

**PROCESSING OF GRAMMATICAL GENDER  
IN NORMAL AND APHASIC SPEAKERS OF RUSSIAN**

**Tatiana Akhutina<sup>1</sup>, Andrei Kurgansky<sup>1,2</sup>, Marina Kurganskaya<sup>1,2</sup>, Maria Polinsky<sup>3</sup>,  
Natalya Polonskaya<sup>1</sup>, Olga Larina<sup>4</sup>, Elizabeth Bates<sup>3</sup> and Mark Appelbaum<sup>3</sup>**

<sup>1</sup>Moscow State University, Moscow, Russia; <sup>2</sup>University of California, San Francisco;  
<sup>3</sup>University of California, San Diego; <sup>4</sup>Moscow Speech Pathology Center, Moscow, Russia)

*Cortex*, 2001, 37(3), 295-326.

ABSTRACT

Sensitivity to grammatical gender was investigated in 22 Russian-speaking aphasic patients, compared with young controls. Experiment 1 used a cued shadowing paradigm to assess gender priming (facilitation and/or inhibition of lexical access by a prenominal modifier with congruent, incongruent or neutral gender). Experiment 2 used a grammaticality judgment paradigm with similar stimuli. Normals showed significant interactions between gender and priming in Experiment 1 (facilitation for feminine and neuter nouns but not for masculines) and Experiment 2 (larger effects of context on feminine and neuter nouns) that we interpret as a Markedness Effect. Patients showed significant priming in Experiment 1 and above-chance accuracy in Experiment 2, but failed to show reduced effects for the least-marked masculine gender (the Markedness Effect) in either experiment. Context effects were not related to specific aphasic symptoms or subtypes in either experiment. However, canonical correlation revealed differential effects of specific aphasic symptoms on judgment accuracy (false alarms vs. misses). We conclude that knowledge of grammatical gender is spared in Russian aphasics, but gender processing is deviant. A possible model to account for these differences is discussed.

Key words: psycholinguistics, grammatical gender, aphasia, agrammatism, syntactic processing, grammatical priming, Russian

INTRODUCTION

Effects of grammatical gender on lexical access, or *gender priming*, have been demonstrated across several languages and tasks: in Italian using picture naming (Bentrovato, Devescovi, D'Amico et al., 1999), cued shadowing (repetition of an auditory word target—Bates, Devescovi, Hernandez et al., 1996) and gender monitoring (classification of noun targets as masculine or feminine—Bates et al., 1996); in French using auditory lexical decision and gating (Grosjean, Dommergues, Cornu et al., 1994); in German using picture naming (Hillert and Bates, 1996; Jacobsen, 1999) and cross-modal visual naming (Jacobsen, 1999); in Spanish using picture naming and semantic judgment (Wicha, Bates, Hernandez et al., 1997); in Russian using cued shadowing (Akhutina, Kurgansky, Polinsky et al., 1999), reading time (Taraban and Kempe, 1999) and a forced-choice gender agreement task for verbs (Taraban and Kempe, 1999). All these results indicate that prenominal modifiers embedded in a short auditory phrase or a longer sentence context can prime the nouns they modify. Specifically, modifiers matching in grammatical gender can facilitate lexical access (decreasing reaction time relative to a neutral baseline)

while modifiers with mismatching gender can inhibit or interfere with lexical access (increasing reaction times relative to the same baseline). Such findings for gender complement and extend earlier studies showing that case and gender marking on prenominal adjectives can prime word access in Serbo-Croatian (Gurjanov, Lukatela, Lukatela et al., 1985; Gurjanov, Lukatela, Moskovljevic et al., 1985; Lukatela, Kostic, Feldman et al., 1983).

In contrast with the robust gender-processing effects observed in normal speaker-listeners, studies of grammatical gender in brain-injured patients have yielded mixed results. Offline (untimed) studies by other investigators have shown that some Italian-speaking patients can identify the gender of a target noun at above-chance levels, even though they are unable to retrieve the phonetic form of the word (Badecker, Miozzo, and Zanuttini, 1995). Online (timed) studies using a metalinguistic grammaticality judgment task have also shown that Italian aphasics can detect violations of gender agreement (Devescovi, Bates, D'Amico et al., 1997; see also Lu, Bates, Li et al., 2000, for preserved classifier-noun agreement in Chinese aphasics). These results would indicate that

gender knowledge is preserved in aphasia (including both fluent and nonfluent patients in the study by Devescovi et al.). However, a recent study of Italian patients using the gender-monitoring task failed to uncover evidence for gender priming, even though the same patients (both fluent and nonfluent) were able to classify nouns as masculine or feminine, and detect violations of adjective-noun agreement at above-chance levels (Bates, Marangolo, Pizzamiglio et al., 2000). Hence the real-time effects of gender may be fragile in aphasic patients, a result similar to findings in the literature on semantic priming in aphasia. Some semantic-priming studies have shown loss or attenuation of priming, especially in nonfluent Broca's aphasics (Milberg, Blumstein and Dworetzky, 1987, 1988a, 1988b; Milberg, Blumstein, Katz et al., 1995; Prather, Shapiro, Zurif et al., 1991; Prather, Zurif, Stern et al., 1992; Utman and Bates, 1998; Utman, Blumstein and Sullivan, 1998). However, other studies have shown that semantic priming is preserved in aphasia, at least for some patients, tasks, materials and timing parameters (Swaab, Brown and Hagoort, 1997, 1998; Tyler, Ostrin, Cooke et al., 1995). Given this range of results for semantic priming, we should not be surprised to find that gender-priming effects also "come and go" under different methodological conditions.

To learn more about gender processing in aphasia, we need an expanded database that includes a broader range of languages, and a richer array of gender systems. For reasons that we will explain in more detail below, an investigation of gender processing in Russian aphasics could be especially informative regarding the interplay between grammatical and lexical processing in aphasia. In this paper, we present and analyze the results of gender priming and gender processing in 22 Russian-speaking aphasic patients and controls, using short modifier-noun phrases, in two tasks: cued shadowing (also called auditory word naming) and grammaticality judgment (also called violation detection). The cued shadowing task meets most of the criteria that have been proposed to define and elicit automatic priming (short time windows, relatively fast latencies, no metalinguistic component, with a neutral baseline to permit inferences about facilitation vs. inhibition). The grammaticality judgment task was chosen to yield information about the patients' knowledge of gender class and gender agreement, and the time required to detect violations of agreement. Before these experiments are presented, we must briefly review relevant properties of the Russian gender system.

#### OVERVIEW OF RUSSIAN GENDER

Russian has three main genders: masculine, feminine, and neuter (see Corbett, 1982, 1991 for the so-called subgenders). Masculine nouns constitute about 46% of the nominal lexicon, feminines are at 41%, and neuters are at 13% (statistics based on the 34,000-word

count; see also Comrie, Stone and Polinsky, 1996, p. 109). Despite the predominance of masculine and feminine nouns, neuter is a robust category in Russian: about 10% of new words and recent borrowings are neuters, whereas 48% are masculine nouns and 42% are feminine nouns (Comrie et al., 1996, p. 107). Noun assignment to the masculine or feminine is determined by semantics for the majority of nouns denoting humans and for some animate nouns (biological gender of the referent). For the rest of the nouns, their assignment to one of the three genders is determined by their declensional type and phonology. Russian has four main declensional types with different endings in six noun cases. For many but not all nouns, their ending in the nominative singular acts as an indicator of their declensional type and consequently gender. The following generalizations concerning the nominative singular approximate the basic gender assignment principles (for details, see Corbett, 1991; Comrie et al., 1996, pp. 105-112):

- (a) nouns ending in a nonpalatalized consonant (zero-ending) are masculine
- (b) nouns ending in *-a* are feminine
- (c) nouns ending in *-o*, *-e* are neuter

We will refer to such nouns, which "wear their gender on their sleeve" as phonologically transparent. A robust minority of nouns are "phonologically opaque", with endings in the nominative that do not predict their gender. For example, nouns ending in a palatalized consonant are either masculine or feminine, depending on their declension type. The two experiments reported below were restricted entirely to phonologically transparent nouns. We note, however, that Akhutina et al. (1999, Experiment 3) report significant gender priming with opaque Russian nouns. We will return to this issue in the final discussion below.

In Russian, agreement with the gender of a given noun is manifested on the following forms: (prenominal) modifying adjectives and possessive pronouns; modifying participles; past tense form of verbs; some numerals. In the experiments below, the crucial agreement type is the first one, in which a modifying adjective or possessive pronoun agrees with the head noun. Participants were presented with modifier-noun pairs where the first constituent was a possessive pronoun. Three pronouns that we used had overt and unambiguous gender agreement, as shown in (1), and three pronouns had homophonous forms for all the genders, as shown in (2):

- (1) Declinable possessive pronouns
  - a. *moj* (masc.), *mojá* (fem.), *mojó* (neuter) 'my'
  - b. *tvoj* (masc.), *tvojá* (fem.), *tvojó* (neuter) 'your (sg.)'
  - c. *ničéj* (masc.), *nič'já* (fem.), *nič'jó* (neuter) 'no one's'

- (2) Nondeclinable possessive pronouns
  - a. *jejó* ‘her’
  - b. *jego* ‘his’ (pronounced *jevo*)
  - c. *ix* ‘their’

The pronouns in (1) were used in our experiments to create gender-concordant or gender-discordant modifier-noun pairs, while the pronouns in (2) served as a gender-neutral baseline. For example, a concordant item might be *moj* (masc.) *bilet* (masc.) ‘my ticket’, a discordant item might be *moja* (fem.) *bilet* (masc.) ‘my ticket’, and a neutral item might be *jejo bilet* ‘her ticket’.

#### EXPERIMENT 1: CUED SHADOWING

Experiment 1 was designed to tap into automatic gender priming, using the cued shadowing task (repetition of auditory target words—Bates and Liu, 1996). This technique elicits rapid reaction times in normals, including facilitation relative to a neutral baseline, and it requires no metalinguistic judgments or other operations irrelevant to lexical access. Performance by aphasic patients is first compared with the “optimal profile” displayed by a control group of healthy young adults. Then patients are compared with each other along theoretically relevant symptom dimensions.

#### *Materials and Methods*

##### *Participants*

Twenty-two aphasic patients took part in the experiments. Twenty-one patients were from the Moscow Speech Pathology Center, and one was from the Neurology Clinic, Moscow Medical Academy. Demographic and neurological information is summarized in Table I. These patients represent a broad range of age and etiology. However, all were at least 6 months beyond lesion onset, and all had participated in an intensive course of speech-language therapy prior to testing. Linguistic performance close to time of testing was assessed by the Method of Speech Assessment in Aphasics (Tsvetkova, Akhutina, Polonskaya et al., 1979; Tsvetkova, Akhutina and Pylaeva, 1982). Results of this assessment are summarized in Table II, divided into receptive and expressive scales. In addition, Table II contains a separate expressive agrammatism score based on the patients’ descriptions of the Cookie Theft picture from the Boston Diagnostic Aphasia Examination (Goodglass and Kaplan, 1968). This agrammatism scale assigns patients one of four levels: (1) single-word speech only (no connected discourse), (2) severe agrammatism (characterized by speech fragments and frequent omissions, although some nonformulaic phrases and sentences may be present), (3) mild agrammatism (connected speech but with some omissions and grammatical substitutions), and (4) normal fluency (with or without lexical symptoms, e.g. circumlocutions, phonemic and semantic paraphasias).

For one patient (# 8), the Cookie Theft session was missing; in this case the agrammatism score was assigned based on other free-speech evidence, including the “story” component of the language battery in Table II. Total aphasia severity scores indicate that all patients were in the mild to moderate range. The observed fluency range was 2-4 (i.e. there were no patients with only single-word speech). For the fluency analyses below, we maximized power by combining patients with a ranking of 2 or 3 as “nonfluent” and patients with a ranking of 4 as fluent.

Based on a full neuropsychological battery that includes nonlinguistic measures of memory and sensory processing, patients were classified into traditional categories described by Luria and Jakobson (Akhutina, in press; Jakobson, 1964; Luria, 1970; see also Akhutina and Glozman, 1995; Ryabova (Akhutina), 1967). Seventeen patients were classified as motor aphasics, 10 as sensory aphasics, and 5 as semantic aphasics. These classifications are often translated approximately into the classic Western categories of Broca’s aphasia (motor), Wernicke’s aphasia (sensory) and anomic aphasia (semantic). However, as can be seen from examination of Table II, the sensory and semantic patients did not differ substantially in their receptive language abilities ( $p > .05$  by a 2-tailed t-test on all receptive subscales in Table II). For comparison with the Luria taxonomy, a second diagnosis designed to approximate Western classifications was made based on all the linguistic subscales in Table II. Nonfluent patients (fluency scores of 2 or 3) were classified as Broca’s aphasics. Based on a median split on the total comprehension scores for these 22 patients, fluent patients (fluency = 4) with low comprehension were classified as Wernicke’s, and fluent patients with high comprehension were classified as anomic aphasics. These classifications are also listed in Table II.

Twenty college students (10 males and 10 females) also took part in Experiment 1 as a control group.

##### *Materials and Design*

Russian nouns of all three genders were used, in three conditions: concordant, neutral and discordant. All the nouns were gender transparent (see above). Each item consisted of an auditory modifier-noun pair in which the first constituent was a possessive pronoun. For the concordant and discordant conditions, possessive pronouns were overtly marked for gender. For the neutral-control condition, we used nondeclinable possessive pronouns that are overtly ambiguous for gender (see (2a-c) above). A complete list of stimuli is provided in the Appendix.

The experiment contained 216 possessive-pronoun/noun pairs. Within each pair, noun targets began immediately after the offset of the pronoun prime (Stimulus Onset Asynchrony = 0 ms). Within the constraints of the design, pairs were constructed randomly from a list

of 72 nouns (24 masculine, 24 feminine, and 24 neuter nouns—see Appendix), which ranged from 4-6 phonemes in length (the only exception is *uxo* ‘ear’, 3 phonemes), and 12 possessive pronouns (see (1a-c) and (2a-c) above). Possessive-pronoun modifiers and nouns were audio-recorded and digitized separately. To achieve the natural intonation of Russian modifier-noun phrases, possessives were recorded in a rising tone and nouns in a falling tone. Mean waveform duration for the target nouns was 531 ms (s.d. = 88). The respective masculine, feminine, and neuter nouns were balanced for waveform duration [feminine, 522 ms, s.d. = 91; masculine, 553 ms, s.d. = 89; and neuter, 517 ms, s.d. = 83 —  $F(2, 69) = 1.1937, p > .3$ ], frequency of use, based on Zazorina (1977) and Štejnfeldt (1963) [feminine, 197.7, s.d. 260.4; masculine, 173.9, s.d. 195.1; neuter, 191.1, s.d. 235.4 —  $F(2, 69) < 1, p > 0.1$ ], and word length in syllables [feminine, 4.6667, s.d. = .7614; masculine, 4.7083, s.d. = .7506; neuter, 4.6250, s.d. = .8242].

The overall design for Experiment 1 was a 3 (masculine, feminine, neuter noun)  $\times$  3 (gender-concordant, gender-discordant, gender-neutral possessive pronoun), with 24 items in each of the nine cells of the design. Within the constraints of this design, possessive pronouns and nouns were chosen at random by the PsyScope shell to produce unique pairings for each subject (Cohen, MacWhinney, Flatt et al., 1993). Because individual nouns were uniquely and randomly assigned to conditions for each participant, analyses of variance over items would be inappropriate (and unnecessary, since the individual randomization process assures that specific noun items are completely unconfounded with priming conditions). All analyses based on this design will be conducted over subjects only.

### *Procedure*

Pronoun-noun pairs were presented auditorily via loudspeakers, at a comfortable amplitude. The trial did not end until a response was registered (but see scoring, below, for data trimming). 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 are 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.

### *Scoring and Data Analysis*

There were no trials on which participants failed to register a response, but response times were occasionally very long (especially for aphasic patients). Prior to data analysis, all outlying response times ( $> 1500$  ms from target onset) were eliminated (see Akhutina et al., 1999 for details). The 1500-ms criterion was determined from examination of the RT distribution ( $> 2$  standard deviations from the group mean). Trials with frank repetition errors were also eliminated. Means for each participant, for each cell, were based on the remaining items.

### *Results and Discussion: Normal Controls*

#### *Errors*

Normal controls made very few errors on this task. The average error rate (including repetition errors and RTs exceeding 1500 ms) was below 2%, and no individual normal control exceeded 3%. Therefore the error data were not subjected to further analyses.

#### *Reaction times*

Collapsed across all cells of the design, the mean for normals was 692 ms (s.d. 139, s.e. = 31). This mean RT value is close to the means observed in four previous experiments with young normals (Akhutina et al., 1999). When measured from the end of the word, RTs averaged 161 ms, which is compatible with values reported by Bates et al. (1996) for Italian and Akhutina et al. (1999) for Russian.

We conducted a 3 (masculine, feminine, neuter target)  $\times$  3 (concordant, neutral, discordant possessive pronoun) within-subjects analysis of variance on these reaction times. In this analysis (and in all analyses reported below), alpha levels are based on Greenhouse-Geisser corrections for heterogeneity of variance. This analysis of normal controls yielded a significant main effect of gender congruence [ $F(2, 18) = 29.80, p < .0001$ ], and a significant congruence by noun gender interaction [ $F(4, 16) = 11.15, p < .0001$ ]. No main effect of noun gender was found [ $F(2, 18) = 1.67, p > .21$ ]. Hence young normals do not display a general advantage (independent of context) for unmarked masculine nouns. Cell means are summarized in Table III.

The main effect of congruency reflects an overall average of 25.5 ms of priming (discordant – concordant). As expected, RTs for indeclinable pronouns fell midway between discordant and concordant conditions, which means that these pronouns do constitute an appropriate neutral baseline. Planned comparisons confirmed that the facilitative component (neutral – concordant) of 16.5 ms was significant ( $F = 24.15, p <$

.0002), as was the small inhibitory component (discordant – neutral) of 9.0 ms ( $F = 7.28, p < .02$ ).

Inspection of Table III shows that the congruence by gender interaction is due primarily to the difference between masculine items and the other two genders. Specifically, masculine nouns (the unmarked gender) showed much smaller effects of gender context, and no evidence at all for gender facilitation. For masculine nouns, planned contrasts indicate that the small overall priming effect did reach significance ( $F = 7.03, p < .02$ ). However, this effect comes entirely from the inhibitory component ( $F = 10.56, p < .005$ ). For feminine nouns, the overall priming effect was significant ( $F = 82.88, p < .0001$ ), as were both the facilitative component ( $F = 37.44, p < .0001$ ) and the inhibitory component ( $F = 8.91, p < .01$ ). For neuter nouns, the overall priming effect was also significant ( $F = 73.45, p < .0001$ ), although this large effect came entirely from gender facilitation ( $F = 57.75, p < .0001$ ), and not from inhibition ( $F < 1.0, n.s.$ ). Figure 1 illustrates this interaction using facilitation and inhibition differences scores for each of the three genders.

#### *Interim conclusions*

We found significant gender-priming effects in young Russian-speaking adults using the cued shadowing paradigm. These results are compatible with findings for four other cued shadowing experiments by Akhutina et al. (1999), using various combinations of Russian gender stimuli. Like the results of Experiment 4 in Akhutina et al. (which is closest to the present experiment in materials and design—see final discussion), gender priming reflects a combination of facilitation (by a concordant modifier) and inhibition (by a discordant modifier), relative to a gender-neutral baseline.

We also found an interaction between grammatical context and noun gender for normal listeners in this experiment, which differs from the patterns that have been reported to date for two-gender languages. In our view, this effect can be explained in the following way. From the standpoint of class size and inflectional endings in the nominative, the masculine is the unmarked gender in Russian. It is numerically the largest class of nouns, and it is the one in which most words in the nominative have a zero ending. For the phonologically transparent masculine nouns that we have used here, the effects of a gender-marked modifier appear to be minimal: there was no evidence for gender facilitation, although a small inhibitory effect did reach significance. In contrast, the feminine and neuter nouns both showed robust gender facilitation, and feminine nouns also showed significant inhibition. These are the results that we would expect if normal listeners start each new trial with the assumption that the next noun they hear will belong to the default class, i.e. masculine. Hence a concordant masculine modifier does not move

them any closer to the upcoming noun than they would be with no gender-relevant information at all. By contrast, the time required to recognize and reproduce a feminine or neuter noun can be facilitated by gender-congruent information.

The same logic applies to the finding that inhibition is significant for feminines but not for neuters. That is, because neuters are the least likely of the three Russian genders, the default assumption may be that a neuter noun will *not* occur. Hence it is possible to facilitate neuters beyond baseline expectations, but much more difficult to inhibit them beyond baseline. Results for feminine nouns fall in between the masculine and neuter extremes. That is, feminines can be facilitated relative to baseline (because they are not the default form), but they are also common enough in the Russian language that they are subject to inhibition. In what follows, we will refer to the asymmetry between genders in terms of their class size, token frequency, and the marking of transparent nouns that belong to a given gender as the Markedness Effect. The main idea behind this concept is that nouns in larger/more frequent/least morphologically marked gender classes are more likely to occur, thus their non-occurrence leads to inhibition; nouns in smaller/less frequent/morphologically marked gender classes are unlikely to occur, so their occurrence leads to facilitation. The Markedness Effect allows us to separate the masculine from the two other genders, but as we just explained with respect to the feminine, the Markedness Effect may also involve more fine-grained expectations, which can result in differences between feminines and neuters as well.

An everyday metaphor to capture these effects can be found in the simple two-dimensional space defined by a tennis court: all other things being equal, a good player moves back to center court as quickly as possible after each move, because that is the position from which one can reach the largest number of volleys in the smallest amount of time. In the same vein, our normal listeners are able to use gender context to anticipate the region in which a target word will occur. However, a competent listener (like a competent tennis player) moves back to the central (unmarked) region when each trial is over, in order to maximize efficiency when the gender of the next target is unknown. For this reason, masculine gender context has little effect, and the least common forms (neuter nouns) are the ones that profit most from context.

This tennis metaphor can be related to the connectionist framework proposed by Elman (1990) for syntactic processing, and extended to gender by Betrovato et al. (1999) and Wicha et al. (1997). Within this framework, the lexicon is defined as a high-dimensional space of both sound and meaning. Every lexical form occupies a unique position in this high-dimensional space, and similarity (semantic, phonological and morphosyntactic) is defined in terms of relative distance

between items along multiple dimensions. Lexical access can be viewed as movement through space into the position that defines the target word. Prior information can be said to prime lexical access if it moves the listener closer to or farther away from the position occupied by the upcoming target. An unmarked/default position is defined as a relatively central position or region within this hyperspace (“center court”), from which it is relatively easy to arrive at a large number of possible destinations (assuming no prior information). We will return to the tennis court metaphor and the theoretical framework behind it after we examine the data for aphasic patients in the same experiment.

Before proceeding, we note that these results for normals are compatible with (though not identical to) results for normals in the four cued shadowing experiments reported by Akhutina et al. (1999). The morphophonological characteristics of Russian modifiers and nouns make it extremely difficult to set up a full congruence (congruent, incongruent, neutral) by phonological transparency (opaque vs. neutral) by three-gender design in Russian. Opaque nouns do exist for all three genders, but are rare and often low in frequency, thus a full design would lead to confounds in length and frequency that would complicate interpretation of results. For this reason, Akhutina et al. approached these issues one at a time, in separate experiments. In Experiment 1, they used phonologically transparent masculine and feminine nouns with phonologically transparent adjectival modifiers, with no neutral baseline. Significant priming was observed for both masculine and feminine nouns, but there was also a significant priming by gender interaction reflecting more priming for feminine items—in line with the results that we have reported here. In Experiment 2, they used feminine and neutral nouns only, with a neutral adjectival baseline that would not work for masculine nouns. Hence this experiment did not address the markedness issue that concerns us here. Experiment 3 used phonologically opaque masculine and feminine nouns, but there was no neutral baseline. Masculines were faster than feminines, and significant priming was observed for both genders, but there was no interaction (in contrast with Experiment 1). Experiment 4 in the Akhutina et al. (1999) study is the closest to the present study in design and purpose. That experiment used phonologically transparent masculines and feminines only. Primes were restricted to versions of the Russian word for ‘simple’: *prostoj* (masculine zero-marked adjective), *prostaja* (feminine marked adjective) and *prostoj* (an adverbial form that is not marked for gender, and should be compatible with any noun). Results included a significant main effect of gender (masculines faster than feminines—a result that we did not observe here), a significant main effect of priming (congruent < neutral < incongruent—similar to results of the present study). In contrast with the

present study, the gender by priming interaction was not significant. However, planned comparisons indicated that masculine nouns elicited significant inhibition but no facilitation, while feminine nouns elicited significant facilitation but no inhibition. This last result is quite consistent with findings in the present study. The current design is superior insofar as it permitted a consideration of all three noun genders, with the same neutral baseline. Unfortunately, given the limited number and range of phonologically opaque nouns in Russian (especially feminines), a complete 3 (congruent, incongruent, neutral) by 3 (masculine, feminine, neuter) by 2 (opaque vs. transparent) design was simply not practical. In the final discussion, we will consider whether phonological similarity could have contributed to priming effects for normal controls and/or aphasic patients.

### *Results and Discussion: Aphasic Patients*

#### *Errors*

Two patients (# 8 and # 20 in Tables I-II) were unable to complete the task with enough response times below the 1500-ms criterion. Omissions, repetition errors and/or delayed responses for the remaining 20 patients did not exceed 20% and were not processed statistically.

#### *Reaction Times*

Reaction time analyses were based on the cell means for all correct trials (excluding omissions, errors and outlying RTs). We will begin by reporting results for aphasic patients as a group, and follow this with a series of exploratory analyses to determine whether specific aphasic syndromes or symptoms contribute to these priming results. These will include two  $3 \times 3 \times 3$  analyses partitioning the group by aphasia classification (the Luria system, and the Western system). These will be followed by a succession of  $2 \times 3 \times 3$  analyses partitioning aphasic patients by fluency (nonfluent vs. fluent) and by median splits along the other major symptom dimensions in Table II. The symptom dimensions that we will use in these analyses include total aphasia severity (total production + total comprehension), total production, total comprehension, a lexical production subscore (noun naming + verb naming), a lexical comprehension subscore (noun comprehension + verb comprehension), a grammatical production subscore (sentential descriptions of single pictures + story telling), and a grammatical comprehension subscore (sentence comprehension + comprehension of commands). Finally, we will present multivariate results using factor analysis to reduce the number of predictors to a manageable set, followed by canonical correlation analyses to determine whether there is a significant mapping between aphasia symptom factors and performance in the priming experiment.

First the reaction time data were subjected to a 3 (gender congruence) by 3 (noun gender) within-subjects analysis of variance for all 20 patients. This analysis yielded a significant main effect of gender congruence or priming [ $F(2, 18) = 26.82, p < .0001$ ], and a smaller main effect of noun gender [ $F(2, 18) = 3.73, p < .041$ ]. In contrast to results for normal controls, the interaction between congruence and gender in the aphasic patients was not significant [ $F(4, 16) = 1.17, n.s.$ ]. In other words, these patients do not demonstrate the Markedness Effect observed in normals.

Table III presents cell means for the aphasic patients. These are broken down by congruence and noun gender, for comparison with the significant interaction observed in normal controls, but we must underscore that the interaction for patients did not reach significance. Planned contrasts indicate that the gender-priming effect for aphasics includes both significant facilitation (neutral – concordant,  $F = 21.66, p < .0001$ ) and significant inhibition (discordant – neutral,  $F = 6.61, p < .015$ ). The main effect of noun gender for the aphasic patients reflects faster RTs on masculines, slower RTs on feminines, with neuters falling in between—a result that is only partially compatible with a markedness hypothesis (which would predict masculine < feminine < neuter). In contrast with results for normal controls, Table III shows that the masculine nouns elicited faster RTs than the other two genders in all three priming conditions. Finally, Table III shows that priming effects for aphasic patients are roughly equivalent in magnitude for all three genders (masculine = 51.5 ms, s.e. = 13.45; feminine = 34.1 ms, s.e. = 8.0; neuter = 38.4 ms, s.e. = 11.23). Planned contrasts show that all three of these priming effects are significant ( $F = 24.85, p < .0001$  for masculines;  $F = 10.88, p < .006$  for feminines;  $F = 13.79, p < .003$  for neuters). Hence, in the absence of a Markedness Effect, which reduces priming for masculine nouns in young controls, these aphasic patients actually show larger-than-normal gender priming overall.

To compare results for patients and controls directly, we conducted an omnibus 2 (group)  $\times$  3 (congruence)  $\times$  3 (noun gender) analysis of variance, with group as a between-subjects factor, congruence and gender as within-subjects factors. Not surprisingly, this result yielded a large main effect of group [ $F(1, 38) = 29.145, p < .0001$ ], reflecting substantially slower RTs in aphasic patients. There was also a large main effect of congruence [ $F(2, 38) = 52.23, p < .0001$ ], a small but significant main effect of gender [ $F(2, 38) = 4.77, p < .015$ ], and a significant three-way interaction of group by congruence by gender. The main effect of gender reflects faster reaction times for masculine nouns, slower reaction times for feminine nouns, with neutral nouns falling in between. The three-way interaction is the most important result for this omnibus analysis, because it confirms that the gender-priming profile is

significantly different for normals and aphasic patients. Specifically, aphasic patients failed to display the interaction between gender and context that we found in healthy young controls who are “optimal listeners” in Russian. This contrast between patients and controls is illustrated in Figure 2, using difference scores for facilitation vs. inhibition within each gender.

Having established that gender priming does exist in these aphasic patients, and that the priming profile appears to be aberrant relative to performance by young controls, we then carried out a series of exploratory analyses of variance for aphasic patients only. In each analysis, the 20 patients were partitioned along one of the classifications or dimensions in Table II. We can summarize results of these analyses briefly, concentrating on main effects and interactions involving the between-subject groupings. First, a main effect of priming was confirmed in each of the analyses (all  $F > 20.0$  and all  $p < .0001$ ), but this priming effect did not interact with any of the patient groupings (all  $F < 2.23$ , and all  $p > .12$ ). Hence gender priming per se seems to be independent of aphasia type or specific aphasia symptoms. Second, there were no significant interactions between group and gender (all  $F < 1.53$ , all  $p > .22$ ). Third, the gender by congruency interaction (i.e. the Markedness Effect) failed to appear in any of the analyses (all  $F < 2.00$  and all  $p > .13$ ), nor did we ever obtain a significant group by gender by congruency interaction (all  $F < 1.27$  and all  $p > .29$ ). In other words, we cannot use aphasia classifications (including median splits on classic symptom dimensions) to predict who will or will not show priming effects, and/or the markedness interaction.

We did find three main effects of group on cued shadowing reaction times: in the fluency analysis ( $F = 18.8, p < .0001$ ), in the analysis using the Luria classification ( $F = 12.52, p < .0001$ ) and in the analysis using the Western classification ( $F = 9.43, p < .002$ ). All of these results are due to the same pattern: non-fluent patients display substantially slower cued shadowing latencies (mean = 1184 ms, s.d. = 154) compared with the more fluent groups (mean = 891 ms, s.d. = 132). Post hoc t-tests comparing Wernicke’s and anomics did not reach significance by either classification [ $t(12) = 2.07, p < .06$  for the Luria classification;  $t(12) < 1, n.s.$ , for the Western classification].

It is of course not surprising to find that nonfluent aphasics are slow in a task that requires repetition of target words. Indeed, these patients tended to perform quite slowly and deliberately, sometimes chanting the words one syllable at a time. To determine whether gender priming reached significance even within these subgroups (including nonfluent patients), we conducted the 3  $\times$  3 within-subjects analysis separately, within each of the three syndromes (based on the Western classification). The main effect of priming reached significance in all three groups ( $p < .001$  for Broca’s,

Wernicke's and anomics, respectively). Within the nonfluent Broca group (comprising only 6 patients), planned comparisons for the facilitative and inhibitory components of priming did not reach significance. Within the other two groups (6 Wernicke's, and 8 anomics) both the facilitative and the inhibitory effects were significant. None of the groups displayed a congruency by gender interaction. Because separate estimates of facilitation and inhibition failed to reach significance in Broca's aphasics, one might conclude that gender priming is less reliable in this group, in line with previous studies demonstrating weak or absent priming effects in Broca's aphasia. However, in the absence of a significant group by priming interaction, this hypothesis has very weak support.

In a further attempt to characterize those patients who failed to show the normal gender-priming profile, we divided the 20 patients into three groups based on their cell means: (1) a group who showed no priming at all (discordant – concordant < 10 ms, N = 3); (2) a group who resembled the mean performance by young normals in showing overall priming, as well as smaller priming for masculine nouns compared to the average for feminine and neutral nouns, N = 9; (3) a group who also showed priming but displayed priming for masculines that was equal to or greater than priming for the other two genders — N = 8), thus no Markedness Effect. We then conducted one-way analyses on each of the variables in Table II, using this three-level priming classification as an independent variable. None of these analyses reached significance. Hence the absence of the normal gender by congruence interaction in aphasic patients does not seem to be due to any specific symptom or syndrome.

A problem with all the subgroup analyses that we have conducted so far is that they involve an arbitrary division along one or more continuous dimensions. This is true of the classic aphasia groupings and all our median splits on specific symptom dimensions, and it is also true of the division that we have just described of patients who do or do not show normal priming or a normal Markedness Effect. With any of these arbitrary divisions, potentially important information may be lost. It would therefore be useful to apply a technique that preserves continuous information on both sides of the analysis: continuous information about patient symptoms, and continuous information about behavior within the experimental design. One analytic approach that preserves this continuous information is canonical correlation, which tests for significant relationships between one set of variables (the symptom space) and the various main effects and interactions in the  $3 \times 3$  within-subject design (linear contrasts). This is the last approach that we tried here to uncover potential causes for the aberrant priming profile observed in these Russian aphasics.

We began by conducting a factor analysis (principal-components analysis) to determine whether the 10 aphasia scales reduce to a more manageable set of latent variables. Results of this factor analysis are summarized in Table IV. Three latent factors with eigenvalues greater than one emerged from this analysis. An examination of the factor loadings in Table IV suggests that the first factor (accounting for 56.4% of the variance) loads highly on the various word and sentence production scales. We will refer to this as the Production Factor, although it is interesting to note that this factor also strongly loads on one measure of sentence comprehension (compatible with the suggestion that sentence comprehension sometimes involves covert production). The second factor (accounting for 14.4% of the variance) loads primarily on the various word and sentence comprehension scales, and will be referred to as the Comprehension Factor. The third factor (accounting for 10.4% of the variance) receives its greatest contributions from single-word comprehension and production scales. We will refer to this as the Lexical Factor, although we note that this factor also loads highly on comprehension in dialogue (suggesting that such comprehension in context may be strongly based on lexical as well as grammatical processing).

These three factors were then used in a canonical correlation analysis, to determine their relationship to critical contrasts in the  $3 \times 3$  priming by noun gender design. For this purpose, the following linear contrasts were selected: (1) two gender contrasts within the main effect of gender (masculine vs. feminine; masculine vs. neuter), (2) two gender contrasts within the main effect of congruency (concordant vs. neutral; discordant vs. neutral), and (3) four contrasts that represent the key components of the gender by congruency interaction (masculine-feminine by concordant-neutral; masculine-feminine by discordant-neutral; masculine-neuter by concordant-neutral; masculine-neuter by discordant-neutral). The omnibus canonical correlation failed to reach significance (likelihood ratio,  $p = .102$ ). Separate assessments also failed to reach significance for gender only (likelihood ratios for each contrast,  $p > .80$ ) and priming only (likelihood ratios for each contrast,  $p > .50$ ), and for the contrasts that define the interaction (likelihood ratios,  $p > .30$ ). At this point, we must admit defeat: *we cannot account for the aphasia profile in gender priming by any of the specific symptom patterns available for these 20 patients.*

#### *Interim Conclusions*

In contrast with previous reports for Italian, we found significant gender priming in these Russian-speaking aphasics. However, the profile of priming for patients appears quite different from the one that we observed in healthy young controls (i.e. inhibition only for the unmarked masculine gender; significant facilitation for the other two genders). This deviance in



gender priming was unrelated to any of the classic aphasia symptoms and/or classifications listed in Table II, nor did it correlate with the continuous aphasia factor scores.

To account for the Markedness Effect in normal controls, we offered the metaphor of a tennis court. A competent Russian listener operates like a skilled tennis player on a two-dimensional court: after each volley, the player returns to a relatively central position, so that s/he can react to the largest number of volleys in the shortest possible time. To explain the absence of this Markedness Effect on gender priming in aphasic patients, we can make use of the same metaphor. An unskilled tennis player fails to anticipate her partner's volleys. This may include lingering too long in the area where the last ball fell, and failing to return to center court. We know that aphasic patients are prone to perseveration, demonstrating various kinds of interference from previously activated words (Dell, Schwartz, Martin et al., 1997). In much the same fashion, these aphasic listeners may not snap back to an optimal (unmarked) position between trials, due perhaps to lingering activation from the last noun that they encountered. This is of course a speculative account, but it may serve as a working hypothesis for future research as it brings together the phenomenon of aphasic perseveration and the concept of the default or unmarked gender, which in Russian is the masculine.

For all the reasons outlined in the introduction, priming effects observed with this cued shadowing technique meet criteria for the establishment of automatic priming. Experiment 2 was designed to learn more about sparing and impairment of gender processing in these aphasic patients, using a metalinguistic violation detection task that requires a different kind of processing.

## EXPERIMENT 2: GRAMMATICALITY JUDGMENT

### *Materials and Methods*

Experiments 1 and 2 used the same overall design. However, the grammaticality judgment task is not a priming experiment, because the correct answer in the discordant condition is “no” while the correct answer in the concordant condition is “yes,” responses that may tap into very different cognitive operations. We include the concordant/discordant comparison as a factor in our analyses not as an estimate of priming, but to determine whether response profiles differ across the three genders for correct acceptances (“yes”) and correct rejections (“no”).

### *Participants*

The same group of aphasic patients that took part in the cued shadowing experiment (Exp. 1) also participated in this experiment (see Tables I and II for details). A different group of college controls was used

for Experiment 2 (N = 21 college students, 9 males and 12 females).

### *Materials and Design*

A new set of noun stimuli was selected for Experiment 2, to avoid stimulus repetition effects for the aphasic patients. Stimuli for this study were 216 individually randomized pronoun-noun pairs, constructed from a new list of 72 nouns (24 masculine, 24 feminine, and 24 neuter nouns) and 12 pronouns (the same as in Experiment 1). A complete list of stimuli is presented in the Appendix. The masculine, feminine, and neuter nouns were balanced for frequency, [236.7, s.d. = 259 for feminine; 168.6, s.d. = 161.3 for masculine; 212.7, s.d. = 222.1 for neuter;  $F(2, 69) < 1$ ,  $p > 0.5$ ], and word length in phonemes ranging from four to seven phonemes, each size represented equally across the three genders [5.0000, s.d. = .8847, 5.0417, s.d. .8065, 4.9167, s.d. .9286, respectively;  $F(2, 69) < 1$ ]. However, in contrast with Experiment 1, we were unable to balance the Experiment 2 stimuli for waveform duration [574, s.d. = 85 ms for feminine, 602, s.d. = 69 for masculine, and 522, s.d. = 101 for neuter,  $F(2, 69) = 5.32$ ,  $p < .007$ ]. Hence, all other things being equal, we should expect a main effect of gender reflecting slower RTs (measured from word onset) for masculine forms.

The overall design for Experiment 2 was identical to Experiment 1, although different items were employed: a 3 (masculine, feminine, neuter noun)  $\times$  3 (gender-concordant, gender-discordant, gender-neutral pronoun), with 24 items in each of the nine cells of the design. Pronouns and nouns were chosen at random by the PsyScope shell to produce unique pairings for each subject. The neutral baseline items were included in Experiment 2 to maximize the procedural match between Experiments 1 and 2). Note, however, that neutral items have a very different status in the grammaticality judgment task. For the concordant and discordant conditions, it is not possible to classify an item as “good” or “bad” until the noun is recognized. For the neutral condition, any legal noun can serve as a grammatical completion; hence it is possible to classify neutral stimuli as “grammatical” as soon as processing of the possessive pronoun is complete, even before the onset of the target noun. For this reason, they are relatively uninformative for our purposes here. Hence we will treat neutral modifier-noun pairs as fillers and exclude them from the analyses of errors and reaction times described below.

### *Task and Procedure*

In Experiment 2, participants were tested across four sessions (54 stimuli per session). All participants were tested individually, in a quiet room. Participants were instructed to press the green button if the phrase is “correct” and the red button if the phrase is “incorrect”.

Pronoun-noun pairs were presented auditorily via loudspeakers, at a comfortable amplitude. They were allowed to respond until a response was registered. Each trial was followed by a fixed intertrial interval of 1000 ms. The main experiment was preceded by a short practice session (10 trials). The stimuli used for the practice trials did not occur in the experimental trials.

#### *Scoring and Data Analysis*

There were no cases in which the participant failed to respond at all. Because grammaticality judgment requires substantially longer latencies than cued shadowing, the criterion used to eliminate outlying data points was longer than the criteria applied in Experiment 1. For controls, any response exceeding 3000 ms was eliminated. For patients, we dropped any response longer than 6000 ms.

#### *Results and Discussion: College Controls*

##### *Errors*

One individual made a very large number of errors across the board, suggesting that she had failed to attend carefully to instructions. Data for this individual were excluded, and all analyses presented below are based on the remaining 20 controls.

With the exclusion of the single subject just described, normal controls made very few errors on this task. The average error rate (including incorrect button choice, failure to respond, or RT exceeding 3000 ms) ranged from 0 to 2 errors for any individual participant. These results were not subjected to further statistical analysis; reaction time analyses below are based on correct responses only.

##### *Reaction Times*

Collapsed across the concordant and discordant conditions, the mean RT from word onset for normals was 1033 ms (s.d. = 162). Taking mean word duration into account, this means that participants were responding an average of 400-500 milliseconds after the end of the target word, considerably slower than the RTs observed in cued shadowing with similar stimuli.

Results of the 3 (masculine, feminine, neuter target)  $\times$  2 (concordant, discordant conditions) within-subjects analysis of variance yielded a significant main effect of concordance [ $F(2, 18) = 25.01, p < .0001$ ], reflecting longer latencies on discordant items (a “no” response) compared with concordant items (a “yes” response). This is the direction that is usually reported in timed grammaticality judgment studies of normal listeners (Blackwell and Bates, 1995; Blackwell, Bates and Fisher, 1996; Devescovi et al., 1997). The main effect of noun gender was not significant ( $F < 1.0, n.s.$ ), even though (as noted above) the masculine target nouns in Experiment 2 were slightly but significantly longer than feminines or neuters. However, there was a significant concordance by noun gender interaction [ $F(4, 16) = 7.2,$

$p = .008$ ]. Cell means for the full interaction are summarized in Table V.

Examination of cell means indicates that this interaction may be due to yet another manifestation of the Markedness Effect. Specifically, RTs to accept correct masculine-noun pairs (1007 ms, s.d. = 171) were substantially *slower* than RTs to accept correct feminine-noun pairs (967 ms, s.d. = 151) or neuter-noun pairs (965 ms, s.d. = 142). By contrast, RTs to reject incorrect masculine-noun pairs were *faster* (1063 ms, s.d. = 151) than RTs to reject feminines (1099, s.d. = 193) or neuters (1099, s.d. = 215). In other words, correct masculines take longer to accept but incorrect masculines are easier to reject. This complete reversal of the gender effect for “good” vs. “bad” stimuli cannot be due to the fact that masculines are slightly (though significantly) longer than feminine or neuter nouns, because these nouns were counterbalanced across the “good” and “bad” conditions.

To understand this finding, recall the word recognition effects in Experiment 1. In the cued shadowing task (Experiment 1), masculine nouns, our unmarked class, were relatively unaffected by the preceding gender cue. To the extent that there was any effect at all, it involved inhibition by the “wrong” modifier but no facilitation by the “correct” modifier. In the judgment task (Experiment 2), the difference between concordant masculines (1007 ms) and discordant masculines (1063 ms) is also relatively small: a 56-ms difference that may reflect nothing more than the usual disadvantage for a “no” response in grammaticality judgment tasks. For feminine and neuter nouns in the cued shadowing task (Experiment 1), the effects of gender context were much larger: strong facilitation for both feminine and neuter nouns, and significant inhibition for feminines. In the judgment task (Experiment 2), the difference between “accept” and “reject” items for feminine and neuter nouns is also very large: 132 ms for feminines and 133 ms for neuters.

Hence the Markedness Effect on judgment is compatible with the Markedness Effect on word recognition. This is what we would expect, because word recognition (Experiment 1) is a prerequisite to grammaticality judgment (Experiment 2). It is interesting that the markedness effect on judgments (Experiment 2) is actually much larger than the corresponding result in word recognition (Experiment 1). However, the two tasks differ in so many ways that we can only speculate about the reason for this difference in magnitude. For present purposes, we are satisfied that the two experiments (with different tasks and different sets of normal controls) yield parallel results.

We turn now to results for aphasic patients, to determine (1) whether they are able to make above-chance judgments of grammaticality, and (2) whether they will show Markedness Effect in Experiment 2 (a

metalinguistic task) even though they failed to display it in Experiment 1 (an automatic-priming task).

### *Results for Aphasic Patients*

#### *Errors*

All 22 patients were able to complete the experiment. However, some aphasic patients made a large number of errors in the judgment task, and there was substantial variation in error rates. On the assumption that grammatical errors comprise the “signal” to be detected in a grammaticality judgment task, we define “hit” as a correct rejection of a bad sentence, “miss” as an incorrect acceptance of a bad sentence, and “false alarm” as an incorrect rejection of a good sentence. The average rate of incorrect acceptances (misses) was 13% (s.d. = 18), with a range from 0% to 51%. The average rate of incorrect rejections (false alarms) was also 13% (s.d. = 16), with a range from 0% to 64%. Based on a binomial test, above-chance performance for individual patients would correspond to an error rate of no more than 24 errors per cell for concordant and discordant items, respectively (33%). By this criterion, 20 out of 22 patients displayed above-chance sensitivity on concordant items, and 18 out of 22 were above chance on discordant items. When concordant and discordant are considered together, only two patients (# 19 and # 20) were performing below chance.

To determine whether errors are affected differentially by concordance and gender, a 2 (concordant vs. discordant) by 3 (masculine, feminine, neuter) within-subjects analysis of variance was conducted for all 22 patients. This analysis yielded no main effect of congruence [i.e. no difference in misses vs. false alarms,  $F(1, 18) < 1.00$ , n.s.], no main effect of gender [ $F(2, 18) = 1.84$ , n.s.], and no significant interaction [ $F(2, 18) < 1.00$ , n.s.]. In other words, misses and false alarms are equally probable and appear to be distributed evenly across conditions.

Additional analyses were undertaken to determine whether the pattern of false alarms and misses differed for these patients along any of the symptom dimensions and classifications in Table II. In all these analyses, the  $2 \times 3$  within-subjects analysis was repeated with between-subject groupings based on the Luria classification and the Western classification ( $3 \times 2 \times 3$ ), as well as two-way between-subject groupings ( $2 \times 2 \times 2$ ) based on fluency, overall severity, overall production, lexical production, grammatical production, overall comprehension, lexical comprehension and grammatical comprehension. In the interests of brevity, all of these exploratory analyses can be summarized briefly as follows. There were no significant main effects of congruency or gender in any analysis, and no interactions between congruency and gender (all  $F < 2.0$ ,  $p > .05$ ). Nor did we find any significant interactions involving the various between-subjects groupings. The

only effects that reached significance were main effects of group. This includes a main effect for the aphasia classification based on Western criteria [ $F(2, 20) = 4.31$ ,  $p < .03$ ], reflecting higher overall error rates for Broca’s aphasics (18.3%, s.d. = 13) and Wernicke’s aphasics (20.4%, s.d. = 22), compared with anomics (2.2%, s.d. = 1.2). The main effect of fluency failed to reach significance [ $F(1, 20) = 2.15$ ,  $p < .16$ ], but all of the other two-way groupings were highly significant (ranging from  $p < .015$  to  $p < .0001$ ). In other words, accuracy in grammaticality judgment is affected by severity of aphasia, regardless of the specific way in which it is measured.

As we noted in Experiment 1, all of these subgroup analyses involve an arbitrary division of patients along one or more symptom dimensions. Canonical correlation is a technique that permits us to assess the effect of continuous dimensions on performance in a within-subjects experimental design. For this purpose, we used the three orthogonal aphasia factors (production, comprehension, lexical) derived in Experiment 1 (see Table IV) to determine whether any of these factors are related to the main effects or interactions in the  $2 \times 3$  experimental design. The contrasts of interest include (1) the main effect of concordance (concordant vs. discordant), (2) two contrasts that define the main effect of gender (masculine vs. feminine; masculine vs. neuter), and (3) two contrasts that define the gender by concordance interaction (masculine vs. feminine by concordant vs. discordant; masculine vs. neuter by concordant vs. discordant). An omnibus canonical correlation between the three aphasia factors and these experimental contrasts did reach significance [Wilks’ Lambda = 0.19,  $F(15, 39) = 2.15$ ,  $p < .028$ ], which means that aphasic symptoms do predict a differential breakdown of error rates across the main effects and interactions in this design. To explore this omnibus effect further, separate canonical correlations were conducted for each of the principal contrasts. For the main effect of concordance, the canonical correlation was significant [Wilks’ Lambda = 0.41,  $F(3, 18) = 8.67$ ,  $p < .0009$ ]. The corresponding analysis for the two gender contrasts did not reach significance [Wilks’ Lambda = 0.57,  $F(6, 34) = 1.84$ ,  $p < .12$ ], nor did the canonical correlation against the contrasts that define the gender by concordance interaction [Wilks’ Lambda = 0.63,  $F(6, 34) = 1.48$ ,  $p < .22$ ]. This tells us that the significant result in the omnibus analysis was coming primarily from the contribution of aphasic symptoms to the difference between concordant items (where errors are incorrect rejections, or false alarms) and discordant items (where errors are incorrect acceptances, or misses).

Because we know that misses and false alarms may reflect different mental operations, it is interesting that the aphasia symptom factors do make differential contributions to each error type. To investigate this finding

in more detail, the three derived aphasia factors were used in stepwise regression analyses, conducted separately for percent false alarms (incorrect rejections of "good" items) and percent misses (incorrect acceptances of "bad" items). Results of these analyses are summarized in Table VI. In the analysis of false alarms, the three factors together accounted for a very large 72.4% of the variance ( $p < .0001$ ). Furthermore, Table VI shows that two of the aphasia factors made a significant independent contribution to these judgment scores when it was entered on the last step: 31.0% from the production factor ( $p < .0003$ , with a zero-order partial correlation of  $-.73$ ) and 35.2% from the lexical factor ( $p < .0001$ , with a zero-order partial correlation of  $-.75$ ). The contribution from the comprehension factor on the last step just missed significance (6.0%,  $p < .07$ , zero-order partial correlation =  $-.42$ ). In other words, the ability to accept a "good" gender item is disrupted by virtually all aspects of aphasia, not just total severity but also the separate dimensions of production, lexical processing and (to a lesser extent) comprehension. Results for misses (incorrect acceptances of "bad" items) showed a very different pattern. The three variables together accounted for 61.5% percent of the variance ( $p < .0005$ ). However, the only unique contribution to reach significance was a substantial 54% ( $p < .0001$ ) added by the comprehension factor on the last step (zero-order partial correlation =  $-.75$ ). Putting these two distinct patterns together, we may conclude that "false alarms" (incorrect rejections) reflect noisy performance that can have multiple causes, but "misses" (failure to detect a grammatical error) reflect specific deficits in receptive language processing. A visual examination of the data for individual patients indicated that a high proportion of these misses came from three patients who were among the most severe expressive agrammatic patients in the group. Nevertheless, these dimensional analyses indicate that it is their receptive processing (including receptive agrammatism) that is most responsible for this particular kind of error.

#### *Reaction Times*

RT analyses were based only on trials that elicited a correct response, under 6000 ms. For comparison with performance by normal controls, cell means for all conditions are included in Table V, for all 22 aphasic patients. As Table V shows, judgment times for these aphasic patients were very slow (1716 ms,  $s.d. = 535$ ), approximately 700 ms slower than performance by young controls (1033,  $s.d. = 161$ ).

Judgment times for all patients were subjected to a 2 (concordance) by 3 (gender) analysis of variance. This analysis yielded no main effect of congruence [ $F(1, 20) < 1.0$ , n.s.], no main effect of gender [ $F(2, 20) < 1.0$ , n.s.], and no interaction [ $F(2, 20) = 1.14$ , n.s.]. Note in particular that patients do not display the Mark-

edness Effect (i.e. a gender by congruence interaction) that we reported above for normal controls. To determine whether these null effects were due to patients who performed at chance, the analysis was repeated with #19 and #20 removed. Again, none of the effects reached significance (all  $F < 1.00$ , n.s.). Finally, we repeated the  $2 \times 3$  analysis of variance excluding the six patients with error rates at or above 30% (close to chance: #4, #6, #7, #8, #19, #20). Results were the same: no significant main effects and no interaction.

If the normals and aphasic patients are performing differently, then this difference should show up as a 3-way interaction of patient group, concordance and gender (similar to results for Experiment 1). To ask this question, we conducted a 2 (normals vs. aphasics)  $\times$  2 (concordant vs. discordant)  $\times$  3 (masculine, feminine, neuter) analysis of variance with patient group as a between-subjects variable, and concordance and gender as within-subjects variables. Not surprisingly, there was a large main effect of patient group [ $F(1, 40) = 30.04$ ,  $p < .0001$ ], reflecting the fact that judgment times are almost 700 ms slower for aphasics. The congruence by patient group interaction just missed significance [ $F(1, 40) = 3.97$ ,  $p < .053$ ], reflecting a larger difference between concordant and discordant items for normals. However, there was no interaction between patient group and gender [ $F(1, 40) < 1.0$ , n.s.], nor did the predicted 3-way interaction of patient group, gender and concordance reach significance ( $F < 1.00$ , n.s.). Although the main point of this analysis was to look for interactions with patient group, it is worth noting that the interaction between gender and congruence also failed to reach significance in this analysis. In other words, the Markedness Effect reported above for normals is washed out when results are combined for aphasic patients and controls [ $F(1, 40) < 1.81$ , n.s.].

To determine whether these results are due to patients who are responding close to chance, the same  $2 \times 2 \times 3$  analysis was conducted excluding patients #19 and #20. Results did not change. Finally, we repeated the analysis excluding all patients with error rates of 30% or above (near chance). This analysis yielded the expected main effect of patient group [ $F(1, 34) = 40.69$ ,  $p < .0001$ ]. The interaction between patient group and concordance just missed significance, similar to results with the full sample [ $F(1, 32) = 4.05$ ,  $p < .052$ ], reflecting a larger concordance effect in normals. The group by noun gender interaction once again failed to reach significance ( $F < 1.0$ , n.s.), nor did we find a significant 3-way interaction of patient group by gender by concordance [ $F(2, 32) = 1.34$ , n.s.]. However, the congruence by gender interaction did emerge in this analysis [ $F(2, 32) = 3.41$ ,  $p < .04$ ]. In other words, we do see a Markedness Effect when data for normals and aphasics are combined, but only if we exclude patients with near-chance performance on the judgment task.

To determine whether there are any patterns in the judgment time data due to aphasia subtype or specific symptom dimensions, we again conducted a series of exploratory analyses for patients only, partitioning the groups according to the symptom classifications and dimensions summarized in Table II. Data for all 22 patients were included in these analyses. We will restrict our attention only to main effects or interactions involving the between-subject groupings, with particular emphasis on a 3-way interaction involving group, congruency and gender (i.e. a test for the missing Markedness Effect). These can be summarized easily: the three-way interaction did not even approach significance in any of the analyses (all  $F < 1.0$ , n.s.). The two-way interactions of group by gender and group by congruency also failed to reach significance in any analysis. Nor were there any main effects of group, although the fluency analysis approached significance ( $F = 3.36$ ,  $p < .09$ ), reflecting slightly slower RTs in this button-pressing task for nonfluent patients.

Finally, we used the three symptom factors derived in Experiment 1 (comprehension, production, lexical) as predictors in a canonical correlation analysis. The relevant contrasts within the  $2 \times 3$  design are the same ones described above for the error analysis. Results of the omnibus analysis did not even approach significance [Wilks' Lambda = 0.32,  $F(15, 39) = 1.32$ ,  $p > .23$ ]. No further explorations were attempted.

We conclude that reaction times to make a correct grammaticality judgment are not differentially affected by these specific symptoms of aphasia, a result similar to findings for reaction times in the priming task (Experiment 1). This contrasts with the robust and differential effects of aphasic symptom dimensions that we observed for judgment accuracy.

#### *Interim Conclusions*

In Experiment 2, normals showed a significant interaction between grammatical context and gender, which we interpret as a manifestation of the Markedness Effect (based on class size and zero/overt endings). Specifically, we found smaller effects of context on the unmarked masculine nouns, much larger effects of context on the marked feminine and neuter nouns. This result parallels our findings for Experiment 1 with a different group of young adults. Aphasic patients failed to show the Markedness Effect in Experiment 2 (i.e. no gender by context interaction), even though their error scores indicate performance well above chance on the judgment task. This result parallels findings for the same patients in Experiment 1, in which aphasics also failed to show the Markedness Effect despite robust evidence of gender priming. The difference between normals and aphasics showed up as a significant three-way interaction in Experiment 1. The corresponding three-way interaction failed to reach significance in

Experiment 2, although we must underscore that three-way interactions are not easy to detect.

In Experiment 2 (as in Experiment 1), we conducted exploratory analyses to determine whether accuracy, reaction times or the missing Markedness Effect are associated with aphasia classifications or with any of the specific symptoms recorded in Table II. We did find a general effect of severity on error scores, but these groupings did not account for the presence or absence of the Markedness Effect. Our applications of the canonical correlation technique did not reveal interesting relationships between aphasia factors and reaction times in the grammaticality judgment task. However, this approach did reveal robust and selective effects of aphasia factors on accuracy scores. Follow-up analyses showed that false alarms (incorrect rejection of "good" items) were significantly and uniquely affected by both the lexical factor and the production factor, while the contribution from the comprehension factor just missed significance. Hence this kind of error can be induced in various ways, by deficits that cut across comprehension and production. By contrast, misses (incorrect acceptance of "bad" items) were affected significantly and uniquely only by the comprehension factor. The differential vulnerability of "saying yes" and "saying no" underscores that these decisions involve different mental operations, which can be differentially disrupted in aphasia. In other words, specific aphasic symptoms can have specific effects on judgment accuracy (especially the ability to detect and reject violations), but they do not differentially affect reaction times, nor do they account for the absence of a Markedness Effect in either of these experiments.

#### SUMMARY AND CONCLUSIONS

In Experiment 1, we found robust gender priming in both normals and aphasic patients. However, normals also displayed a gender Markedness Effect that did not appear in aphasics: inhibition for unmarked masculine nouns, robust facilitation and inhibition for feminine nouns, and significant facilitation only for neuter nouns (the smallest gender class in Russian). Although Experiment 2 was not a priming study, normals displayed another Markedness Effect on grammaticality judgment that did not appear in aphasics. This involved large context effects on feminine and neuter nouns (i.e. large difference in RTs between correct acceptance and correct rejection), and much smaller context effects on masculine nouns, so that masculines completely reversed positions between the two judgment types.

To explain the Markedness Effect in normals and its absence in aphasics, we proposed an account based on Elman's (1990) recurrent network model of grammatical processing. Specifically, lexical items are defined by positions within a high-dimensional lexical space, whose dimensions reflect both sound and meaning.

Lexical access involves movement into the correct position within this hyperspace. Similarity (phonological, semantic and grammatical class) is defined by relative distance between items within various planes or regions. An unmarked region (including unmarked gender) would be a region that is relatively central along one or more dimensions, from which the largest number of moves can be made in the shortest possible time. Within this framework, context can prime lexical access by moving the listener closer to or farther away from the region where the lexical target will occur. In the absence of such information (e.g. at the beginning of a new trial), competent Russian listeners move back into a more central position (like a competent tennis player moving back to center court). Incompetent listeners (including these Russian patients with aphasia) can use context to assist them in lexical identification (i.e. they can show priming). However, like incompetent tennis players, they fail to move back toward a maximally efficient central position, lingering too long in the region where they had moved on the last trial.

As a group, the aphasic patients displayed robust gender priming in Experiment 1 and above-chance judgment accuracy in Experiment 2, which suggests to us that knowledge of gender agreement is preserved despite their many and varied linguistic problems. The grammaticality judgment results are compatible with reports by other investigators for Italian (Badecker et al., 1995; Bates et al., 2000; Devescovi et al., 1997). The performance deficits observed in all these studies are probabilistic in nature, and seem to reflect disruptions in the process by which this gender knowledge is accessed and deployed. The Elman model may be useful in explaining the global difference in performance between normal and aphasic Russian listeners, including the presence of a Markedness Effect in normals and the absence of a Markedness Effect in patients. However, it would be useful if we could identify more specific correlates (and potential causes) of deviant gender processing. Although we looked at the evidence from many points of view, we were unable to find systematic effects of aphasia type (by the Luria system or the Western classification), nor did we find systematic effects of specific aphasic symptoms (in analyses of variance with median splits on multiple dimensions, or in canonical correlations that preserve continuous symptom variation). In fact, the only specific effect on reaction time in either experiment involved global slowing. In Experiment 1, nonfluent patients were slower in repeating target words (even though they still show significant gender priming). In Experiment 2, nonfluent patients were marginally slower in button pressing ( $p < .09$ ). Neither result is surprising.

In contrast with the null results for reaction time, we did find specific effects of symptom dimensions on judgment accuracy in Experiment 2. Predictably, most indices of severity (in comprehension or production)

were associated with higher error rates overall. In addition, we found interesting effects of symptom dimensions on judgment accuracy, illustrating the value of canonical correlation as a method for exploring the mapping between aphasic symptoms and performance in a within-subjects design. However, these results for accuracy do not help us to explain deviant reaction time profiles. It is possible that the missing Markedness Effect will prove to be a very general symptom in Russian listeners with suboptimal language abilities. It may prove to be a general effect of cognitive aging (though we stress that many of the patients in our sample here were relatively young), in dementia, and perhaps in developmental syndromes like dyslexia and specific language impairment. This is a question for future research.

Why did we find strong evidence for gender priming in these Russian patients (including Broca's aphasics), when Bates et al. failed to find significant gender priming in Italian patients? There are a number of differences between these studies that could explain this discrepancy.

First, the difference between the findings for Russian and the findings for Italian may be due to more general differences between a three- and a two-gender system. In Russian, masculine is the largest class, and it also includes most nouns with zero-marked endings in the nominative singular. These two features together justify the proposal that masculines are the unmarked form. To date, there is no evidence to suggest that markedness plays a role in the two-gender system like that of Italian. Note that the two gender classes are roughly equal in size in Italian, and neither class is zero-marked (i.e. masculines and feminines both end in a vowel for all but a handful of foreign loan words). We also note in this regard that the Russian markedness effect is not a simple binary advantage for masculine nouns (although there was a trend in that direction for aphasic patients). To some extent, markedness results in Experiment 1 actually reflect a continuum: inhibition only for masculines (as we would expect if listeners start with the expectation that they will hear a masculine noun), facilitation only for neuters (as we would expect if listeners do not expect to hear these low-probability nouns), with feminine nouns eliciting both facilitation (reflecting their non-default status) and inhibition (reflecting the fact that feminines are a very large class even though they are less likely than the masculine nouns). It may be that markedness effects are more likely to emerge in a three-gender language than in two-gender languages like Italian, where masculines and feminines are both very common (and neither are zero-marked). Hence gender-priming effects may also differ, in nature and magnitude, for three- vs. two-gender languages.

Second, we used the cued shadowing task in this experiment, in contrast with the gender-monitoring task

employed by Bates et al. (2000). The gender-monitoring task yields much slower reaction times, requires a metalinguistic judgment, and (in research on normals) it tends to yield significant inhibition but no facilitation relative to a neutral baseline. By contrast, cued shadowing yields faster reaction times, requires no metalinguistic judgment, and (in the Bates et al., 2000, study of Italian normals, and in the Akhutina et al., 1999, studies of Russian normals) it results in significant facilitation relative to neutral baseline. Hence the kind of gender priming that we have observed in Russian may be due to automatic processes, while the gender priming that Bates et al. failed to find in Italian may be due to a more strategic, postlexical effect.

There is a third possibility that merits especially serious consideration. The stimuli that we employed in the present study were restricted to phonologically transparent nouns and their modifiers, i.e. items that “wear their gender on their sleeve”. Bates et al. used both transparent and opaque items in their study, and failed to find priming in aphasic patients for either one. However, they did find significantly higher accuracy levels for aphasic patients when gender was transparently marked. In their studies of semantic priming, Milberg, Blumstein, Utman and colleagues have shown that phonological clarity can affect priming in aphasics. In particular, even modest amounts of phonological degradation can reduce or eliminate priming effects, especially for nonfluent patients (Milberg et al., 1987, 1988a, 1988b; Milberg et al., 1995; Utman and Bates, 1998; Utman et al., 1998). By using phonologically transparent gender stimuli, we have tested for gender priming under ideal circumstances, analogous to the conditions that optimize semantic priming in English-speaking aphasics.

This last point can be interpreted in two ways: (1) transparency of marking makes gender priming easier for aphasic patients; (2) transparency of marking turns a grammatical priming experiment (based on gender congruency) into a phonological priming experiment (based on phonological similarity). If the second possibility is correct, it would greatly reduce the importance of our findings for the argument that grammatical knowledge is preserved in aphasia. However, we think it is quite unlikely that these results are due to phonological similarity alone, independent of grammatical knowledge, because the phonological similarity between primes and targets in this experiment is much smaller and less consistent than one might think at first glance. First, transparent masculine nouns and transparent masculine modifiers are only similar at a very abstract level: they are both zero-marked, ending in final consonants, but these consonants can be very different (e.g. *moj* (masc.) *bilet* (masc.) = ‘my ticket’), and while the possessive pronoun always ends in *-j*, the noun following it can end in a variety of consonants. Hence any priming for masculines that relies on a phonological match will

have to occur at an abstract level. Second, transparent feminine and neuter nouns and their congruent modifiers are similar only in the final vowel (*-o* or *-a*). There is very little evidence in the literature on phonological priming to suggest that overlap on a single phoneme (other than the initial phoneme) is sufficient to bring about significant facilitation. Third, the possessive pronouns used as primes in this experiment all carry final stress, which makes the vowel in the feminine and neuter forms clear and unreduced. However, only half of the noun targets have the final stress; in the absence of a final stress, the final vowel of the stimuli gets reduced (see Appendix). Hence any tendency toward rhyme priming based on the final vowel would be disrupted or degraded by a phonetic mismatch between the stressed and reduced unstressed vowel. Finally, an explanation based on phonological priming would predict very different results for neuter nouns from those that we have observed here. Phonologically transparent neuter nouns and their modifiers end in a final vowel *-o*. Hence a listener could try to match neuter nouns and neuter adjectives based on their final vowel. However, two of the three gender-neutral possessive pronouns that served as the neutral baseline for this experiment also end in a final *-o* (*jejó* = ‘her’; *jego* = ‘his’; vs. *ix* = ‘their’). If reaction times were based entirely on phonological similarity, then we would expect Russian listeners to display similar behavior for neutral and congruent neuter-noun items, while displaying larger differences between neutral and incongruent neuter-noun items. This ought to show up as a large inhibitory effect but very little facilitation—exactly the opposite of what we observed for neuter nouns in this experiment. Putting these arguments together, it seems quite unlikely that the gender-priming results observed here for Russian-speaking normals and aphasics are based only on phonological similarity. To the extent that phonological similarity does have an effect (increasing the magnitude of priming), it seems to work as a mnemonic within a system that is defined in lexical and grammatical terms. From this point of view, our results are compatible with a growing literature indicating that grammatical knowledge is preserved in aphasia, which means that an explanation of aphasic symptoms must rest in the processes by which this knowledge is accessed and deployed. Precisely because gender lies at the intersect of phonology, grammar and the lexicon, it is fertile ground for future investigations of normal and impaired language processing.

*Acknowledgements.* Support for this research was provided by a grant to Elizabeth Bates, "Cross-linguistic studies of aphasia" (NIDCD R01-DC00216). Our thanks to Robert Buffington and Meiti Opie for technical support, and to Fred Dick, Laura Kemmer and Nicole Wicha for assistance during manuscript preparation. Please address correspondence to Elizabeth Bates, Center for Research in Language 0526, University of California, San Diego, La Jolla, CA 92093-0526 (email: bates@crl.ucsd.edu).

#### REFERENCES

- Akhutina, T. *Poroždenie Reči. Nejrolingvističeskij Analiz Sintaksisa* [Language Production: The Neurolinguistic Analysis of Syntax]. Moscow: Moscow State University, 1989.
- Akhutina, T. *Roman Jakobson i Razvitie Russkoi Nejrolingvistiki* [Roman Jakobson and the Development of Russian Neurolinguistics]. In H. Baran, S. I. Gindin et al., (Eds.), *Roman Jakobson. Texts, Documents, Studies*, pp. 382-401. English language translation, *Journal of Russian and East European Psychology*, in press.
- Akhutina, T., and Gluzman, J. The neurolinguistic study of semantics. *Aphasiology*, 9: 143-152, 1995.
- Akhutina, T., Kurgansky, A., Polinsky, M., and Bates, E. Processing of grammatical gender in a three-gender system: Experimental evidence from Russian. *Journal of Psycholinguistic Research*, 28: 695-713, 1999.
- Badecker, W., Miozzo, M., and Zanuttini, R. The two-stage model of lexical retrieval: Evidence from a case of anomia with selective preservation of grammatical gender. *Cognition*, 57: 193-216, 1995.
- Bates, E., Devescovi, A., Hernandez, A., and Pizzamiglio, L. Gender priming in Italian. *Perception & Psychophysics*, 58: 992-1004, 1996.
- Bates, E., and Liu, H. Cued shadowing. In F. Grosjean and H. Frauenfelder (Eds.), *A Guide to Spoken Word Recognition Paradigms*. Hove, England: Psychology Press, 1996, pp. 577-581.
- Bates, E., Marangolo, P., Pizzamiglio, L., and Dick, F. *Linguistic and nonlinguistic priming in aphasia* (Tech. Rep. CRL-0005). La Jolla: University of California, San Diego, Center for Research in Language, 2000.
- Bentrovato, S., Devescovi, A, D'Amico, S., and Bates, E. The effect of grammatical gender and semantic context on lexical access in Italian. *Journal of Psycholinguistic Research*, 28: 677-693, 1999.
- Blackwell, A., and Bates, E. Inducing agrammatic profiles in normals: Evidence for the selective vulnerability of morphology under cognitive resource limitation. *Journal of Cognitive Neuroscience*, 7: 228-257, 1995.
- Blackwell, A., Bates, E., and Fisher, D. The time course of grammaticality judgment. *Language and Cognitive Processes*, 11: 337-406, 1996.
- Cohen, J., MacWhinney, B., Flatt, M., and Provost, J. PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, & Computers*, 25: 257-271, 1993.
- Comrie, B., Stone, G., and Polinsky, M. *The Russian Language in the Twentieth Century*. Oxford: Oxford University Press, 1996.
- Corbett, G.G. Gender in Russian: An account of gender specification and its relationship to declension. *Russian Linguistics* 6: 197-232, 1982.
- Corbett, G.G. *Gender*. Cambridge: Cambridge University Press, 1991.
- Dell, G.S., Schwartz, M.F., Martin, N., Saffran, E.M., and Gagnon, D.A. Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104: 801-838, 1997.
- Devescovi, A., Bates, E., D'Amico, S., Hernandez, A., Marangolo, P., Pizzamiglio, L., and Razzano, C. Grammaticality judgments in patients with and without expressive agrammatism. *Aphasiology*, 11: 543-579, 1997.
- Elman, J.L. Finding structure in time. *Cognitive Science*, 14: 179-211, 1990.
- Goodglass, H., and Kaplan, H. *Boston Diagnostic Aphasia Examination* (2nd ed.). Philadelphia: Lea and Febiger, 1983.
- Grosjean, F., Dommergues, J-Y, Cornu, E., Guillemon, D., and Besson, C. The gender-marking effect in spoken word recognition. *Perception & Psychophysics*, 56: 590-598, 1994.
- Gurjanov, Lukatela, G., Lukatela, K., Savic, M., and Turvey, M. Grammatical priming of inflected nouns by the gender of possessive adjectives. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 11: 692-701, 1985.
- Gurjanov, M., Lukatela, G., Moskovljevic, J., and Turvey, M.T. Grammatical priming of inflected nouns by inflected adjectives. *Cognition*, 19, 55-71, 1985.
- Hillert, D., and Bates, E. *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, 1996.
- Jacobsen, T. Effects of grammatical gender on picture and word naming: Evidence from German. *Journal of Psycholinguistic Research*, 28: 499-515, 1999.
- Jakobson, R. Toward a linguistic typology of aphasic impairments. In A.V.S. de Reuck and M. O'Connor (Eds.), *Ciga Foundation Symposium: Disorders of Language*. London: J. & A. Churchill, Ltd., 1964, pp. 21-42.
- Kempe, V., and MacWhinney, B. Processing of morphological and semantic cues in Russian and



- German. *Language and Cognitive Processes*, 14: 129-171, 1999.
- Lu, C-C., Bates, E., Li, P., Tzeng, O., Hung, D., Tsai, C.-H., Lee, S-E., and Chung, Y-M. Judgments of grammaticality in aphasia: The special case of Chinese. *Aphasiology*, 14: 1021-1954, 2000.
- Lukatela, G., Kostic, A., Feldman, L., and Turvey, M. Grammatical priming of inflected nouns. *Memory & Cognition*, 11: 59-63, 1983.
- Luria, A.R. *Traumatic Aphasia: Its Syndromes, Psychology and Treatment*. The Hague and Paris: Mouton, 1970.
- Milberg, W., Blumstein, S.E., and Dworetzky, B. Processing of lexical ambiguities in aphasia. *Brain and Language*, 31: 138-150, 1987.
- Milberg, W., Blumstein, S., and Dworetzky, B. Phonological factors in lexical access: Evidence from an auditory lexical decision task. *Bulletin of the Psychonomic Society*, 26: 305-308, 1988a.
- Milberg, W., Blumstein, S., and Dworetzky, B. Phonological processing and lexical access in aphasia. *Brain and Language*, 34: 279-293, 1988b.
- Milberg, W., Blumstein, S., Katz, D., and Gershberg, F. Semantic facilitation in aphasia: Effects of time and expectancy. *Journal of Cognitive Neuroscience*, 7: 33-50, 1995.
- Neely, J.H. Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner and G.W. Humphreys (Eds.), *Basic Processes in Reading: Visual Word Recognition*. Hillsdale, NJ: Erlbaum, 1991, pp. 264-336.
- Prather, P., Shapiro, L., Zurif, E., and Swinney, D. Real-time examinations of lexical processing in aphasics. *Journal of Psycholinguistic Research*, 20: 271-281, 1991.
- Prather, P., Zurif, E., Stern, C., and Rosen, T.J. Slowed lexical access in non-fluent aphasia: A case study. *Brain and Language*, 43: 336-348, 1992.
- Rjabova (Akhutina) T. Mexanizm poroždenija reči po dannym afaziologii [The mechanism of speech production according to aphasiological data]. In *Questions of Language Production and Language Teaching*. Moscow, Moscow University Press, 1967, pp.76-94.
- Štejnfeldt, E. *Častotnyj Slovar' Sovremennogo Russkogo Literaturnogo Jazyka. 2500 Naibolee Upotrebiteľ'nyx Slov* [Frequency Dictionary of Contemporary Standard Russian. 2500 Most Commonly Used Words]. Tallin: n.p., 1963.
- Swaab, T., Brown, C., and Hagoort, P. Spoken sentence comprehension in aphasia: Event-related potential evidence for a lexical integration deficit. *Journal of Cognitive Neuroscience*, 9: 39-66, 1997.
- Swaab, T., Brown, C., and Hagoort, P. Understanding ambiguous words in sentence contexts: Electrophysiological evidence for delayed contextual selection in Broca's aphasia. *Neuropsychologia*, 36: 737-761, 1998.
- Taraban, R., and Kempe, V. Gender processing in native and non-native Russian speakers. *Applied Psycholinguistics*, 20: 119-148, 1999.
- Tsvetkova, L.S., Akhutina, T., Polonskaya, N.N., and Pylaeva, N.M. A method for evaluation of speech dynamics in aphasia. *Soviet Psychology*, 17: 39-45, 1979.
- Tsvetkova, L.S., Akhutina, T., and Pylaeva, N.M. *Metod Ocenki Reči pri Afazii* [The Method of Evaluating Speech under Aphasia]. Moscow: Moscow State University, 1982.
- Tyler, L., Ostrin, R., Cooke, M., and Moss, H. Automatic access of lexical information in Broca's aphasics: Against the automaticity hypothesis. *Brain and Language*, 48: 131-162, 1995.
- Utman, J.A., and Bates, E. Effects of acoustic degradation and semantic context on lexical access: Implications for aphasic deficits (Abstract). *Brain and Language*, 65: 516-518, 1998.
- Utman, J.A., Blumstein, S.E., and Sullivan, K. From sound to meaning: Reduced lexical activation in Broca's aphasics. Manuscript submitted for publication.
- Wicha, N.Y., Bates, A., Hernandez, A., Reyes, I., and Gavaldón de Barreto, L. G. *When zebras become painted donkeys: The interplay between gender and semantic priming in a Spanish sentence context* (Technical Report 9701). La Jolla: University of California, San Diego, Center for Research in Language, 1997.
- Zasorina, L.N. (Ed.). *Častotnyj Slovar' Russkogo Jazyka: Okolo 40,000 slov* [Frequency Dictionary of Russian: About 40,000 Words]. Moscow: Russkij jazyk, 1977.

APPENDIX:  
STIMULUS MATERIALS FOR EXPERIMENTS 1 and 2

Nouns (Experiment 1)

<b>Feminine</b>		<b>Masculine</b>		<b>Neuter</b>	
muká	‘flour’	úžin	‘supper’	jajcó	‘egg’
izbá	‘peasant cottage’	ètáž	‘floor’	mjásó	‘meat’
váza	‘vase’	urók	‘lesson’	vinó	‘wine’
zimá	‘winter’	úžas	‘horror’	góre	‘grief’
bedá	‘trouble’	obéd	‘dinner’	bjuró	‘bureau’
gubá	‘lip’	udár	‘blow’	úxo	‘ear’
gorá	‘mountain’	átom	‘atom’	léto	‘summer’
reká	‘river’	úgol	‘corner’	nébo	‘sky’
nogá	‘leg’	jazyk	‘tongue’	útro	‘morning’
vodá	‘water’	ópyt	‘experience’	móre	‘sea’
igrá	‘play’	ogón’	‘fire’	oknó	‘window’
zemljá	‘earth’	górod	‘city’	licó	‘face’
vojná	‘war’	zavód	‘plant, factory’	méstó	‘place’
vesná	‘spring’	mésjac	‘month’	zernó	‘grain’
lámpa	‘lamp’	muzéj	‘museum’	dobró	‘grace, goodness’
vódka	‘vodka’	bilét	‘ticket’	zvenó	‘link’
búkva	‘letter’	rýnok	‘market’	máslo	‘butter’
lénta	‘ribbon’	zabór	‘fence’	górlo	‘throat’
nítka	‘thread’	bokál	‘goblet’	vedró	‘bucket’
lózka	‘spoon’	mešók	‘bag’	metró	‘subway’
minúta	‘minute’	millión	‘million’	dérevo	‘tree’
rakéta	‘rocket’	análiz	‘analysis’	zóloto	‘gold’
monéta	‘coin’	urožáj	‘harvest’	molokó	‘milk’
lopáta	‘spade’	ogoród	‘orchard’	bolóto	‘marsh’

Nouns (Experiment 2)

**Feminine**

síla	‘strength’
dušá	‘soul’
báza	‘base’
rýba	‘fish’
šéja	‘neck’
kóža	‘skin’
róza	‘rose’
túča	‘storm cloud’
kníga	‘book’
vojná	‘war’
škóla	‘school’
fórna	‘form’
stená	‘wall’
tájna	‘secret’
sména	‘change’
šápka	‘cap’
nedélja	‘week’
idéja	‘idea’
múzyka	‘music’
zapíska	‘note’
teléga	‘cart’
posúda	‘crookery’
berjóza	‘birch tree’
océnka	‘estimate’

**Masculine**

gólos	‘voice’
uspéx	‘success’
úgol’	‘coal’
itóg	‘sum total’
ótdyx	‘rest’
ijúl’	‘July’
pójas	‘belt’
ovjós	‘oats’
sojúz	‘union’
sovét	‘advice’
véčer	‘evening’
béreg	‘bank’
pálec	‘finger’
román	‘novel’
tánc	‘dance’
piróg	‘pie’
interés	‘interest’
okeán	‘ocean’
úroven’	‘level’
apparát	‘apparatus’
ókorok	‘gammon’
ogonjók	‘light’
osádok	‘sediment’
ideál	‘ideal’

**Neuter**

ímja	‘name’
póle	‘field’
télo	‘body’
čúdo	‘miracle’
seló	‘village’
kinó	‘cinema’
pívo	‘beer’
fóto	‘photo’
slóvo	‘word’
sólnc	‘sun’
sérdce	‘heart’
pis’ mó	‘letter’
čisló	‘number’
tepló	‘warmth’
pal’tó	‘coat’
rebró	‘rib’
načálo	‘beginning’
rádio	‘radio’
koléno	‘knee’
ózero	‘lake’
óblako	‘cloud’
želézo	‘iron’
kolesó	‘wheel’
jábloko	‘apple’

**Table I: Neurological and Demographic Characteristics of Aphasic Patients**

<b>ID</b>	<b>Gender</b>	<b>Age at Test</b>	<b>Years of School</b>	<b>Years Post Lesion</b>	<b>Etiology</b>	<b>Fluency Level</b>	<b>Luria Classification</b>	<b>Western Classification</b>
8	M	60	15	3	CVA	2	Motor	Broca
7	M	22	8	3	Trauma	2	Motor	Broca
6	F	21	9	7	CVA	2	Motor	Broca
21	M	51	15	0.5	CVA	3	Motor	Broca
4	M	22	14	3	Trauma	3	Motor	Broca
23	M	33	11	2	Trauma	3	Motor	Broca
9	M	54	15	5	CVA	3	Motor	Broca
20	M	37	10	4	Trauma	4	Sensory	Wernicke
13	M	47	10	5	CVA	4	Sensory	Wernicke
19	F	45	8	2	CVA	4	Sensory	Wernicke
12	F	44	15	3	Tumor	4	Sensory	Wernicke
5	F	48	11	2	CVA/right hemisphere	4	Sensory	Wernicke
24	F	42	20	0.5	Hemorrhage	4	Sensory	Anomic
18	F	50	15	0.5	CVA	4	Sensory	Anomic
14	F	63	15	0.5	CVA	4	Sensory	Anomic
15	F	33	15	0.5	CVA	4	Sensory	Anomic
3	F	48	12	1	Hemorrhage	4	Sensory	Anomic
17	M	39	15	1	CVA	4	Semantic	Wernicke
1	M	60	10	0.5	CVA	4	Semantic	Anomic
10	F	45	15	1	Hemorrhage	4	Semantic	Wernicke
22	M	67	15	0.5	CVA	4	Semantic	Anomic
16	F	63	15	2	CVA	4	Semantic	Anomic

**Table II: Results of Language Battery for Aphasic Patients**  
(maximum possible = 30 for all scales except story production, which can go higher)

ID	EXPRESSIVE LANGUAGE SCALES						RECEPTIVE LANGUAGE SCALES					
	Answers in Dialogue	Nouns: Picture Naming	Verbs: Picture Naming	Sent. Prod.	Story Prod.	Total	Comp. in Dialogue	Nouns: picture choice	Verbs: picture choice	Sent. Comp: picture choice	Comp. of commands	Total
8	22.5	26	15	15	2.5	<b>81</b>	30	27	25.5	22	27	<b>131.5</b>
7	22.5	29	24.5	20	5	<b>101</b>	30	25	21	14	22.5	<b>112.5</b>
6	27	29	27.5	18	9.5	<b>111</b>	30	27.5	27	8	22.5	<b>115</b>
21	22.5	25	22	16.5	9	<b>95</b>	30	30	27.5	20	30	<b>137.5</b>
4	30	28	25	21	20.5	<b>124.5</b>	30	26	19.5	15	27	<b>117.5</b>
23	27	28.5	28	24	13	<b>120.5</b>	30	30	29	20	30	<b>139</b>
9	28.5	29.5	28.5	27	24	<b>137.5</b>	30	27	27	30	30	<b>144</b>
20	27	23.5	15	19	12	<b>96.5</b>	27	16.5	16.5	9	10.5	<b>79.5</b>
13	22.5	25	19.5	13	6	<b>86</b>	25.5	24.5	22	12	18	<b>102</b>
19	22.5	25	16	19	15.5	<b>98</b>	25.5	19	17.5	20	21	<b>103</b>
12	25.5	29	21.5	20	15.5	<b>111.5</b>	27	24.5	24	26	27	<b>128.5</b>
5	27	30	29	22	18.5	<b>126.5</b>	30	27	21.5	23	22.5	<b>124</b>
24	27.5	28.5	28	24	27	<b>135</b>	30	26.5	27.5	22	28.5	<b>134.5</b>
18	30	28	28	21	18	<b>125</b>	30	30	28	28	30	<b>146</b>
14	25.5	30	27.5	27	23.5	<b>133.5</b>	27	30	27.5	26	30	<b>140.5</b>
15	30	30	30	23	33	<b>146</b>	30	27	25.5	30	24	<b>136.5</b>
3	30	28.5	29	25	31	<b>143.5</b>	30	30	30	30	30	<b>150</b>
17	26	29	26	19	19	<b>119</b>	30	29.5	26	18	19.5	<b>123</b>
1	30	27	25	22	16	<b>120</b>	30	29.5	28	18	27	<b>132.5</b>
10	28.5	30	28.5	22	24	<b>133</b>	30	27	24	16	24	<b>121</b>
22	30	29	29.5	26	24	<b>138.5</b>	30	29	30	24	27	<b>140</b>
16	30	26	26.5	22	41	<b>145.5</b>	30	29	29	24	27	<b>141</b>

**Table III: Cell Means for Experiment 1 for Normal Controls and Aphasic Patients**

	Normal Controls (N = 20)			Aphasic Patients (N = 20)		
	Mean	Standard Deviation	Standard Error	Mean	Standard Deviation	Standard Error
Concordant:	677.81	137.41	30.73	956.20	188.55	42.16
--masculine	687.47	136.56	30.54	936.55	182.79	30.53
--feminine	676.34	136.41	30.50	975.51	186.14	41.62
--neuter	669.63	140.44	31.40	956.56	204.20	45.66
Neutral:	694.27	137.69	30.79	982.84	197.05	44.06
--masculine	685.22	138.48	30.97	967.09	201.41	45.04
--feminine	699.36	138.78	31.03	1010.26	202.25	45.22
--neuter	698.22	137.35	30.71	971.16	200.40	44.81
Discordant:	703.30	141.64	31.67	997.55	195.59	43.73
--masculine	697.45	141.58	31.66	988.09	204.97	45.83
--feminine	710.58	142.03	31.76	1009.61	185.33	41.44
--neuter	701.87	143.11	32.00	994.95	211.07	47.20
Overall RT	691.79	138.66	31.00	978.86	193.20	43.20
--masculine	690.05	138.22	30.91	963.91	194.38	43.46
--feminine	695.43	138.72	30.02	998.46	189.81	42.44
--neuter	689.91	139.87	31.28	974.22	203.82	45.58
Overall Priming:	25.49	18.30	4.09	41.35	26.38	5.90
--facilitation	16.45	9.28	2.08	26.63	27.04	6.05
--inhibition	9.04	15.85	3.54	14.71	23.19	5.19
--masculine	9.97	24.55	6.16	51.54	60.14	13.45
--feminine	34.25	20.71	4.63	34.10	35.90	8.03
--neuter	32.23	19.92	4.46	38.39	50.22	11.23

**Table IV: Factor Analysis of 10 Scales from the Language Battery for 22 Aphasic Patients**

Language Subscales	Factor 1 (Variance explained = 56.4%) “Production factor”	Factor 2 (Variance explained = 14.4%) “Comprehension factor”	Factor 3 (Variance explained = 10.4%) “Lexical factor”
Answering in dialogue	+.77	+.01	+.40
Noun naming (pictures)	+.49	+.17	+.58
Verb naming (pictures)	+.61	+.25	+.69
Sentence production (picture description)	+.82	+.30	+.20
Story telling (complex picture description)	+.88	+.22	+.06
Dialogue comprehension	+.14	+.19	+.84
Noun comprehension (picture choice)	+.06	+.67	+.68
Verb comprehension (picture choice)	+.15	+.73	+.51
Sentence comprehension (picture choice)	+.52	+.76	-.10
Comprehension of commands with objects	+.17	+.86	+.30

**Table V: Cell Means for Experiment 2 for Normal Controls and Aphasic Patients**

	Normal Controls (N = 20)			Aphasic Patients (N = 22)		
	Mean	Standard Deviation	Standard Error	Mean	Standard Deviation	Standard Error
Concordant:	979.72	151.98	33.98	1735.97	555.06	118.34
--masculine	1006.93	170.86	38.21	1796.95	586.54	125.05
--feminine	966.75	150.55	33.66	1722.76	621.65	132.54
--neuter	965.48	141.99	31.75	1688.20	555.06	118.34
Discordant:	1087.04	183.48	41.03	1696.61	560.59	119.52
--masculine	1062.68	151.09	33.78	1681.70	612.32	130.55
--feminine	1099.39	193.64	43.30	1646.58	497.76	106.12
--neuter	1099.07	215.47	48.18	1761.55	713.38	152.09
Overall RT	1033.38	161.49	36.11	1716.29	534.99	114.06



**Table VI: Results of regression analyses using aphasia factor scores in Experiment 2**

	Joint Prediction (all three factors together, last step)			Production Factor (entered on the last step)			Comprehension Factor (entered on the last step)			Lexical Factor (entered on the last step)		
	% Var	p <	r =	% Var	p <	r =	% Var	p <	r =	% Var	p <	r =
Error rates:												
% Error: congruent	.724	.0001	.85	.310	.0003	-.73	.060	.07	-.42	.352	.0001	-.75
% Error: incongruent	.615	.0005	.78	.071	.09	-.40	.540	.0001	-.77	.000	n.s.	-.03