

# Grammatical gender in spoken word recognition in German

JENS BÖLTE

*Westfälische Wilhelms-Universität, Münster, Germany*

and

CYNTHIA M. CONNINE

*State University of New York, Binghamton, New York*

The role of grammatical gender for auditory word recognition in German was investigated in three experiments and two sets of corpus analyses. In the corpus analyses, gender information reduced the lexical search space as well as the amount of input needed to uniquely identify a word. To test whether this holds for on-line processing, two auditory lexical decision experiments (Experiments 1 and 3) were conducted using valid, invalid, or noise-masked articles as primes. Clear gender-priming effects were obtained in both experiments. Experiment 2 used phoneme monitoring with words and with pseudowords deviating from base words in one or more phonological features. Contrary to the lexical decision latencies, phoneme-monitoring latencies showed no influence of gender but did show similarity mismatch effects. We argue that gender information is not utilized early during word recognition. Rather, the presence of a valid article increases the initial familiarity of a word, facilitating subsequent responses.

As in many other languages, German employs a gender system where noun gender is (can be) indicated by a preceding article and in some constructions, gender-specific morphological affixes. The German gender-marking system not only signals the syntactic gender of a word but also provides information concerning number and case. The experiments reported here investigated the role of gender marking in spoken word recognition. We focus on the role of a preceding gender-specific article in restricting lexical hypotheses during spoken word recognition. Statistics from a corpus analysis of German were utilized to determine the degree to which gender information could, in principle, restrict lexical hypotheses.

Previous research on the role of gender marking in spoken word recognition has focused on a number of issues. In general, the effects of gender marking are usually robust and have been demonstrated in many languages across a variety of tasks (but see Friederici & Jacobsen, 1999, and discussion below for modality and task differences). The role of gender marking has been investigated with respect to where and how gender marking influences word recognition processes. The debate

mirrors that concerning the relationship between speech perception and lexical representations (Connine, Titone, Deelman, & Blasko, 1997; Norris, McQueen, & Cutler, 2000). One view holds that the influence of gender marking is subsequent to lexical activation. According to this view, gender marking is used to determine congruency relations among words. For example, Colé and Segui (1994) investigated gender marking in French using a lexical decision task. In their experiment, gender was indicated by a valid or invalid gender marking on a preceding adjective. At both a long (500-msec) and a short (150-msec) stimulus onset asynchrony (SOA), the valid adjective marking resulted in faster reaction times (RTs). Given the maintenance of the gender validity effect at the long SOA, Colé and Segui argued for a postlexical locus involving incongruency between lexical and syntactic processing. Schmidt (1986) came to similar conclusions in his investigation of the recognition of masculine and neuter words in German with visual lexical decision. He obtained inhibitory effects for invalid article–noun pairs with a list containing valid and invalid article–noun pairs. Processing was facilitated for valid article–noun pairs when the experimental lists contained only valid article primes and control primes (“xxx”). When all primes (valid article, invalid article, or control prime “xxx”) had to be produced prior to the lexical decision response, invalid article–noun pairs showed an inhibitory effect, and valid article–noun pairs a facilitatory effect. Schmidt’s results showed that gender priming facilitatory effects depend on list construction specifically for situations where the gender cue is of large predictive value for the response. Schmidt suggested that the results support an article–noun gender congruency check that operates with-

---

The phoneme-monitoring experiment was conducted in Münster while C.M.C. was an Alexander von Humboldt fellow at the Westfälische Wilhelms-Universität, Münster, and was supported by NIH Grant R01 DC02134. We thank Dennis Norris and two anonymous reviewers for their helpful comments on the manuscript. Parts of this research were presented at the 38th Annual Meeting of the Psychonomic Society, Philadelphia. Correspondence should be addressed to J. Bölte, Psychologisches Institut II, Westfälische Wilhelms-Universität Münster, Fliegerstr. 21, 48149 Münster, Germany (e-mail: boelte@psy.uni-muenster.de).

out attention. If an incongruity is detected, attention may be directed to this incongruity, slowing down task performance.

In contrast, other researchers have argued that gender information influences lexical activation. Grosjean, Domergues, Cornu, Guillelmon, and Besson (1994) investigated the role of gender marking in recognizing French words using both gating and lexical decision tasks. The gating task revealed that listeners could identify the intended word earlier (9% less signal information) when a gender-valid article was available. The influence of gender information was also evident in the set of hypotheses generated by listeners—all hypotheses shared the gender of the intended word. Similarly, lexical decision times were faster for a word preceded by a valid article. Grosjean et al. considered both a lexical locus where gender information influences word recognition processes and a postlexical locus where gender information is utilized in a syntactic checking mechanism acting on the output of lexical processing, but favored a lexical locus.

Others have argued that gender-marking effects are strongly indicative of an interactive process, whereby grammatical features (such as gender-marking articles) serve to cue lexical entries (Bates, Devescovi, Hernandez, & Pizzamiglio, 1996; see also Bates, Devescovi, Pizzamiglio, D'Amico, & Hernandez, 1995). Bates et al. (1995) suggested that this would be advantageous for richly inflected languages such as Italian, where gender agreement is marked on almost all modifiers. In the work of Bates and colleagues, the approach has been to investigate the facilitatory and inhibitory influences of gender and to use the direction of effects to make inferences concerning the automatic or controlled nature of its use. This approach follows Neely's (1991) suggestion that facilitation serves as a marker for automatic processes and inhibition serves as a marker for controlled processes.

Bates et al. (1996) utilized a battery of tasks to investigate both inhibitory and facilitatory effects of gender marking in Italian. In one task, repetition of an auditory word was faster when the target was preceded by a valid gender-marked adjective, relative to an invalid gender-marked adjective. A neutral condition showed intermediate naming times. A similar design using gender monitoring showed only facilitatory effects. Grammaticality judgments showed faster acceptance for gender-consistent phrases relative to rejection latencies for gender-inconsistent phrases. Bates et al. (1996) concluded that processing of gender information did not require that attention be specifically directed to that feature, and that its processing involved a mix of automatic (facilitation) and controlled (inhibition) processes. In similar experiments, effects of gender marking in Italian were examined in conjunction with influences of sentence context (Bentrovato, Devescovi, D'Amico, & Bates, 1999). Inhibitory effects were found only when both dimensions, gender marking and sentential context, were violated.

Colé, Pynte, and Andriamamonjy (2003) investigated the influence of gender information on word recognition in French. With visual lexical decision, they found that

masculine nouns with high-predictive endings (final bigram matching the final phoneme's orthographic transcription, e.g., *-et* signals only masculine words; *-ette* signals predominantly feminine words) had a processing advantage over masculine words with a less predictive ending (*-le*), at least for low-frequency words. No such difference was obtained for feminine words. Colé et al. attributed this difference to the fact that feminine words have overall less predictive endings than masculine words. Colé et al. also found that eye movements during reading of the article–noun combinations showed more word ending fixations on high-predictive than on low-predictive endings, regardless of noun gender. The article was fixated less often for feminine high-predictive words than for feminine low-predictive words. No such difference was observed for masculine words. Colé et al. argued that this difference is due to a gender-orthography checking mechanism that is easier for feminine words (search for *-e*) than for masculine words (search for the absence of *-e*). Overall, they suggest that French readers use gender information at different levels of word recognition that include a prelexical (predictive value of ending), a lexical (gender category), and a postlexical (morphosyntactic) level.

In an earlier eye-movement study, Dahan, Swingley, Tanenhaus, and Magnuson (2000) found that auditorily presented gender-marked articles restrict the set of lexical hypotheses in French. They also proposed that recognition is guided by sound-based statistics of the language (see Colé et al., 2003, for the influence of orthography-based regularities). Hearing the gender-marked article and the first phonemes of the target word enhances the probability of the target word (e.g., *bouton*<sub>masc</sub>, *bouteille*<sub>fem</sub>; *le bou* → *bouton*).

Our survey of the literature showed that gender information might influence processing at a number of different levels within the word recognition system. At least three different levels (in this, we follow Friederici & Jacobsen, 1999) that include lexical access (i.e., the activation of lexical units) can be distinguished, with gender information either available to influence this initial activation process or used postlexically when lexical processing is complete (see Tanenhaus & Lucas, 1987, for a discussion of various context effects in lexical processing). A third view suggests that gender information serves to preactivate gender-consistent lexical entries as, for instance, Bates et al. (1995) have suggested. Facilitatory effects can be expected under this processing view. This view, termed *intralexical effect of gender information*, is also found in production models such as that proposed by Levelt, Roelofs, and Meyer (1999). In this model, a single gender representation is activated by a gender-bearing lemma that can activate all gender-congruent lemma representations (see Schriefers & Jescheniak, 1999, for a review of gender effects in speech production).

A separation of potential influences of gender-marking effects into prelexical, intralexical, and postlexical levels has provided important guidance in this area of research. However, these distinctions are difficult to maintain as

truly separate levels when one considers the broader domain of models in auditory word recognition. Current models of auditory word recognition do not have components that fit neatly into pre-, intra- and postlexical processes. In some models, partial input serves to simultaneously activate representations at independent semantic and phonological levels (Gaskell & Marslen-Wilson, 1997). In other models, partial input is used to activate initial perceptual and lexical representations with an interactive relationship (see, e.g., McClelland & Elman, 1986; Norris, 1994). In both types of models, partial activation at a number of levels simultaneously blurs the boundaries among levels of processing.

In summary, a growing literature on processing of gender-marked languages focuses on the role of gender marking during word recognition and language comprehension. Orthogonal to these central questions is the pattern of effects as a function of modality of presentation. The literature review by Friederici and Jacobsen (1999) yielded a number of important insights. Friederici and Jacobsen noted that facilitation (in RT tasks) for correct gender marking has been a more consistent pattern for auditory presentation. In addition, Friederici and Jacobsen argued that a facilitation effect will most likely emerge with tasks with a postlexical checking component and with phonologically transparent gender-marking languages (e.g., French, Italian). Facilitation as a consequence of gender information was argued to reflect mainly lexical selection rather than lexical access (Marslen-Wilson, 1987) or preactivation (Bates & MacWhinney, 1989). In contrast, inhibition for incorrect gender marking has been found (along with facilitation) across visual and auditory presentation. The consistency of inhibitory effects across tasks led Friederici and Jacobsen to argue that inhibition reflects the operation of a postlexical gender congruency check that takes longer for incongruent pairs than for congruent pairs.

German, the language of interest here, employs three genders—masculine, feminine, and neuter. Native speakers of German can readily identify the gender of a noun in isolation, but the occurrence of one of the three articles is not always an unambiguous marker of a specific gender. Definite articles (masculine *der*, feminine *die*, neuter *das*) mark gender, number, and case. For example, the definite article *der* is used for singular masculine nominative, for singular feminine genitive and dative, and for plural genitive of all syntactic genders. Thus, depending on the particular syntactic construction, a given definite article can signal gender only with reference to the syntactic structure. Another characteristic of German is that an article might be separated from its noun by intervening words, some of which can be nouns with a different gender marking. Consider the following example: *Das<sub>neut</sub> aus höchster Not<sub>fem</sub> gerettete Kind<sub>neut</sub> freute sich* (literally translated as “The from extreme difficulties rescued child was glad”). A feminine noun, *Not*, follows the definite neuter article *das*. In this example, the article is not relevant to a following gender-bearing word but signals that there must be a neuter word occurring later

in the utterance. Adjectives and quantifiers (and verbs if used as adjectival participles) are also gender-marked: *herzlicher<sub>masc</sub> Mann*, *herzliche<sub>fem</sub> Frau*, *herzliches<sub>neut</sub> Kind*. Adjectives and quantifiers, in contrast to nouns, can take any gender marking.

The question addressed in the present experiments is whether a gender-marked article is utilized to restrict lexical hypotheses, thereby serving to decrease the lexical activation space during activation of lexical form. As a precursor to addressing this psychological processing question, we conducted a corpus analysis to determine whether gender could, in principle, serve to restrict lexical search space. Along a number of linguistic dimensions, corpus analyses have been extremely useful in providing a characterization of experiential language use for both grammatical (Connine, Ferreira, Jones, Clifton, & Frazier, 1984) and phonological dimensions of language use (Deelman & Connine, 2001; Patterson & Connine, 2001; Patterson, LoCasto, & Connine, 2003). The corpus analysis described here was intended to determine any potential reduction in lexical activation space afforded by an article. A second, related goal was to determine whether any asymmetries in distribution exist across the gender system of German. If so, more precise processing predictions may be formulated on the basis of the degree to which a given article restricts lexical hypotheses. An article that permits fewer words could provide a more restrictive set of lexical hypotheses. If this information is utilized during word recognition, more efficient processing may result than would occur with less restrictive articles. The latter hypotheses were tested in Experiments 1, 2, and 3. Third, given German word order flexibility, we wanted to determine how often article–noun combinations agree in gender. To anticipate the corpus analysis results, the corpus statistics indicate that gender information could potentially result in earlier word recognition as measured in uniqueness point (UP). Moreover, this varies across the three gender categories of German (see Colé et al., 2003, for differential gender effects in French). Furthermore, article–noun combinations very often agree in gender. The behavioral experiments tested whether the statistical likelihood for the convergence of article and noun gender information is utilized by listeners to narrow down the lexical search space.

We sought to determine whether any gender-marking effects obtained with the lexical decision task (Experiments 1 and 3) generalized to phoneme monitoring (Experiment 2). Of particular relevance is whether any facilitation found for gender-consistent trials with lexical decision will also be found for phoneme monitoring. Although lexical decision has been argued to reflect a postlexical familiarity checking component (Balota & Chumbley, 1984, 1990), there is general agreement that phoneme monitoring reflects the integration of acoustic-phonetic and lexical representations (Connine & Titone, 1996; Norris et al., 2000). If, as suggested by Friederici and Jacobsen (1999), facilitation can be shown only when a postlexical component is part of the task configuration, then an advantage for gender-consistent stimuli

should be found relative to a neutral baseline only in lexical decision (Experiment 1). Our benchmark for whether gender marking can facilitate recognition is twofold in the design of Experiment 2. One aspect parallels the logic of Experiment 1, where a neutral baseline is compared with a gender-consistent condition. In addition, Experiment 2 used nonword stimuli that varied in degree of phonological match relative to a base word. If gender operates at a prelexical level, then it may serve to mitigate the influence of mismatching information. In Experiment 3, we examined whether the nature of the available gender information influences its use.

## CORPUS ANALYSES

### Corpora

Our analyses utilized two lexical databases, CELEX (1995) and the NEGRA corpus (Brants, 1996). The former corpus consists of 51,728 German types and provides frequency counts as well as morphological, syntactic, and phonological information about the types. The NEGRA corpus consists of approximately 176,000 tokens (10,000 sentences) of German newspaper text. Morpho-syntactic information (number, case, and gender) has been determined for approximately 20,000 words. We used the NEGRA corpus despite its small size because some characteristics of the German gender system require context coding (see below).

### Analyses of the CELEX Corpus

The distributional analysis of types from the CELEX database is summarized in Table 1. If gender information is considered in defining the possible search space, the number of lexical possibilities is reduced to 11.6%–26.5% of all entries when only nouns are considered. The figures become less impressive if nouns, adjectives, and verbs are included (masculine 58.4%, feminine 64.0%, neuter 49.1%). Common for both analyses was that neuter gender limited the set of compatible lexical hypotheses more than either masculine or feminine articles (for results using only monomorphemic nouns, see Schiller & Caramazza, 2003).

Table 1 also shows the additional decrease in the lexical hypothesis space when word frequency was consid-

ered. In this analysis, each word type was weighted by its token frequency. The percentage of words remaining in the search space ranged from 5% to 9% (nouns only) when word frequency was considered.

These analyses suggest that gender information together with word frequency results in a reduction in the number of lexical hypotheses, and depending on the structure of the search space, the reduction can be quite large. However, the contribution of any type of information, for example, gender information, has to be evaluated in the context of the overall recognition process. Spoken word recognition is a process that extends over time, and the contribution of any type of context information can also be evaluated in terms of how the information becomes available to the listener over time. As pointed out by Carter (1987), percentages tend to overestimate the contribution of any information source in restricting likelihoods in selection processes. Gender information may narrow the search space to gender-consistent types, but this is clearly only a first step in selecting the correct lexical hypothesis (i.e., the one lexical entry that matches the input). Table 1 shows that knowing a feminine noun's gender reduces the set of lexical hypotheses from 51,728 to 13,667 (26.5%). However, word recognition must continue in order to select a match to the input from the remaining 13,708 lexical hypotheses. If we assume that the following processing steps result in a similar reduction, an additional six reductions are needed to finally come to one entry (13,708 → 3,622 → 960 → 254 → 67 → 18 → 1). Thus, knowing the syntactic gender reduces the search space by 73.5%, but this simplified example illustrates that knowing the syntactic gender constitutes only 12.5% of the work.

In order to consider sequential constraints in conjunction with gender information more closely, we computed the lexical search space on the basis of uniqueness points (UPs) using a phonemic transcription (DISC) from the CELEX database. The UP is the point from word onset at which only a single word remains a complete match with a sequence of segments (Bölte & Uhe, 2004; Gaskell & Marslen-Wilson, 2002; Marslen-Wilson, 1993).

Table 2 shows the average number of phonemes required for recognition on the basis of UP alone and the combination of UP and gender information. The average

**Table 1**  
Word Class and Gender Information in German

Word Class	Gender	Percentage	
		Unweighted	Weighted With Word Frequency
Nouns	Masculine	20.9	7.7
	Feminine	26.5	9.1
	Neuter	11.6	5.0
Adjectives, quantifiers	Masc., fem., neut.	19.3	10.8
Verbs	Gender marked if adjectival participle	18.2	17.4
Adverbs, conjunctions, etc.	No gender	3.1	49.8

Note—Some nouns (approximately 222) take two or all genders—for example, *See<sub>masc</sub>* (lake), *See<sub>fem</sub>* (sea). They make up the percentages missing to 100%. There were 51,728 entries in total.

**Table 2**  
**Uniqueness Points in Phonemes, as a Function of Different Partitionings of the CELEX Database**

Word Class	Gender	Type of Partitioning							
		Unpartitioned		Nouns vs. Adjectives + Verbs + Rest		Nouns + Adjectives vs. Verbs + Rest		Nouns, Adjectives, Verbs vs. Rest	
		<i>N</i>	<i>SD</i>	<i>N</i>	<i>SD</i>	<i>N</i>	<i>SD</i>	<i>N</i>	<i>SD</i>
Nouns	Masculine	6.5	1.9	5.4	1.7	5.9	1.9	6.1	1.9
	Feminine	7.0	2.2	5.5	1.8	6.4	2.2	6.7	2.2
	Neuter	6.2	1.9	4.9	1.6	5.6	1.8	5.7	1.8
Adjectives, quantifiers	Masc., fem., neut.	7.0	2.2	5.8	1.8				
Verbs	Gender marked if adjectival participle	6.5	1.8	5.7	1.7	5.2	1.6		
Adverbs, conjunctions, etc.	No gender	5.9	1.9	5.8	2.1	5.3	1.8	4.7	1.8

Note—Some nouns (approximately 222) take two or all genders and are not included in the UP calculation [e.g., *See<sub>masc</sub>* (*lake*), *See<sub>fem</sub>* (*sea*)].

length of an entry in number of phonemes in this lexicon is 8.6 (*SD* 2.8). The average UP is 6.7 (*SD* 2.1). If we consider only nouns, gender specification moved the UP on average 1.3 phonemes (or approximately 11% of the average word length) closer to a word's onset. If adjectives, verbs, and quantifiers are included along with nouns, the average reduction was .4 phonemes (approximately 3% of the average word length). A similar analysis conducted by Grosjean et al. (1994) for French showed values intermediate with respect to those reported here. However, it is important to keep in mind that different languages may in fact yield different statistical properties on the basis of both their gender and lexical structure.

In summary, less phonological information is required to decrease the lexical hypothesis set to one when a word's gender is available. Obviously, this depends on the initial set of words under consideration.

### NEGRA Corpus Analyses

The preceding analyses considered the potential role of gender specification in reducing lexical hypotheses when lexical variables of frequency and UP are taken into account. These analyses assumed an unambiguous specification of grammatical gender for a given word. However, the German word order system permits a great deal of flexibility and in some instances, connected speech or written text will contain sequentially mismatching gender information for paired nouns and adjectives. For example, as previously described, a noun (or adjective) following an article may have a different gender than that indicated by the article. In the following section, we report a corpus analysis that examines the unfolding of gender specification for words in sentences using the NEGRA corpus. Specifically, we examined the frequency with which an article is followed by a matching adjective or noun. One limitation of the NEGRA corpus is that it contains an incomplete morpho-syntactic analysis, and as a result the available tokens with sufficient gender specification numbered 4,605 (24.9% of 18,447 tokens). Thus, we view this analysis as only an approxi-

mation of the input Germans receive. Despite this limitation, it is to our knowledge the only available corpus with the necessary coding.

Of the 4,605 tokens, there were 1,502 masculine, 2,268 feminine, and 809 neuter word pairs. The article and the following word agreed in gender (as well as in number and case) in all but 26 cases. These exceptions were distributed approximately evenly across the three articles (2, or .5%, masculine; 17, or .8%, feminine; and 7, or .9%, neuter tokens). Thus, the possible configuration in German of separating an article and noun (or adjective) by intervening gender-mismatching words seems to be rarely used. It appears that Germans receive a highly consistent input where article and noun in close proximity nearly always agree.

### Discussion

In summary, the database analyses showed that an article can constrain the gender of the following word despite the flexibility of the German word order. Furthermore, the distribution across masculine, feminine, and neuter tokens indicated a somewhat better convergence of article and noun gender for neuter gender than for either masculine or feminine gender. In order to examine the use of gender in word recognition, three experiments were conducted. Experiment 1 presented German nouns preceded by a valid article or by signal-correlated noise (neutral baseline) in a lexical decision task. The purpose of Experiment 1 was to determine whether the valid (gender-matching) article condition provided an advantage for lexical decision, relative to a neutral noise baseline. Experiment 2 investigated whether gender-marking effects generalize to phoneme monitoring. Experiment 3 returned to the lexical decision task using both a valid (gender-matching) and an invalid (gender-mismatching) article in order to determine whether the nature of the available gender information influenced its use.

It is important to be aware that the selection of a baseline condition is controversial and, as Jacobsen (1999) and Friederici and Jacobsen (1999) have discussed in de-

tail, fraught with difficulties. In selecting noise as the baseline condition in Experiments 1 and 2, we sought to utilize a baseline that was neutral. Noise is certainly neutral in that it does not provide gender information, but it does not fulfill the requirement outlined by Jacobsen that a neutral baseline be matched on all other dimensions (obviously, noise is not linguistic). As a consequence, the noise baseline may introduce elements that influence the emergence of a gender-marking effect. For example, a noise baseline may enhance any gender-marking effect—by virtue of its nonlinguistic nature it may slow responses on subsequent linguistic information. Experiment 3 addresses aspects of these concerns by utilizing an invalid cue. However, a similar difficulty may arise in using a gender-invalid article in that this may discourage the use of gender or reflect processing of ungrammaticality (see Jacobsen, 1999). The idea of a neutral baseline that is matched in all but one dimension follows Donders's (1969) suggestion of "pure" insertion. This assumption, however, is not without detractors (Friston et al., 1996; Luce, 1986). Kinoshita (2002), for instance, shows that overall semantic priming effects remain constant while the RTs to neutral trials depend on list composition. Since our experiments (given the constraints on word selection) do not provide a perfect solution to this issue, we are appropriately cautious in our conclusions.

## EXPERIMENT 1

As previously discussed, Experiment 1 utilized German words that were preceded by a valid gender article or by noise. Listeners made a lexical decision to each word. The goal of the experiment was to determine whether a valid gender article facilitated RTs to words, relative to the neutral noise baseline.

### Method

**Participants.** This study included 28 native speakers of German, mostly students of the University of Münster, Germany, who received course credit for their participation (mean age 22.8 years,  $SD$  4.8; 7 men, 21 women).

**Materials.** A set of 60 gender-unambiguous common nouns, 20 of each syntactic gender, was selected (see Appendix A) in such a way that the frequency of occurrence, number of phonemes, and number of syllables did not differ between the three groups. A set of pseudoword fillers was constructed by altering the initial consonant of an existing noun.

All items (nouns, pseudowords, and articles) were spoken by a male speaker of German on a Sony TDC-D8 DAT recorder. The recording was digitized and stored on a computer disk for further processing. Onset and offset of each item were marked under visual and auditory control using CSRE 4.5. Nouns were combined with either an appropriate article or a neutral stimulus. Care was taken to ensure that the transition between article and noun was as natural sounding as possible (e.g., articles had the same length in milliseconds as articles produced in combination with nouns, and onsets and offsets were cut at zero crossings to prevent the introduction of extraneous noise). The neutral stimulus was constructed by multiplying every other sample of a valid article for a given noun with  $-1$ . The resulting signal is a noise stimulus with the same amplitude contour as the original article. Pseudoword targets were also preceded by articles. Half of the pseudoword targets were valid,

with reference to the noun the pseudoword had been made of, or neutral (noise replaced). A pulse added to the left channel marked the onset of each target. All items were transferred on DAT tapes for later presentation.

Two lists were prepared, with valid or neutral article counterbalanced across list and items. Pseudoword targets were randomly interspersed among the test trials so that no more than three word or three pseudoword answers followed each other. The same set of pseudowords was used on both lists. Both lists were preceded by 24 practice trials.

**Procedure and Apparatus.** Participants were tested in groups up to 4. Participants were instructed to perform a lexical decision on the second auditory stimulus as fast and as accurately as possible, and that their decision should be made independently of the nature of the preceding information (article or noise).

The experiment was controlled by the experimental software program NESU running on a PC (Baumann, Nagengast, & Wittenburg, 1992). All stimuli were presented from DAT tape (Sony DTC 59 ES) via headphones (Sennheiser HD 565, mono signal). A trial was set up as follows. A 250-msec warning tone signaled the start of a trial. The auditory gender cue appeared 250 msec later and was immediately followed by the auditory target. A pulse on the audio channel inaudible to the participants triggered a millisecond timer that was stopped by the participant's keypress. RT was measured from target onset, and a trial was terminated if no response was made within 1,500 msec after target onset.

### Results

A total of eight word targets were excluded from all analyses because of more than 30% erroneous answers or time-outs.<sup>1</sup> The overall error rate for words was 6.15%. Errors and time-outs (RTs longer than 1,500 msec) were excluded from all analyses. Mean RTs were submitted to separate analyses of variance (ANOVAs) with participants ( $F_1$ ) and items ( $F_2$ ). Item analyses utilized syntactic gender as a between-items variable with cue type as a within-items variable. Table 3 shows RT and percent correct as a function of syntactic gender (masculine, feminine, neuter) and cue type (valid vs. neutral).

A two (cue type)  $\times$  three (syntactic gender) ANOVA was conducted across participants ( $F_1$ ) and across items ( $F_2$ ). The ANOVA yielded a significant result for the main effect of cue type [valid vs. neutral;  $F_1(1,27) = 20.08$ ,  $MS_e = 5,081$ ,  $p < .001$ ;  $F_2(1,49) = 19.78$ ,  $MS_e = 2,854$ ,  $p < .001$ ]. Targets preceded by a valid article were responded to some 50 msec faster than targets preceded by a noise-masked article. The main effect of gender was also significant [ $F_1(2,54) = 25.73$ ,  $MS_e = 2,078$ ,  $p < .001$ ;  $F_2(2,49) = 3.32$ ,  $MS_e = 10,191$ ,  $p < .05$ ]. Masculine words (832 msec) resulted in faster RTs than did

**Table 3**  
Gender Priming With Valid and Noise Cue Types: Reaction Times (RTs, in Milliseconds),  $SD$ , and Percent Error (PE) as a Function of Condition, Experiment 1

Cue Type	Syntactic Gender											
	Masculine			Feminine			Neuter			$M$		
	RT	$SD$	PE	RT	$SD$	PE	RT	$SD$	PE	RT	$SD$	PE
Valid	813	92	2.1	863	88	5.6	851	92	8.1	842	5.3	
Neutral	851	76	5.0	918	90	7.6	906	70	8.9	892	7.2	
$M$	832		3.6	891		6.6	879		8.5			

feminine (891 msec) or neuter (879 msec) words. Feminine words did not differ from neuter words [ $t_1(27) = 1.47, p = .153; t_2(33) < 1$ ], but RTs to masculine words were significantly faster than to feminine words [ $t_1(27) = 6.534, p < .001; t_2(33) = 2.328, p = .026$ ] or to neuter words [ $t_1(27) = 5.163, p < .001; t_2(32) = 2.588, p = .014$ ]. The interaction was not significant (all  $F_s < 1$ ).

In summary, a valid article resulted in faster lexical decisions to auditory words relative to a noise-masked article. Notable was that the effect of having a preceding article provided equivalent facilitation for each gender. Thus, the asymmetry in the distribution of gender indicated by the corpus analysis was insufficient to provide an additional advantage for the more restrictive neutral gender.

## EXPERIMENT 2

At issue in Experiment 2 was the locus at which the contextual constraint provided by an article operates. As previously discussed, the lexical decision task has been argued to reflect postlexical consistency effects. The purpose of Experiment 2 was to determine whether the facilitation found with lexical decision replicates with a phoneme-monitoring task. Phoneme monitoring is a task that is very sensitive to the degree of lexical activation. Previous research has demonstrated that this manipulation serves to alter the relative activation of a lexical representation, and phoneme monitoring is sensitive to these changes in activation (Connine et al., 1997; Deelman & Connine, 2001). For example, Connine et al. (1997) found that phoneme-monitoring RTs increased as the similarity of the input decreased relative to a real word. Phoneme monitoring is also sensitive to other lexical dimensions, such as associative relationships (Frauenfelder & Segui, 1989). If the gender-consistent facilitation effect found with lexical decision does not replicate with phoneme monitoring, then this would support Friederici and Jacobsen's (1999) analysis that gender is utilized in a postlexical checking mechanism, but not during initial activation.

Experiment 2 used a set of pseudowords that mismatched a word by a single phoneme. The mismatched phoneme was either very similar (minimal pseudowords) or dissimilar (maximal pseudowords) relative to the real word from which it was derived (base word). We expected that the similarity effect found by Connine et al. (1997) would replicate with German materials (for experiments using lexical decision, see Bölte, 2001; Bölte & Coenen, 2002). With respect to a potential influence of gender in this task, the mismatch effect provides a means to establish that lexical information has been utilized in the task. A mismatch effect in the absence of a gender effect would suggest that lexical activation took place without an influence from gender information. Alternatively, the presence of a preceding article may influence phoneme-monitoring times as a function of similarity. If the presence of a valid article operates to "preactivate" only those words consistent with that gender (or to "deactivate" all gender-

incongruent words), then it may mitigate the decreased lexical activation afforded by the minimal and maximal pseudowords. This would suggest that lexical activation is influenced by gender information.

## Method

**Participants.** Seventy-nine participants of the same population as in Experiment 1 served in this experiment (18 men, 61 women).<sup>2</sup> They received course credit for their participation.

**Materials.** Three groups of German words of at least two syllables were selected. The target phoneme was the last consonant of the stimulus. In some cases, the consonant was followed by a schwa (see Appendix B). One group of 24 base words ended with the phoneme /k/, a second group of 48 base words ended with /t/ as the last consonant, and a third group of 24 base words had the phoneme /s/ as the last consonant. Because insufficient numbers of words ending in /s/ were available, this set included a subset of words in which the /s/ was preceded by a /t/. The status of this complex offset is unclear for German with analyses that categorize it both as monophonemic and as biphonemic (Kloecke, 1982; Ramers & Vater, 1992). Participants should find /s/ detection easier if /ts/ has a biphonemic status than if the phoneme sequence has monophonemic status.

The base words (e.g., *diplomat*) were used to construct two types of pseudowords (minimal: \**biplomat*; maximal: \**miplomat*). Minimal pseudowords were constructed by altering the initial phoneme by one phonological feature. The initial phoneme of a maximal pseudoword differed by two or more features from the initial phoneme of the base word.

Base word, minimal, and maximal pseudowords of each phoneme target group were counterbalanced over three lists so that one of an item set (base word, minimal pseudoword, maximal pseudoword) appeared only once on a list. For instance, *diplomat* appeared on List 1, \**biplomat* on List 2, and \**miplomat* on List 3. There were an even number of base words, minimal, and maximal pseudowords on each list. A valid article preceded 50% of the items. The remaining 50% were preceded by a neutral cue constructed as in Experiment 1 (signal-correlated noise).

A set of no-target filler words was selected for each target phoneme with the appropriate proportions of words and pseudowords, stimuli with and without target, and stimulus initial and stimulus final targets on each list. For /t/, there were 32 words and 32 pseudowords with initial targets, 16 words with final targets, and 64 words and 64 pseudowords without targets. For /s/ and for /k/, we chose 16 words and 16 pseudowords with initial targets, 8 words with final targets, and 32 words and 32 pseudowords without targets, respectively. Again, 50% were preceded by the valid article, and the other half by signal-correlated noise. Trials were randomized without restrictions. There were 20 practice trials.

Stimuli were recorded by a female speaker of German on a Sony TDC-D8 DAT recorder. This recording was digitized at 22.05 kHz with 16-bit resolution and stored on a computer disk for further processing. A waveform editor (CSRE 4.5) was used to mark the onset and offset of each stimulus using both auditory and visual information. Timing marks for phoneme onset were conventionalized to the closure onset for plosives and the frication onset for fricatives. Items were then converted to 20 kHz (12-bit resolution) for later presentation.

**Procedure and Apparatus.** One participant was tested per session. Participants were informed using written instructions that they would be hearing a list of unrelated auditory stimuli. Their task was to listen for a specified target phoneme and to press a button as fast as possible when the target phoneme was detected. Accuracy and speed of detection were both emphasized. Items were presented via headphones (Sennheiser HD 565, mono signal) from computer disk. The target phoneme was presented visually to the participants

using a computer monitor (Sony Multiscan 100 SF). Each participant received three blocks, with the target phoneme held constant within a block. Block order was counterbalanced across participants as much as possible. Twenty warmup trials preceded each block. Feedback about the error rate and average RT were presented after each warmup list. There was a short break after each block. A trial was organized in the following manner: A 250-msec warning tone was presented, followed 250 msec later by the valid article or the neutral noise-masked article and then the test item. RTs were measured from target onset for 1,750 msec.

**Results**

Only target-final items were considered in the following analyses. Five participants were excluded from the analyses, 1 because of computer failure and 4 because of error rates greater than 15%. Two *t*-final item sets were discarded because they had minimal pseudowords erroneously classified as maximal pseudowords. Because of the undetermined phonemic status of /ts/, we first examined whether items ending in /ts/ were different from items ending in /s/. The error rates for /t/, /k/, and /s/ (9.5%, 8%, and 12.1%) patterned similarly, while the error rate for /ts/ was much higher (34%). This pattern for /ts/ suggests that the representation of /ts/ is as a single unit rather than as a sequence of /t/ and /s/. Since these stimuli appeared to be treated differently and had very high error rates, we did not include them in any further analyses. Table 4 displays the error rates and phoneme-monitoring RT from phoneme target onset. All RTs less than 100 msec were excluded from the analyses (5.8% of all valid responses), and neither time-outs nor erroneous responses were replaced.

An ANOVA with the variables target phoneme (/t/, /k/, /s/), cue type (valid article, neutral noise-masked article), and degree of deviation (word, minimal pseudoword, maximal pseudoword) was conducted. The independent variables were all within factors in the  $F_1$  analysis. The  $F_2$  analysis was a mixed design with the factor phoneme as a between factor, and cue type and degree of deviation as within factors.

Table 5 shows the complete ANOVA output. No main effect of target phoneme (/t/ 437 msec, /k/ 445 msec, /s/ 453 msec) or cue type (valid, 442 msec; neutral, 448 msec)

was found.<sup>3</sup> If we assume a directionality of cue type based on the results of Experiment 1, one-direction statistical analyses showed that cue type effect missed significance in  $F_1$  [ $F_1(1,73) = 1.326, MS_e = 10,285, p = .127$ ], but was significant in  $F_2$  [ $F_2(1,85) = 3.347, MS_e = 5,363, p = .036$ ].

Table 5 also shows a main effect of degree of deviation. Phonemes were detected faster in words (391 msec) than in minimal pseudowords (443 msec) and maximal pseudowords (501 msec). A Tukey HSD revealed a critical difference of 41.8 msec (participants,  $df = 3,142, p \leq .05$ ) or 40.8 msec (items,  $df = 3,170, p \leq .05$ ). This main effect was further qualified by a significant interaction of degree of deviation with target phoneme. For /t/, phoneme-monitoring time increased by 70 msec from word to minimal pseudoword [ $t_1(73) = 8.905, p < .001; t_2(45) = 6.411, p < .001$ ] and an additional 20 msec in the maximal pseudoword condition [ $t_1(73) = 2.545, p = .013; t_2(45) = 2.560, p = .014$ ]. Monitoring times for /k/ were 37 msec slower in minimal pseudowords than in words [ $t_1(73) = 3.245, p = .002; t_2(23) = 4.523, p < .001$ ] and showed a further increase of 137 msec for maximal pseudowords [ $t_1(73) = 11.483, p < .001; t_2(23) = 2.852, p = .009$ ]. An /s/ was detected faster in words than in minimal pseudowords [ $t_1(73) = 4.122, p < .001; t_2(17) = 3.908, p = .001$ ]. Again, monitoring times were slowest in maximal pseudowords [ $t_1(73) = 2.177, p = .033; t_2(17) = 2.441, p = .026$ ]. No other effects were significant.

The ANOVA showed no significant main effect for cue type or any significant interaction with cue type. An inspection of the RTs suggests a facilitation (valid vs. neutral) for maximal pseudowords across all phonemes (26 msec, 18 msec, and 25 msec for /t/, /k/, and /s/, respectively). No effect at all appeared for minimal pseudowords (/t/, 8 msec; /k/, -2 msec; /s/, 4 msec). For word targets, a mixed picture of facilitation and inhibition emerged (/t/, 19 msec; /k/, -24 msec; /s/, -15 msec). This pattern suggests that the size and the direction of the effect of a valid cue might depend on the degree of mismatch. That is, a valid cue supports phoneme detection for maximal pseudowords, but hinders phoneme detection for words. However, the results of the ANOVA do not

**Table 4**  
**Phoneme-Monitoring Reaction Times (RTs, in Milliseconds), SD, and Percent Error (PE)**  
**as a Function of Target Phoneme (t, k, s), Cue Type (Valid vs. Neutral), and Degree of Deviation**  
**(Word, Minimal, and Maximal Mismatch Pseudoword), Experiment 2**

Target Phoneme	Cue Type	Degree of Deviation										
		Word			Minimal			Maximal			<i>M</i>	
		RT	SD	PE	RT	SD	PE	RT	SD	PE	RT	PE
<i>t</i>	Valid	376	99	6.5	450	122	13.0	461	100	10.9	429	10.1
	Neutral	395	97	6.0	458	91	8.9	487	108	11.5	446	8.8
<i>k</i>	Valid	387	158	4.5	413	124	8.3	539	167	10.9	446	7.9
	Neutral	363	122	6.1	411	143	9.3	557	134	9.0	444	8.1
<i>s</i>	Valid	422	155	7.8	463	125	14.4	469	161	16.8	442	13.0
	Neutral	407	135	6.9	467	149	15.0	494	180	11.6	456	11.2
<i>M</i>	Valid	395		6.3	442		11.9	490		10.3	442	9.5
	Neutral	388		6.3	445		11.1	513		9.4	449	8.9
Mean		392		6.3	444		11.5	502		11.8		



**Table 5**  
ANOVA Results ( $F_1$  and  $F_2$ ) Over Reaction Times in  
Phoneme Monitoring With Gender Information

	$df$	$F$	$MS_e$	$p$
Main effects				
Phoneme	2,146	1.729	16,578	.181
	2,85	<1	57,855	
Cue type	1,73	1.326	10,285	.253
	1,85	3.347	5,363	.071
Deviation	2,146	113.145	11,803	.001
	2,170	45.479	13,361	.001
Interactions				
Phoneme $\times$ cue type	2,146	1.258	9,320	.287
	2,85	<1	5,363	
Phoneme $\times$ deviation	4,292	8,822	22,573	.001
	4,170	13,361	3,862	.005
Cue type $\times$ deviation	2,146	11,899	2,085	.128
	2,170	3,278	1,550	.215
Phoneme $\times$ cue type $\times$ deviation	4,292	<1	12,301	
	4,170	<1	3,278	

Note—Results of the  $F_1$  analysis are mentioned first; results of the  $F_2$  analysis are mentioned second.  $F_1$  was a within design,  $F_2$  a mixed design (between—phoneme; within—cue type, degree of deviation).

support this interpretation, and none of the nine paired  $t$  tests reached significance after Bonferroni correction.<sup>4</sup>

Overall, the phoneme-monitoring results replicated other findings showing graded RT effects as a function of similarity to a real word (Connine et al., 1997; Deelman & Connine, 2001), along with some variability across phoneme type. This is a useful cross-language replication, but it also served to demonstrate that listeners were sensitive to lexical information in the task. Despite this evidence that lexical information influenced phoneme-monitoring RTs, the overall advantage for a valid cue relative to the neutral baseline was small and approached significance only assuming a one-tailed test. Further, there was no indication that the valid cue served to mitigate the reduction in activation due to mismatching segments. This indicates that gender information was not utilized during early spoken word recognition.

### EXPERIMENT 3

The combined results of Experiments 1 and 2 support the notion that gender information is utilized during word recognition but only subsequent to lexical access. In Experiment 3, we sought to investigate the nature of the late use of gender information demonstrated in Experiment 1. Specifically, Experiment 3 investigated whether the use of gender marking shown in Experiment 1 is influenced by the nature of the surrounding stimuli. Recall that Experiments 1 and 2 used noise as a neutral baseline. Despite the advantages of a noise baseline, one concern is that this may influence the degree to which gender is used. Previous work suggests that the gender-marking effect can be influenced by list composition (e.g., Schmidt, 1986, for the visual domain). Experiment 3 utilized a different baseline, an invalid preceding article, to determine whether the usage of gender is influenced by the consis-

tency of gender-marking information. The valid condition, a valid preceding article, remained the same for Experiments 1 and 3. We anticipated that, as in Experiment 1, the valid condition would result in faster lexical decision RTs than the invalid condition. However, if the presence of the invalid article results in decreased utilization of gender, then the overall validity effect should be smaller in Experiment 3 than in Experiment 1. Relatedly, a cross-experiment comparison of the valid condition should show slower lexical decision RTs in Experiment 3.

### Method

**Procedure, Design, and Apparatus.** We used the same set of targets in this experiment. The noise condition was replaced with an invalid article—that is, an article inconsistent with the gender of the following noun. The procedure, design, and apparatus were identical to those of Experiment 1.

**Participants.** There were 30 participants in this experiment from the same population as in Experiment 1 (7 men and 23 women; mean age 24.8 years,  $SD$  9.1).<sup>5</sup> Each participant received course credit for participation.

### Results

Nine word targets were discarded in this experiment because of high error rates ( $>30\%$  erroneous answers). Seven of these nine targets had also been excluded in the previous experiment due to high error rates. RTs were discarded on the basis of the same criteria as in Experiment 1 (5.82%). Table 6 shows the mean lexical decision times along with standard deviations and error rates as a function of cue type (valid, invalid) and syntactic gender (masculine, feminine, neuter).

Overall, a valid cue (847 msec) resulted in faster RTs than an invalid cue (919 msec). A two-way ANOVA with cue type (valid–invalid) and syntactic gender as factors showed a significant main effect of cue type [ $F_1(1,29) = 34.15$ ,  $MS_e = 6,766$ ,  $p < .001$ ; and  $F_2(1,48) = 63.60$ ,  $MS_e = 1,808$ ,  $p < .001$ ]. As in Experiment 1, RTs to masculine nouns were faster than to either feminine or neuter nouns (859, 892, and 898 msec, respectively). This main effect of syntactic gender was significant across participants [ $F_1(2,58) = 10.59$ ,  $MS_e = 2,519$ ,  $p < .001$ ], but not across items [ $F_2(2,48) = 1.81$ ,  $MS_e = 11,872$ ,  $p = .175$ ]. Finally, the cue type effect showed some variation across gender type in that the priming effect was larger for feminine (85 msec) and neuter (79 msec) cues than for masculine (51 msec) cues. The ANOVA yielded a significant interaction across participants

**Table 6**  
Gender Priming With Valid and Invalid Cues: Reaction Times  
(RTs, in Milliseconds),  $SD$ , and Percent Error (PE) as a  
Function of Condition, Experiment 3

Cue Type	Syntactic Gender											
	Masculine			Feminine			Neuter			$M$		
	RT	$SD$	PE	RT	$SD$	PE	RT	$SD$	PE	RT	PE	
Valid	833	94	.8	850	66	5.9	859	78	8.5	847	5.1	
Invalid	885	87	2.5	935	79	10.2	938	68	6.3	919	6.3	
$M$	859		1.7	893		8.1	899		7.4			

$[F_1(2,58) = 3.35, MS_e = 1,471, p < .05]$  but not across items  $[F_2(2,48) = 1.74, MS_e = 1,808, p = .186]$ .

A comparison of the cue effects across Experiments 1 and 3 shows, first, that the valid condition showed remarkably similar lexical decision times (842 vs. 847 msec, respectively;  $t < 1$ ). Thus, the introduction of invalid cues did not diminish the utilization of gender information. The postlexical congruency component of lexical decision appears to be impervious to the presence of erroneous gender specification. There is, however, a numerical difference between the neutral (892 msec) and the invalid (919 msec) condition. In order to assess whether the invalid and neutral baselines differed in their relative effects, a between-experiments ANOVA was conducted. If the invalid cue was more disruptive than the noise cue, then this would emerge as a larger difference relative to the valid cue. However, the cross-experiment ANOVA showed that the interaction of cue type and experiment was not significant ( $F < 1$ ) and suggests that the baseline did not influence the size of the validity effect. The consequences of the noise and invalid cue were also assessed directly by comparing these conditions across experiments. If the invalid cue functioned to inhibit processing relative to the neutral baseline, then responses should be slower than those of the noise condition. This comparison was also not significant [ $t(56) = 1.35, p = .18$ ].

## Discussion

In summary, valid gender cues facilitate lexical decisions to words. This result is broadly consistent with the lexical statistics reported above, which suggested that articles can narrow down the number of lexical possibilities. Gender appears to be a reliable cue and when it is available, it can influence processing of a following noun in a facilitatory fashion regardless of the type of baseline. Valid gender information also facilitated lexical decision times when compared with (and in the presence of) invalid gender information in Experiment 3. However, cross-experiment comparisons for the invalid and noise control conditions showed no statistical support for an inhibitory effect given mismatching gender information.<sup>6</sup>

## GENERAL DISCUSSION

The present experiments examined the role of article gender marking on spoken word recognition in German. Initial corpus analyses indicated that gender-marked articles can provide a potential source of narrowing lexical activation space, particularly when combined with other assumptions concerning UP and frequency. Experiment 1 showed a facilitation for gender-marked words in lexical decision compared with a neutral noise condition. However, the asymmetry across gender type evident in the corpus statistics did not modulate the gender facilitation effect—the facilitation effect was comparable irrespective of the article's relative constraint. Experiment 2 used phoneme monitoring and revealed little evidence for any facilitation in responding to words presented with an article relative to the noise baseline. The phoneme-monitoring

results replicated other findings showing lexical similarity effects (Connine et al., 1997), but the similarity effect was not modulated by the presence of an article. Taken together, the results suggest that the influence of gender marking does not serve to constrain lexical activation space, nor does it serve to compensate for reduced lexical activation due to mismatching information in the input. Rather, the combined set of results points to an influence of gender specification subsequent to word recognition.

In Experiment 3, we sought to determine whether the gender-marking effects observed in Experiment 1 could be influenced by the nature of the available gender information. The valid condition was virtually identical, irrespective of the nature of the other trials. The cross-experiment comparison of the control conditions in Experiments 1 and 3 showed only weak support for an inhibitory influence of gender marking. The lack of (or weak) inhibitory effects in these experiments contrasts with other findings in the literature. For example, with auditory stimulus presentation, Bates et al. (1996) found inhibitory effects in word repetition and gender-monitoring tasks in a language with two gender classes (Italian). The latter task showed the inhibitory effect in the absence of any facilitatory effect. However, this experiment provided gender information in a marked adjective heard prior to the target. Others have found that a gender-marking effect relative to an invalid baseline depended on the task. Carello, Lukatela, and Turvey (1988), investigating Serbo-Croatian, showed valid versus invalid gender cue effects in visual lexical decision but not in naming (see also Gurjanov, Lukatela, Moskovljevic, Savic, & Turvey, 1985). They suggested that the dependence of the gender-marking effect on using the lexical decision task reflected its postlexical locus. In summary, the available literature presents a mixed picture for inhibitory effects, and the boundary conditions may involve the nature of the gender-marking system in a given language, the invoking of grammatical representations, the nature of the task, and/or the stimulus modality (as suggested by Friederici & Jacobsen, 1999). In the following, we focus on the facilitatory effects of gender found in Experiments 1 and 3, along with its absence in Experiment 2.

Consistent across Experiments 1 and 3 was a facilitatory influence of gender agreement relative to the baseline conditions; we will now consider a postlexical explanation for these findings in light of the lexical decision task. Some researchers have argued that lexical decision involves an assessment of familiarity that occurs independently of the time course of word recognition. This view has been most highly developed in papers by Balota and Chumbley (1984, 1990), where lexical frequency, in particular, is characterized as an information source that is utilized after a word has been recognized. According to Balota and Chumbley, high-frequency words enjoy an advantage in an initial familiarity assessment that obviates any further detailed processing to identify the item as a word. Low-familiarity words require a more detailed analysis (and thus slower lexical decision times) because of their initial low familiarity. A similar mechanism for

gender information would assume that the presence of a valid article increases the initial familiarity of a word and subsequently facilitates responses. Neutral or invalid articles provide no such increase in familiarity. Tasks without a familiarity component (e.g., phoneme monitoring) to their operation will show no gender influences. This account only assumes that the lexicon is organized so that spoken words have a grammatical gender associated with them and that this information is available once a word is recognized.

The comparable pattern of results for Experiments 1 and 3 is important because it suggests that the effect is a consequence of facilitated responding given consistent gender information, and not simply an inhibitory effect in the baseline condition. Recall that in some accounts of gender utilization, processing in the baseline condition is hindered because a gender congruency check is slowed (see Friederici & Jacobsen, 1999). The noise-masked baseline is particularly informative because no gender congruency check is possible (obviously, noise provides no gender information), yet a difference was nevertheless found relative to the gender-consistent condition. In the absence of additional assumptions, a gender congruency check is unable to accommodate this finding. As noted in the introduction, numerous issues have been raised concerning the appropriateness (or lack thereof) of the baseline. In some ways, a noise baseline may be the best compromise possible. Although it differs from the valid gender condition in that it is nonlinguistic, it has the obvious advantage of providing no gender information at all. In this latter respect, a noise baseline is similar to silence. Relevant here is evidence from other domains (language production) that have wrestled with similar baseline issues and have had some success with a noise baseline. Using a language production interference paradigm, Schriefers, Meyer, and Levelt (1990) found that a noise baseline was similar to silence and that both of these conditions were less disruptive than an unrelated linguistic stimulus.

In German, a potential alternative to noise is to use adjectives that do not take a gender-specific ending, such as *lila* (purple). This possible baseline has a limited use, however, since there are very few adjectives that do not take gender-specific endings. Furthermore, the limited number of words available may create semantic anomalies in combination with many nouns. The presence of a semantic anomaly may trigger processes irrelevant to gender processing and complicate interpretation of data patterns. Alternatively, indefinite articles for masculine or neuter nouns could be used, since they share the same indefinite article *ein*. This is a potentially interesting manipulation for further research, and our account predicts that valid definite articles (e.g., *der*) should facilitate lexical decision relative to the less restricting indefinite article—*ein*—condition.

Although the corpus analyses suggested a potential role for gender in narrowing down lexical hypotheses, the phoneme-monitoring results showed no gender in-

fluences. Further, none of the more fine grain predictions based on the distribution of words across genders were borne out in Experiments 1 and 3. These results suggest that gender is not used at an early point during spoken word recognition processes—that is, prelexical processing or lexical access. In this respect, our findings and conclusions differ from those who argue for an interactive and early use of gender in word recognition (Bates et al., 1996). However, our claims are generally consistent with views of language architecture where contextual information (sentence context) is utilized subsequent to word recognition (Connine, 1987; Zwitserlood, 1989). Our findings are also consistent with models whose focus is more specifically on gender information (Friederici & Jacobsen, 1999) and with work in other domains, such as syntactic parsing. Mitchell, Cuetos, Corley, and Brysbaert (1995) have argued that grammatical gender information in Dutch that might serve to disambiguate syntactic ambiguities is not utilized in syntactic parsing. Others have argued that this information is effective but delayed in its utilization relative to its availability (van Berkum, Brown, & Hagoort, 1999). Irrespective of the resolution of this particular debate, potentially constraining information embodied in grammatical gender does not have immediate effects in parsing. From this view, its ineffectiveness in influencing relative activation of lexical representations is not surprising.

Given that gender does not appear to facilitate word recognition processes, does it have a role outside of that tied to a particular task? The familiarity check envisioned by Balota and Chumbley (1984, 1990) may have a parallel in language processing in the absence of an explicit decision. The presence of a gender marker may serve to increase the familiarity of a new role player and to decrease the likelihood that a more intensive analysis is required for that word. Relatedly, an increased familiarity via gender marking of an already mentioned role player may permit processing resources to be devoted to computation of congruency relations, particularly in spoken language, where arguments may be separated by several intervening words. In this view, gender serves a dual processing role—it provides redundant information to assess familiarity, and this in turn facilitates assignment of actions and characteristics to discourse players.

The corpus statistics on the distribution of grammatical gender in German provided a useful basis upon which to confirm the potentially constraining nature of a preceding article in experiential usage. Moreover, the asymmetry evident in the corpus statistics provided a means to generate a set of specific hypotheses concerning variability in the degree to which a given gender may effectively constrain lexical hypotheses. The more general hypothesis was confirmed in the finding of facilitated lexical decisions given valid grammatical gender, but the more specific hypothesis concerning neuter versus male/female gender was not. One possible conclusion based on the inability of corpus analyses to reflect the more precise distributional facts in processing German is that listen-

ers either do not encode these more subtle statistical facts of their language, or that this information is encoded but not utilized. However, the promise of corpus statistics renders this too pessimistic a conclusion. It may be that the distributional facts of German do not permit the more fine grained differentiation of gender to emerge in processing. Other languages with more constraining (or larger numbers of) gender specification may reveal a different picture regarding the relationship between distributional statistics and processing. Clearly, more cross-language research using the corpus-based approach is needed, particularly in languages with a richer gender vocabulary.

## REFERENCES

- BALOTA, D. A., & CHUMBLEY, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 340-357.
- BALOTA, D. A., & CHUMBLEY, J. I. (1990). Where are the effects of frequency in visual word recognition tasks? Right where we said they were! Comment on Monsell, Doyle, and Haggard (1989). *Journal of Experimental Psychology: General*, **119**, 231-237.
- BATES, E., DEVESCOVI, A., HERNANDEZ, A., & PIZZAMIGLIO, L. (1996). Gender priming in Italian. *Perception & Psychophysics*, **58**, 992-1004.
- BATES, E., DEVESCOVI, A., PIZZAMIGLIO, L., D'AMICO, S., & HERNANDEZ, A. (1995). Gender and lexical access in Italian. *Perception & Psychophysics*, **57**, 847-862.
- BATES, E., & MACWHINNEY, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3-73). Cambridge: Cambridge University Press.
- BAUMANN, H., NAGENGAST, J., & WITTENBURG, P. (1992). *A new experimental set-up, a tool for experimenting in the 90s*. Paper presented at the International Congress of Social Science Information Technology, Amsterdam.
- BENTROVATO, S., DEVESCOVI, A., D'AMICO, S., & BATES, E. (1999). Effect of grammatical gender and semantic context on lexical access in Italian. *Journal of Psycholinguistic Research*, **28**, 677-693.
- BERKUM, J. J. A. VAN, BROWN, C. M., & HAGOORT, P. (1999). When does gender constrain parsing? Evidence from ERPs. *Journal of Psycholinguistic Research*, **28**, 555-571.
- BÖLTE, J. (2001). Graded lexical activation by pseudowords in cross-modal semantic priming: Spreading of activation, backward priming, or repair? In J. D. Moore & K. Stenning (Eds.), *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (pp. 86-91). Mahwah, NJ: Erlbaum.
- BÖLTE, J., & COENEN, E. (2002). Is phonological information mapped onto semantic information in a one-to-one manner? *Brain & Language*, **81**, 384-397.
- BÖLTE, J., & UHE, M. (2004). When is all understood and done? The psychological reality of the recognition point. *Brain & Language*, **88**, 133-147.
- BRANTS, T. (1996). *TnT—Statistical part-of-speech tagging* (Tech. Rep.). Saarbrücken, Germany: University of the Saarland, Computational Linguistics.
- CARELLO, C., LUKATELA, G., & TURVEY, M. T. (1988). Rapid naming is affected by association but not by syntax. *Memory & Cognition*, **16**, 187-195.
- CARTER, D. M. (1987). An information-theoretic analysis of phonetic dictionary access. *Computer Speech & Language*, **2**, 1-11.
- CELEX (1995). [German database, Release D25.] Nijmegen: Centre for Lexical Information.
- COLÉ, P., PYNTE, J., & ANDRIAMAMONJY, P. (2003). Effect of grammatical gender on visual word recognition: Evidence from lexical decision and eye movement experiments. *Perception & Psychophysics*, **65**, 407-419.
- COLÉ, P., & SEGUI, J. (1994). Grammatical incongruity and vocabulary types. *Memory & Cognition*, **22**, 387-394.
- CONNINE, C. M. (1987). Constraints on interactive processes in auditory word recognition: The role of sentence context. *Journal of Memory & Language*, **26**, 527-538.
- CONNINE, C. M., FERREIRA, F., JONES, C., CLIFTON, C., JR., & FRAZIER, L. (1984). Verb frame preferences: Descriptive norms. *Journal of Psycholinguistic Research*, **13**, 307-319.
- CONNINE, C. M., & TITONE, D. (1996). Phoneme monitoring. *Language & Cognitive Processes*, **11**, 635-645.
- CONNINE, C. M., TITONE, D., DEELMAN, T., & BLASKO, D. (1997). Similarity mapping in spoken word recognition. *Journal of Memory & Language*, **37**, 463-480.
- DAHAN, D., SWINGLEY, D., TANENHAUS, M. K., & MAGNUSON, J. S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory & Language*, **42**, 465-480.
- DEELMAN, T., & CONNINE, C. M. (2001). Missing information in spoken word recognition: Non-released stop consonants. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 656-663.
- DONDERS, F. C. (1969). Over de snelheid van psychische processen [On the speed of mental processes], *Acta Psychologica*, **30**, 412-431.
- FRAUENFELDER, U. H., & SEGUI, J. (1989). Phoneme monitoring and lexical processing: Evidence for associative context effects. *Memory & Cognition*, **17**, 134-140.
- FRIEDERICI, A. D., & JACOBSEN, T. (1999). Processing grammatical gender during language comprehension. *Journal of Psycholinguistic Research*, **28**, 467-484.
- FRISTON, K. J., PRICE, C. J., FLETCHER, P., MOORE, C., FRACKOWIAK, R. S. J., & DOLAN, R. J. (1996). The trouble with cognitive subtraction. *NeuroImage*, **4**, 97-104.
- GASKELL, G. M., & MARSLÉN-WILSON, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language & Cognitive Processes*, **12**, 613-656.
- GASKELL, G. M., & MARSLÉN-WILSON, W. D. (2002). Representation and competition in the perception of spoken words. *Cognitive Psychology*, **45**, 220-266.
- GROSJEAN, F., DOMMERGUES, J.-Y., CORNU, E., GUILLELMO, D., & BESSON, C. (1994). The gender-marking effect in spoken word recognition. *Perception & Psychophysics*, **56**, 590-598.
- GURJANOV, M., LUKATELA, G., MOSKOVLEVIC, J., SAVIC, M., & TURVEY, M. T. (1985). Grammatical priming of inflected nouns by inflected adjectives. *Cognition*, **19**, 55-71.
- JACOBSEN, T. (1999). Effects of grammatical gender on picture and word naming: Evidence from German. *Journal of Psycholinguistic Research*, **28**, 499-514.
- KINOSHITA, S. (2002, November). *Estimating the baseline in semantic priming effect*. Paper presented at the 43rd Annual Meeting of the Psychonomic Society, Kansas City, MO.
- KLOECKE, W. U. S. (1982). *Phonologie und Morphologie* [Phonology and morphology]. Tübingen, Germany: Niemeyer.
- LEVELT, W. J. M., ROELOFS, A., & MEYER, A. S. (1999). A theory of lexical access in speech production. *Behavioral & Brain Sciences*, **22**, 1-75.
- LUCE, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.
- MARSLÉN-WILSON, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, **25**, 71-102.
- MARSLÉN-WILSON, W. D. (1993). Issues of process and representation in lexical access. In G. T. M. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing: The Second Shillcock Meeting* (pp. 187-210). Mahwah, NJ: Erlbaum.
- MCCLELLAND, J. L., & ELMAN, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**, 1-86.
- MITCHELL, D. C., CUETOS, R., CORLEY, M. M. B., & BRYLSBAERT, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (non-lexical) statistical records. *Journal of Psycholinguistic Research*, **24**, 469-488.
- NEELY, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264-336). Hillsdale, NJ: Erlbaum.

- NORRIS, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, **52**, 189-234.
- NORRIS, D., McQUEEN, J., & CUTLER, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral & Brain Sciences*, **23**, 299-370.
- PATTERSON, D., & CONNINE, C. M. (2001). A corpus analysis of variant frequency in American English flap production. *Phonetica*, **58**, 254-275.
- PATTERSON, D., LoCASTO, P. C., & CONNINE, C. M. (2003). Corpora analyses of frequency of schwa deletion in conversational American English. *Phonetica*, **60**, 45-69.
- RAMERS, K. H., & VATER, H. (1992). *Einführung in die Phonologie* [Introduction to phonology]. Hürth-Efferen, Germany: Gabel Verlag.
- RATCLIFF, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, **114**, 510-532.
- SCHILLER, N., & CARAMAZZA, A. (2003). Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory & Language*, **48**, 169-194.
- SCHMIDT, R. (1986). Was weiß der Artikel vom Hauptwort? Ein Beitrag zur Verarbeitung syntaktischer Beziehungen beim Lesen [What does the article know about the noun? A contribution about the processing of syntactic relations during reading]. *Zeitschrift für Experimentelle Psychologie*, **33**, 150-163.
- SCHRIEFERS, H., & JESCHENIAK, J. (1999). Representation and processing of grammatical gender in language production: A review. *Journal of Psycholinguistic Research*, **28**, 575-600.
- SCHRIEFERS, H., MEYER, A., & LEVELT, W. J. M. (1990). Exploring the time-course of lexical access in language production: Picture-word interferences studies. *Journal of Memory & Language*, **29**, 86-102.
- TANENHAUS, M. K., & LUCAS, M. M. (1987). Context effects in lexical processing. *Cognition*, **25**, 213-234.
- ZWITZERLOOD, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, **32**, 25-64.

## NOTES

1. The frequency counts of the three groups did not differ statistically after deletion ( $F_2 < 1$ ). Means were as follows: masculine, 3.3 ( $SD = .6$ ); feminine, 5.0 ( $SD = 1.4$ ); neuter, 6.7 ( $SD = 7.2$ ). The number of phonemes or syllables also did not differ after item deletion (both  $F_s < 1$ ). Means were as follows for phonemes: masculine, 5.4 ( $SD = 1.4$ ); feminine, 5.2 ( $SD = 1.3$ ); neuter, 5.2 ( $SD = 1.7$ ). For number of syllables, means were as follows: masculine, 1.8 ( $SD = .6$ ); feminine, 2.1 ( $SD = .5$ ); neuter, 2.0 ( $SD = .7$ ).

2. Age was not recorded. We tested more participants than in the previous experiment because of the error rates that became evident during experimentation and to achieve an even distribution of participants across the different successions of the three lists.

3. Given the RT variance, we conducted two further ANOVAs using outlier-corrected data (mean  $\pm 1.5 SD$  and  $200 < RT < 1,200$ ). The main effect cue type was affected most by outlier correction. It was marginally significant [ $F_1(1,68) = 3.927$ ,  $MS_e = 9,821$ ,  $p = .052$ ;  $F_2(1,85) = 3.772$ ,  $MS_e = 5,820$ ,  $p = .055$ ], with  $\pm 1.5 SD$  corrected data. Using the other outlier correction procedure, this main effect was far from significant [ $F_1(1,66) < 1$ ;  $F_2(1,85) < 1$ ]. The same pattern of significant results was obtained for the remaining main effects and interactions. We conclude from this result that cue type does not contribute reliably to the processes underlying phoneme-monitoring latencies, following Ratcliff's (1993) suggestion.

4. Two paired  $t$  tests were significant before Bonferroni correction [word /t/:  $t_1(73) = -2,210$ ,  $p = .03$ ; max pw /t/:  $t_1(73) = -2,548$ ,  $p = .013$ ;  $t_2(45) = -2.164$ ,  $p = .04$ ].

5. Five participants refused to report their age.

6. One might think that combining two baselines in one experiment would be superior to the approach we took. Kinoshita (2002), however, shows that such an approach is not necessarily a better solution for obtaining a neutral baseline.

## APPENDIX A

**Material Used in Experiments 1 and 3**  
**(Frequency per 1 Million [CELEX Database], Number of Phonemes, Number of Syllables)**

Target Word	Syntactic Gender											
	Masculine			Feminine			Neuter					
	Freq.	No. Pho.	No. Syll.	Target Word	Freq.	No. Pho.	No. Syll.	Target Word	Freq.	No. Pho.	No. Syll.	
Bagger	4	5	2	Bande	6	5	2	Delta	2	5	2	
Balkon	15	6	2	Bohne	6	4	2	Gestell	2	6	2	
Bussard	2	6	2	Buche	4	4	2	Handikap	2	8	3	
Galgen	4	6	2	Dose	5	4	2	Hormon	3	6	2	
Gnom	2	4	1	Grafik	7	6	2	Kalb	7	4	1	
Henkel	2	6	2	Hüfte	7	5	2	Kleinod	2	6	2	
Karneval	5	8	3	Justiz	15	6	2	Konto	14	5	2	
Knoten	5	6	2	Kelle	3	4	2	Krokodil	2	8	3	
Kobold	3	6	2	Kontur	5	6	2	Laub	5	3	1	
Modus	2	5	2	Leine	7	4	2	Lot	2	3	1	
Nabel	2	5	2	Minorität	2	9	4	Maul	7	3	1	
Rum	3	3	1	Mission	6	6	2	Mikrofon	5	8	3	
Scheck	10	3	1	Pflaume	3	5	2	Mosaik	5	6	3	
Schimpanse	2	8	3	Pforte	6	5	2	Paket	13	5	2	
Schinken	3	6	2	Schwalbe	3	6	2	Pensum	2	6	2	
Spatz	4	4	1	Seife	7	4	2	Rheuma	5	4	2	
Spion	7	5	1	Skizze	5	5	2	Sofa	7	4	2	
Wimpel	3	6	2	Steppe	14	5	2	Terrain	3	4	2	
Wodka	4	5	2	Traube	3	5	2	Turnier	15	6	2	
Zar	7	3	1	Woge	4	4		Veilchen	2	6	2	
<i>M</i>	4.5	5.3	1.8	<i>M</i>	5.9	5.1	2.1	<i>M</i>	5.3	5.3	2.0	
<i>SD</i>	3.2	1.4	0.6	<i>SD</i>	3.3	1.2	0.4	<i>SD</i>	4.2	1.6	0.6	

Note—Freq. = frequency/1 million,  $F_2 < 1$ ; No. Pho. = number of phonemes,  $F_2 < 1$ ; No. Syll. = number of syllables,  $F_2(2,57) = 1.4$ ,  $p = .255$ .

**APPENDIX B**  
**Target-Bearing Material Used in Phoneme-Monitoring Experiment**  
**Without /ts/ Items**

	Word	Minimal Pseudoword	Maximal Pseudoword
Target phoneme <i>t</i> -final	Bajonett	Pajonett	Fajonett
	Bürokrat	Dürokrat	Sürokrat
	Defizit	Befizit	Lefizit
	Demokrat	Bemokrat	Semokrat
	Diplomat	Biplomat	Niplomat
	Dokument	Bokument	Nokument
	Dynamit	Bynamit	Fynamit
	Favorit	Vavorit	Lavorit
	Fregatte	Wregatte	Pregatte
	Fundament	Wundament	Kundament
	Gazette	Dazette	Lazette
	Kabarett	Gabarett	Mabarett
	Kabinett	Gabinett	Dabinett
	Kajüte	Gajüte	Najüte
	Kamerad	Gamerad	Bamerad
	Klamotte	Glamotte	Slamotte
	Klarinette	Glarinette	Slarinette
	Kokotte	Gokotte	Dokotte
	Kollekte	Gollekte	Dollekte
	Konsonant	Gonsonant	Sonsonant
	Kosmonaut	Gosmonaut	Sosmonaut
	Krawatte	Grawatte	Frawatte
	Labyrinth	Rabyrinth	Kabyrinth
	Lazarett	Razarett	Fazarett
	Leopard	Reopard	Seopard
	Manifest	Nanifest	Panifest
	Manuskript	Nanuskript	Banuskript
	Marionette	Narionette	Karionette
	Marotte	Narotte	Larotte
	Medikament	Nedikament	Kedikament
	Minute	Ninute	Finute
	Monument	Nonument	Fonument
	Parlament	Barlament	Garlament
	Patient	Batient	Natient
	Pergament	Bergament	Lergament
	Rakete	Lakete	Wakete
	Referat	Leferat	Veferat
	Requisit	Lequisit	Mequisit
	Sakrament	Fakrament	Gakrament
	Serviette	Ferviette	Perviette
	Silhouette	Filhouette	Kilhouette
	Surrogat	Furrogat	Burrogat
Vagabund	Fagabund	Pagabund	
Visite	Sisite	Kisite	
Zigarette	Schigarette	Pigarette	
Zölibat	Schölibat	Pölibat	

*(Continued on next page)*

## APPENDIX B (Continued)

	Word	Minimal Pseudoword	Maximal Pseudoword
Target phoneme <i>k</i> -final	Bibliothek	Dibliothek	Ribliothek
	Botanik	Dotanik	Notanik
	Dynamik	Bynamik	Fynamik
	Genetik	Tenetik	Lenetik
	Grammatik	Trammatik	Frammatik
	Gymnastik	Tymnastik	Lymnastik
	Limerick	Rimerick	Fimerick
	Logistik	Rogistik	Pogistik
	Mechanik	Nechanik	Fechanik
	Methodik	Nethodik	Tethodik
	Mosaik	Nosaik	Fosaik
	Pädagogik	Tädagogik	Chädagogik
	Pazifik	Bazifik	Razifik
	Politik	Bolitik	Volitik
	Pragmatik	Bragmatik	Fragmatik
	Privileg	Trivileg	Grivileg
	Problematik	Troblematik	Groblematik
	Republik	Lepublik	Mepublik
	Romantik	Lomantik	Momantik
	Salmiak	Zalmiak	Talmiak
	Synthetik	Zynthetik	Tynthetik
	Systematik	Zystematik	Tystematik
	Tomahawk	Domahawk	Lomahawk
	Touristik	Douristik	Louristik
Target phoneme <i>s</i> -final	Balance	Palance	Kalance
	Bronchitis	Pronchitis	Fronchitis
	Delikatesse	Pelikatesse	Felikatesse
	Devisse	Pevisse	Fevisse
	Finesse	Vinesse	Ginesse
	Gedächtnis	Kedächtnis	Medächtnis
	Karosse	Tarosse	Rarosse
	Kolchose	Golchose	Scholchose
	Kulisse	Gulisse	Schulisse
	Matrose	Natrose	Batrose
	Mayonaise	Nayonaise	Vayonaise
	Mimose	Nimose	Schimose
	Noblesse	Moblesse	Boblesse
	Paradies	Baradies	Garadies
	Radius	Ladius	Madius
	Kompresse	Tompresse	Fompresse
	Thrombose	Prombrose	Brombrose
Virtuose	Firtuose	Kirtuose	

(Manuscript received November 26, 2001;  
revision accepted for publication November 10, 2003.)