs in both tasks. This difference between our study and Bates et al.'s (1995) study of Italian may reflect methodological differences (i.e., frequency estimates based on spoken language corpora in Italian, subjective ratings of frequency in Bulgarian). Indeed, previous studies within English have confirmed that subjective ratings of familiarity and/or age of acquisition are often better predictors of RT than are objective frequency corpora (Brown & Watson, 1987; Carroll & White, 1973; Gilhooly & Gilhooly, 1980).

Another possible methodological difference lies in the fact that the same subjects were used for both tasks in Bulgarian, but not in Italian. Hence, the Bulgarians had more exposure to the target words, which could have reduced automatic effects and enhanced strategic effects. However, this factor is unlikely to account for the presence of frequency effects in Bulgarian and their absence in Italian, since frequency effects are generally believed to be automatic effects that should (if anything) decrease with word repetitions—the opposite of our findings for frequency in these two languages. On the off-chance that some kind of frequency effect was building across the course of the Italian study, we returned to the Italian data and reanalyzed the correlations between frequency and performance with the data broken down into quartiles on the basis of order of presentation (first quarter, second quarter, third quarter, or fourth quarter) in the randomized lists for each subject. There were no significant frequency–performance correlations in any of these comparisons and, hence, no evidence whatsoever for the strategic waxing or waning of a frequency effect in Italian.

Phonological factors. Word length (measured in milliseconds and in number of syllables) was associated with longer RTs in both tasks (Table 3), although this result held up in the regression analyses only for length measured in milliseconds. There was no effect of word length on gender-monitoring accuracy. These results replicate findings for Italian, with one exception: Bates et al. (1995) reported that RTs were actually shorter for words with more syllables when length in milliseconds was controlled, suggesting that the unique structure of multisyllabic words contributes to their recognition. Although one might have predicted a similar effect in Bulgarian, it did not appear in our analyses, even though the strength of the correlation between word length and number of syllables was equivalent in these two languages ($r = +.73$ in both studies). We will return to this point later, in reference to language-specific effects of word structure.

Bates et al. (1995) reported longer RTs for words with an initial fricative in both tasks. Although we did not study the effects of frication as a variable (excluding fricative-initial words from the data set), we did find shorter RTs for words that begin with a consonant, similarly to previous results on gender monitoring in French (Desrochers & Paivio, 1990). However, this variable was confounded with other measures, including word length, and disappeared when other factors were controlled in stepwise regressions.

Semantic factors. The three semantic factors of concreteness, animacy, and humanness were all correlated with one or more aspects of performance on the two tasks: Concrete words were associated with shorter GMRTs, "human" words were associated with longer WRRTs, and animate words tended to elicit faster and more accu-

Table 4
Unique Variance Contributed by Each Predictor Variable When Entered on the Last Step

	% Correct	GMRT (%)	WRRT (%)
Frequency	+3.1‡	−1.4†	−0.4*
Concrete	n.s.	n.s.	n.s.
Human	n.s.	n.s.	+0.3*
Animate	n.s.	n.s.	n.s.
Vowel	n.s.	n.s.	n.s.
Syllable	n.s.	n.s.	n.s.
Marker–gender mismatch	−1.3*	+3.6‡	−1.1‡
Semantic gender	+3.7‡	−6.3‡	n.s.
Generic	n.s.	n.s.	n.s.
Diminutives	n.s.	−0.8*	n.s.
Abstract action nominals	n.s.	n.s.	n.s.
Length	n.s.	+20.2‡	+73.9‡
Masculine	n.s.	n.s.	n.s.
Feminine	n.s.	−2.6‡	n.s.
Neuter	n.s.	n.s.	n.s.

Note—GMRT, gender-monitoring reaction time; WRRT, word repetition reaction time. +, Positive partial correlation for significant individual effects. −, Negative partial correlation for significant individual effects. *$p < .05$. †$p < .01$. ‡$p < .001$. n.s., nonsignificant.

rate responses in gender monitoring but longer RTs in word repetition. However, stepwise regressions (Table 4) indicate that most of these effects disappear when their confounds with other variables are controlled. The only significant contribution that remained on the final step was a modest positive relationship between humanness and WRRT, similar in size and direction to the results in the Italian study (i.e., longer RTs for words with human referents). We agree with Bates et al. (1995) that these semantic effects are minimal in these tasks.

However, a potentially important difference between the two studies emerged when the contribution of semantic gender was examined. Bates et al. (1995) found no effect whatsoever of semantic gender on any of their outcome variables. In our study, semantic gender had no effect on WRRTs (in line with Bates et al., 1995), but did affect both accuracy and RTs in the gender-monitoring task (in contrast with Bates et al., 1995). In particular, words with referents that have biological gender elicited faster and more accurate gender classifications. This may mean that semantic gender plays a more important role in a three-gender language, at least for a task that requires conscious attention to gender.

Recall, however, that semantic gender was defined differently in the Italian study. Bates et al. (1995) used a scientifically precise definition of semantic gender, applying to any word whose referent is capable of sexual reproduction. In our study, semantic gender was defined in accordance with the common understanding of laymen. To determine whether this methodological difference would account for the different results obtained in Bulgarian and Italian, we returned to the original data reported by Bates et al. and recoded items in that study to reflect the same lay definition of biological gender that we have employed here. This recoding had no effect on the Italian results: There was still no effect of semantic gender on gender-monitoring accuracy or RTs when items were redefined according to the Bulgarian criterion. Bates et al. also included in their word list a subset of words that were matched pairs, differing in semantic gender and in phonological form only by the final vowel (e.g., *gatto*, which is either masculine or unmarked for *cat*; *gatta*, which is marked and female for *cat*). We had no such contrasting pairs in our list. To determine whether these items might have somehow influenced the difference in results, we also reanalyzed Bates et al.'s (1995) data, eliminating all matched pairs. It was still the case, with either Italian or Bulgarian coding for semantic gender, that no semantic gender effect could be found in Italian. Finally, we recalculated the correlations between semantic gender and performance within the four quartiles, as reported above for frequency, to determine whether a nascent semantic effect might be detected in Italian toward the end of the study (indicating a build-up of strategies). There were no significant correlations with semantic gender in Italian within any of the four quartiles.

We conclude that the difference between Italian and Bulgarian results probably does not lie in methodological differences alone, but in cross-language differences that may include a meaningful and potentially important difference between a two- and a three-gender language (but see Vigliocco & Franck, 1999, discussed below). In particular, we suggest that semantic gender has a bigger impact in Bulgarian because grammatical gender has higher *cue validity* in a three-gender language (i.e., it is correlated better with semantic gender).

Grammatical gender. Because the grammatical category of gender in Bulgarian is not dichotomous, it was included in the correlation and regression analyses by recoding into three dummy variables (±masculine, ±feminine, and ±neuter). The correlational analysis with accuracy and RTs from the *beginning* of the word did not reveal any reliable correspondence between any of the three genders as predictors of performance in our tasks (Table 3). In contrast, Table 3 shows several significant correlations when RT was measured from the *end* of the word. Specifically, feminine nouns were associated with shorter RTs on both tasks, masculine nouns were associated with longer RTs, but only on gender monitoring, and neuter nouns were associated with longer RTs in word repetition. Stepwise regressions controlling for confounds between gender and a host of predictor variables (including word length) eliminated all but one of these effects: Feminine nouns were associated with shorter RTs after other variables were controlled, but only on gender monitoring. As was noted earlier, the feminine gender is the most marked member of the three-tier gender system in Bulgarian, and as such, it may be more salient in a task that requires explicit attention to gender.

In this experiment, we used the variable of marker–gender mismatch to assess the importance of noun endings as a phonological cue to gender assignment. The meaning of this variable is very similar to that of the transparency of gender marking in the Italian study. It is important, therefore, to point out the similarities and differences between the results for the two languages. In Italian, phonological opaqueness led to lower accuracy and longer RTs in gender monitoring but had no reliable effect on repetition when other factors were controlled. In Bulgarian, marker–gender mismatch played a similar role in gender monitoring (lowering accuracy and increasing RTs), but it had a rather different effect on word repetition. Specifically, words with a mismatching gender marker resulted in significantly *shorter* WRRTs when other predictors were controlled. The explanation for this surprising result probably lies in the details of word formation in Bulgarian. In particular, many of the irregular (mismatch) words are relatively easy to recognize because of their initial segments. For example, if a word starts as an adjective (e.g., *mlad*, which means *young*), it is easy (at least within the other constraints present in our word list) to predict how that word would end—that is, with a nominalizing ending, such as *mladost* (which means *youth*). Hence, the subject can initiate a repetition before the rest of the word has appeared. Presumably the same recognition bias would also operate in the gender-

monitoring task. However, that task involves postlexical decisions that lengthen RTs and force attention to the end of the word, overriding the recognition advantage that appears in the repetition task.

Finally, Table 4 yields little evidence for a unique contribution by the other gender-associated word structure variables, including diminutives, abstract action nominals, and generic nouns. The one exception is a small but significant negative contribution from diminutives, which seem to be associated with shorter RTs in gender monitoring when other variables are controlled. As we saw in our examination of correlations among the predictor variables, diminutives are significantly less likely to be masculine and more likely to be neuter. Hence, native Bulgarian speakers may make use of this information in the gender classification task.

Sex of subject. The presence of natural gender effects in this study of Bulgarian led us to consider a possibility that was not entertained by Bates et al. (1995) in their study of Italian. Could the sex of the subject influence the timing and/or accuracy of gender processing, particularly within a task such as gender monitoring that requires explicit attention to the gender category? There are some rather straightforward reasons why we might expect a finding of this kind. Because gender is marked so pervasively in this language, not only on the noun, but also with agreement markers on many different modifiers (including some forms of the verb), the relative frequency, or "diet," of masculine and feminine forms differs markedly for men and women. Throughout life, a Bulgarian woman makes more first-person references in the feminine than does a Bulgarian man, she hears more second-person feminine references addressed to her in conversation, and overhears more third-person feminine references about her in the conversation of others. Hence, not only is there a baseline difference in frequency of forms, but also the imbalance of masculine versus feminine marking for men versus women tends to occur on references to and about the self—utterances that tend to attract one's attention more than many other kinds of speech.

To determine whether sex of subject might have an effect on performance in these tasks, we conducted three 3×2 mixed analyses of variance over items (for GMRTs, gender-monitoring accuracy, and WRRTs), in which sex of the subject was included as a within-items factor (e.g., female RT vs. male RT) and grammatical gender (three levels) was treated as a between-items factor. Because we know that grammatical gender is confounded with other item characteristics, these analyses were conducted as multivariate analyses of covariance (MANCOVAs), removing the effects of frequency, length in milliseconds, concreteness, and presence/absence of an initial vowel. All effects involving within-items variables were Greenhouse–Geisser corrected.

The first MANCOVA, on GMRTs in Bulgarian, yielded a main effect of grammatical gender [$F(2,392) = 6.61, p < .002$], in the direction feminine < neuter < masculine. There was no main effect of sex of the subject [$F(1,392) = 2.53$, n.s.], but the interaction of grammatical gender and sex of the subject was significant [$F(2,392) = 11.33, p < .0001$]. There were no interactions with covariates, indicating that we met the assumptions for a MANCOVA in this analysis. The interaction is illustrated in Figure 1, which shows that women were especially slow in classifying masculine nouns and especially fast in classifying feminine nouns, as we might expect if a lifetime of differential exposure to masculine versus feminine forms were to have an influence on performance. In contrast, the performance for males was flatter across gender categories, in the direction neuter > masculine > feminine. Separate post hoc ANCOVAs were conducted for males and females, respectively, using the same covariates. For women, there was a significant effect of gender [$F(2,392) = 10.77, p < .0001$]. The analysis for men also reached significance [$F(2,392) = 3.68, p < .03$], albeit attenuated relative to the effect for women (although performance for masculine nouns was slightly faster than that for neuter nouns for men). So the subject's experience as a man or a woman alters the magnitude more than does the absolute shape of GMRTs in Bulgarian.

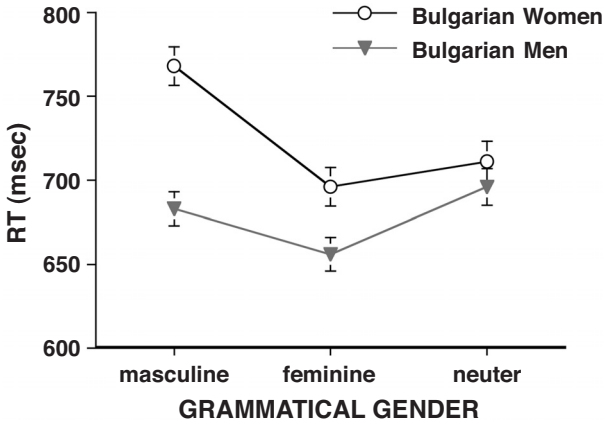

Figure 1. Interaction between grammatical gender and sex of the subject for Bulgarian gender-monitoring reaction times (GMRTs).

The second MANCOVA, on gender-monitoring accuracy in Bulgarian, yielded no significant main effect of grammatical gender ($F < 1$, n.s.) and no main effect of sex of the subject ($F < 1$, n.s.), nor were there any confounding interactions with the covariates. There was, however, a significant interaction between sex of the subject and grammatical gender [$F(2,392) = 6.63, p < .0001$], illustrated in Figure 2, which shows that women made more errors on masculine nouns, whereas men made more of their errors on feminine nouns. This result complements the RT findings, suggesting that personal experience can shape sensitivity to grammatical gender in a gender-monitoring task, at least in this language.

The third MANCOVA, on word repetition, yielded a significant main effect of grammatical gender [$F(2,392) = 3.49, p < .04$], reflecting the fact that feminine nouns (the

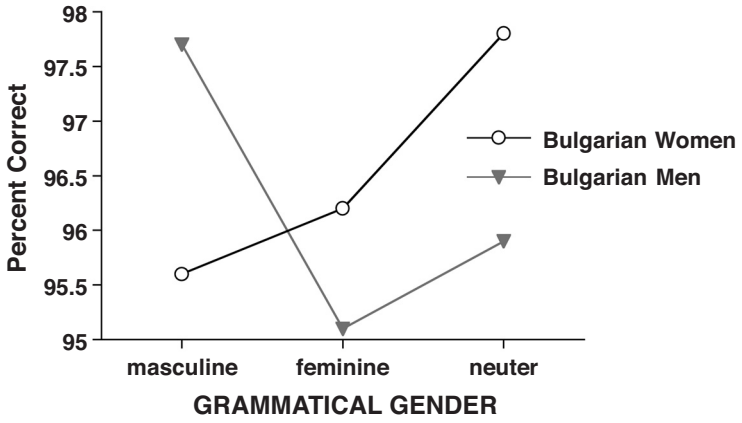

Figure 2. Interaction between grammatical gender and sex of the subject for Bulgarian gender-monitoring accuracies.

most marked form) are especially fast in Bulgarian—a finding that we noted earlier and attributed to the greater salience of marked forms. There was no main effect of sex of the subject ($F < 1$, n.s.), although a gender × sex-of-subject interaction just missed significance [$F(2,392) = 2.82$, $p < .07$], illustrated in Figure 3. We also found an interaction between sex of the subject and word length, which means that we were technically in violation of the assumptions of a MANCOVA. To overcome this problem, we took a regression approach and calculated the difference between RTs for men versus women on each item. These difference scores were used as the dependent variable in a regression analysis using length, concreteness, vowel-initial status, frequency, and gender as predictors (contrasting masculine and feminine only as a dichotomous dummy variable, excluding neuters from the analysis). The overall equation explained 25.3% of the variance in WRRTs ($p < .0001$). Most important for our purposes here, the dichotomous masculine–feminine contrast added a significant 4.3% to the variance accounted for ($p < .0001$) when the other variables were controlled. The direction of this contribution corresponds to the data in Figure 3: The sex difference in RTs was smaller for feminine items. This word repetition effect is weaker than the results for gender monitoring, but it runs in the same direction.

Having found yet another surprising effect of natural gender (in this case, sex of the subject), we decided to return to the Italian data to see whether similar interactions might be found there. We conducted the same three mixed MANCOVAs as those reported above on GMRTs, gender-monitoring accuracy, and WRRTs, covarying for length, frequency, initial frication, and concreteness. Within-items effects were Greenhouse–Geisser corrected.

In the first MANCOVA on GMRTs for Italian, there was no main effect of grammatical gender ($F < 1$, n.s.). There was also no main effect of sex of the subject [$F(1,462) = 1.01$, n.s.], although performance by women

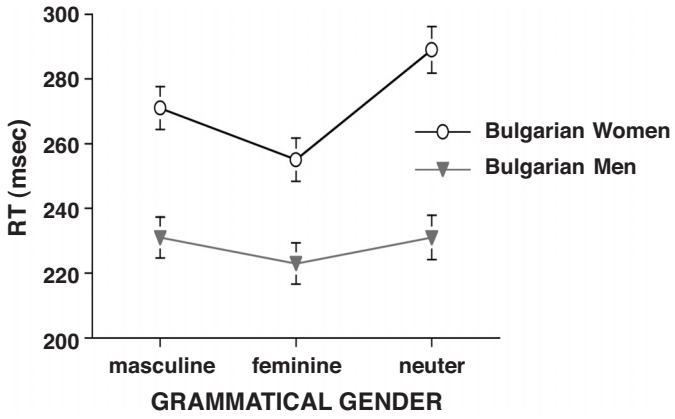

Figure 3. Interaction between grammatical gender and sex of the subject in Bulgarian word repetition.

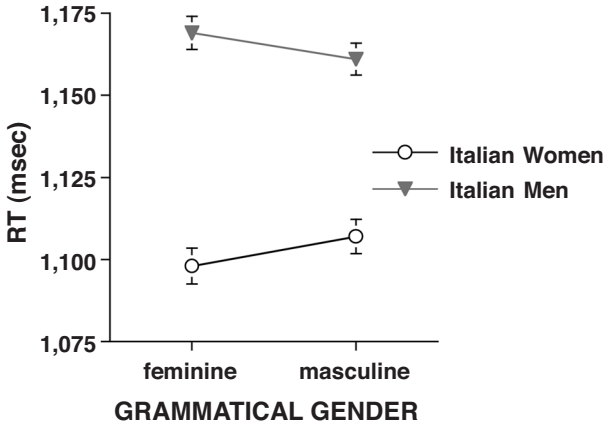

Figure 4. Interaction between grammatical gender and sex of the subject for Italian gender-monitoring reaction times (GMRTs).

tended to be numerically faster (in the opposite direction from that for Bulgarian). Most important for our purposes here, there was a significant interaction between grammatical gender and sex of the subject [$F(1,462) = 6.32$, $p < .02$], illustrated in Figure 4. This figure indicates that women show slightly shorter RTs for feminine nouns, whereas males show shorter RTs for masculine nouns. Aside from this effect, the analysis also revealed a significant interaction of sex with word length ($p < .0001$), which means that we do not meet all the assumptions required for a MANCOVA. We therefore took an alternative approach, calculating difference scores in RTs for men versus women for each item and using these as the dependent variable in a regression analysis, with word length, frequency, concreteness, initial frication, and a dichotomous variable for grammatical gender as predictors. The overall analysis accounted for only 4.7% of the variance ($p < .003$), but gender added a significant 1.3% ($p < .012$) when it was added on the last step, in line with the trends in Figure 4.

Second, the MANCOVA on gender-monitoring accuracy for Italian yielded a main effect of grammatical gender [$F(1,462) = 5.07$, $p < .025$], reflecting higher accuracies for masculine nouns. The interaction just missed significance [$F(2,462) = 3.72$, $p < .06$], but examination of cell means indicated that this trend was due primarily to better performance by Italian women on masculine nouns, which was not in line with predictions.

Finally, the Italian word repetition analysis yielded a main effect of gender [$F(1,462) = 3.96$, $p < .05$], reflecting shorter RTs on masculine nouns (another contrast with Bulgarian, where RTs on the unmarked masculine tend to be relatively long). There was no main effect of sex of the subject and no significant interaction (both Fs < 1, n.s). However, significant interactions with two of the covariates did emerge (with length and frication), violating the assumptions for a MANCOVA. We therefore double-checked this result by taking a regression approach, using male–female differences in RT for each item as the dependent variable, with length, frication, frequency, concreteness, and a dummy variable for grammatical gender as predictors. The overall analysis was significant (8.7% of the variance explained, $p < .0001$), but grammatical gender did not make any significant contribution to sex differences when the other variables were controlled. Hence, we have no evidence for a sex-of-subject effect on WRRTs in Italian.

To summarize, we found the predicted interaction between grammatical gender and sex of the subject in Bulgarian, for both GMRT and gender-monitoring accuracy. There was also a trend in the same direction for word repetition, most evident in the regression analysis (where the unique contribution of masculine–feminine gender to male–female differences did reach significance). However, the result was certainly more robust for gender monitoring, a task that requires explicit attention to gender. In view of all the differences we have found thus far between Italian and Bulgarian, we were somewhat surprised that our reanalyses revealed an analogous interaction between sex of the subject and grammatical gender for Italian, although it was only apparent for GMRT (and not for gender-monitoring accuracy or for WRRT). Even though the overall shape of grammatical gender effects is quite different for Bulgarian and Italian and male–female differences in overall RTs actually run in the opposite direction, the general pattern is one in which (at least for women) there is a bias in favor of the subject's own counterpart in grammatical gender. Hence, a lifetime of hearing more feminine forms may influence the listener's attention and lead to biases toward feminine grammatical gender (and at least in Bulgarian, a male bias toward masculine gender). This seems to be true independent of the presence or absence of "natural" semantic gender effects on the same task (i.e., semantic gender effects in Bulgarian, but not in Italian).

CONCLUSION

This study has revealed several factors that contribute to the processing of gender-bearing Bulgarian nouns, presented out of context in the auditory modality. Many of these effects constitute a replication of research in other languages (notably the Italian study by Bates et al., 1995, on which this work was modeled). However, we also uncovered a number of differences that may reflect interesting differences between languages, including language-specific principles of word formation and word structure and the contrast between two- and three-gender systems.

Word length had a powerful influence on auditory lexical access in Bulgarian in the word repetition task (which requires no metalinguistic reflection) and in the gender-monitoring task (which encourages attention to gender-bearing information at the end of the word). This result is not surprising. However, we found an interesting difference between our two tasks when RTs were compared

at the beginning versus the end of the word. Specifically, GMRTs from these two time points were positively correlated, but WRRTs measured from word onset were negatively correlated with WRRTs measured from the end of the word. This suggests that repetitions are initiated as soon as the word is identified, which may occur before the end of the word, whereas gender classifications are postponed until the end of the word, when postlexical decision processes can be applied (which may include an explicit check of the word's ending).

Most of our semantic factors (animacy, humanness, and concreteness) had little effect on performance in either task, particularly when they were evaluated in stepwise regressions controlling for other confounds. In this regard, our results are similar to those in Bates et al. (1995) for Italian. However, in contrast with the Italian study, we did find significant effects of semantic gender on gender monitoring in Bulgarian. Specifically, words that refer to animate beings with recognizable semantic gender elicited faster and more accurate gender classifications, even after various confounds were controlled through stepwise regression. On the basis of their null results for a two-gender language, Bates et al. concluded that the relationship between grammatical and semantic gender is so arbitrary and indirect that it is ignored by native speakers even in a conscious gender-processing task. Our results suggest that this conclusion is, at the very least, not a language universal. In particular, it may not hold for a three-gender language, due perhaps to the higher correlation (cue validity) between semantic and grammatical gender when the neuter category (with relatively few animate forms) is available.

Mindful that we had used a somewhat different definition of *semantic gender* (based on assessments of biological gender that are obvious and accessible to laymen), we checked the original Italian data to determine whether the results would change if our definition were applied. This did not occur, strengthening our confidence in the hypothesis that processing of grammatical gender can differ over languages, playing a more important role in a three-gender language in which the correlation between grammatical and semantic gender is somewhat higher (although still far from absolute). We also checked to see whether a strategic semantic gender effect might have been building (undetected) in the Italian data, by looking to see whether the expected correlation increased across the course of the task (over quartiles). This was not the case. We should also note, however, that Vigliocco and Franck (1999) did find significant effects of biological/conceptual gender in both French and Italian (both of which are two-gender languages), although they were looking at a very different modality (language production) with markedly different tasks (constrained sentence completion). Hence, the effects of conceptual gender on lexical processing may be a matter of degree, showing up more readily in a three-gender language, such as Bulgarian (e.g., for recognition of single words, out of context), but detectable in two-gender languages, as well under some conditions.

The latter possibility is strengthened by a different kind of natural gender effect that appeared not only in our Bulgarian data, but in reanalyses of the data for Italian. For gender monitoring (but not for word repetition), we found significant interactions in both languages between sex of the subject and grammatical gender. In particular, women seem to be faster and (in Bulgarian) more accurate in the classification of feminine nouns. We suggest that this result reflects a lifetime of experience in producing (in the first person) and listening (in the second and third persons) to references about oneself—a kind of speech that is not only high in frequency, but also high in interest value for most listeners. This interpretation is tentative, and the result itself is so new (and so limited) that it will require much more investigation before firm conclusions can be reached. But the possibility of such sex-of-subject effects would provide strong evidence for the role of episodic memory in language processing, even for the arcane details of grammatical gender.

We examined the effects of several additional variables peculiar to gender marking, word formation, and word structure in Bulgarian, including regularity of phonological marking (i.e., whether the ending of the word matches the typical pattern of marking for each gender). Using a similar regularity variable (which they referred to as *phonological transparency/opacity*), Bates et al. (1995) reported no effects of gender marking on word repetition. However, they did find significant facilitative effects of transparent gender marking on accuracy and RTs in the gender-monitoring task. We found a similar facilitative effect in the monitoring task, but we also found a significant negative effect of a gender match in the repetition task. That is, words with atypical gender marking (i.e., gender mismatch) resulted in faster repetition. We attribute this result not to gender marking per se, but to the peculiarities of word formation patterns in Bulgarian. For example, nouns derived from adjectives have a characteristic beginning and end, which increases their predictability (thereby permitting recognition and initiation of response before the end of the word). Coincidentally, these words tend to end in a gender-atypical suffix. Results like these underscore the importance of considering detailed and language-specific lexical and phonotactic features when techniques for the assessment of auditory lexical access are applied in a new language.

A final difference between our results and those of Bates et al. (1995) revolves around the contribution of word frequency. Bates et al. found no effect of word frequency (on the basis of spoken corpora) on performance in either task when confounding variables were controlled. We found small but significant effects of frequency (on the basis of subjective frequency ratings) in both tasks, in raw correlations, and in stepwise regres-

sions. We suspect that this difference in results has little to do with cross-linguistic differences between Italian and Bulgarian, reflecting instead the well-known (although unexplained) finding that subjective ratings of frequency or familiarity are better predictors of RT than are frequency measures taken from written or spoken corpora (Brown & Watson, 1987).

Although this is only the first exploratory study of gender processing and lexical access in Bulgarian, it offers insight into the specificity of these processes in a less studied language, one with structural characteristics quite different from those of English, German (German, Dutch), and/or the various Romance languages in which gender processing has been investigated to date. Our results suggest that these factors do make a difference.

REFERENCES

Akhutina, T., Kurgansky, A., Kurganskaya, M., Polinsky, M., Polonskaya, N., Larina, O., Bates, E., & Appelbaum, M. (2001). Processing of grammatical gender in normal and aphasic speakers of Russian. *Cortex*, **37**, 295-326.

Akhutina, T., Kurgansky, A., Polinsky, M., & Bates, E. (1999). Processing of grammatical gender in a three-gender system: Experimental evidence from Russian. *Journal of Psycholinguistic Research*, **28**, 695-713.

Aleksieva, N. (1977). Gramaticheski i semantichni faktori pri opredeljane na roda na anglijskite zaemki v bulgarski ezik [Grammatical and semantic factors in assigning gender to English loanwords in Bulgarian]. *Bjuletin za Supostavitelno Izsledvane na Bulgarskija Ezik s Drugi Ezitsi*, 4-5, 44-54.

Bates, E., Devescovi, A., Pizzamiglio, L., D'Amico, S., & Hernandez, A. (1995). Gender and lexical access in Italian. *Perception & Psychophysics*, **57**, 847-862.

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

Brown, G. D. A., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory & Cognition*, **15**, 208-216.

Cacciari, C., Carreiras, M., & Cionini, C. B. (1997). When words have two genders: Anaphor resolution for Italian functionally ambiguous words. *Journal of Memory & Language*, **37**, 517-532.

Cacciari, C., & Cubelli, R. (Eds.) (2003). Special section on the neuropsychology of grammatical gender. *Cortex*, **39**, 377-508.

Carroll, J. B., & White, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology*, **25**, 85-95.

Desrochers, A., & Brabant, M. (1995). Interaction entre facteurs phonologiques et sémantiques dans une épreuve de catégorisation lexicale. *Revue Canadienne de Psychologie Experimentale*, **49**, 240-262.

Desrochers, A., & Paivio, A. (1990). Le phonème initial des noms inanimes et son effet sur l'identification du genre grammatical. *Revue Canadienne de Psychologie Experimentale*, **44**, 44-57.

Deutsch, A., Bentin, S., & Katz, L. (1999). Semantic influence on processing gender agreement: Evidence from Hebrew. *Journal of Psycholinguistic Research*, **28**, 515-536.

De Vincenzi, M. (1999). Differences between the morphology of gender and number: Evidence from establishing coreferences. *Journal of Psycholinguistic Research*, **28**, 537-554.

Dominguez, A., Cuetos, F., & Segui, J. (1999). The processing of grammatical gender and number in Spanish. *Journal of Psycholinguistic Research*, **28**, 485-498.

Friederici, A., & Jacobsen, T. (1999). Processing grammatical gender during language comprehension. *Journal of Psycholinguistic Research*, **28**, 467-484.

Gilhooly, K., & Gilhooly, M. (1980). The validity of age of acquisition ratings. *British Journal of Psychology*, **71**, 105-110.

Heim, S., Opitz, B., & Friederici, A. (2002). Broca's area in the human brain is involved in the selection of grammatical gender for language production: Evidence from event-related functional magnetic resonance imaging. *Neuroscience Letters*, **328**, 101-104.

Jacobsen, T. (1999). Effects of grammatical gender on picture and word naming: Evidence from German. *Journal of Psycholinguistic Research*, **28**, 499-514.

Jakubowicz, C., & Faussart, C. (1998). Gender agreement in the processing of spoken French. *Journal of Psycholinguistic Research*, **27**, 597-617.

Jescheniak, J. D. (1999). Gender priming in picture naming: Modality and baseline effects. *Journal of Psycholinguistic Research*, **28**, 729-737.

Jescheniak, J. D., & Schriefers, H. (1999). In search of gender recency: Evidence from the production of pronouns. *Journal of Psycholinguistic Research*, **28**, 739-745.

Miceli, G., Turriziani, P., Caltagirone, C., Capasso, R., Tomaiuolo, F., & Caramazza, A. (2002). The neural correlates of grammatical gender: An fMRI investigation. *Journal of Cognitive Neuroscience*, **14**, 618-628.

Pashov, P. (1994). *Prakticheska bulgarska gramatika* [Practical Bulgarian grammar]. Sofia: Prosveta.

Plemmenou, E., Bard, E. G., & Branigan, H. P. (2002). Grammatical gender in the production of single words: Some evidence from Greek. *Brain & Language*, **8**, 236-241.

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

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

Stojanov, S. (1980). *Gramatika na bulgarskija knizhoven ezik: Fonetika i morfologija* [A grammar of standard Bulgarian: Phonetics and morphology]. Sofia: Nauka.

Vachkov, V. (1997). Formalni modeli na rodova adaptatsija na substantivnite zaemki ot anglijskija v bulgarskija i v cheshki ezik [Formal models of gender adaptation of substantival loans from English in Bulgarian and Czech]. *Supostavitelno ezikoznanie*, **22**, 19-25.

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

Vigliocco, G., & Franck, J. (1999). When sex and syntax go hand in hand: Gender agreement in language production. *Journal of Memory & Language*, **40**, 455-478.

Vigliocco, G., Lauer, M., Damian, M., & Levelt, W. J. M. (2002). Semantic and syntactic forces in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 46-58.

(Manuscript received August 8, 2002;
revision accepted for publication September 4, 2003.)