

Gender and lexical access in Italian

ELIZABETH BATES

University of California, San Diego, La Jolla, California

ANTONELLA DEVESCOVI

Università di Roma "La Sapienza," Rome, Italy

LUIGI PIZZAMIGLIO

*Università di Roma "La Sapienza," Rome, Italy
and Clinica Neurologica e Centro di Ricerca Santa Lucia, Rome, Italy*

SIMONA D'AMICO

CNR, Istituto di Psicologia, Rome, Italy

and

ARTURO HERNANDEZ

University of California, San Diego, La Jolla, California

Two new procedures were employed to investigate the effects of semantic and grammatical gender on lexical access in Italian and to investigate the interaction of gender with other factors that are known to influence lexical access in other languages. The gender-monitoring task requires a conscious decision about the gender of each noun, whereas the word repetition task does not require explicit attention to gender. In both tasks, single words are presented out of context, under speeded conditions. Both procedures proved to be sensitive indices of word recognition, with reaction times that are closely tied to the point at which words can be uniquely identified (although some processing before and after the uniqueness point was seen). In both tasks, reaction times were strongly affected by phonological factors (e.g., length, number of syllables, and presence of frication on the initial consonant). Phonological transparency of gender marking had a reliable effect on gender monitoring but had no effect on word repetition, suggesting that explicit attention to gender may be a factor affecting utilization of this phonological cue. Semantic factors (including semantic gender) had no effect on performance. Frequency and age of acquisition had very small effects when other factors were controlled. Implications for current models of lexical access are discussed, with special reference to the role of gender.

Gender marking is a pervasive phenomenon in many of the world's languages, but its role in lexical and grammatical processing is still poorly understood. In most of the languages that incorporate some form of productive gender marking, the relationship between grammatical and semantic gender is indirect at best. To offer one widely cited example (e.g., Maratsos, 1981), who can explain why, in the German language, the term for *the flute* (*die Flöte*) is

feminine, whereas the term for *the little girl* (*das Mädchen*) is neuter? Or why the term for *the ocean* is neuter in German (*das Meer*), feminine in French (*la mer*), and masculine in Italian (*il mare*)?¹ If gender serves no systematic semantic function, why do these languages continue to mark gender on nouns and most of their modifiers? This investment is not without a price, since gender marking and gender agreement require a substantial degree of coordination between grammar and lexical selection in real-time processing, and they present a challenge to first- and second-language learning (for discussion of these points, see Bates & Devescovi, 1989; Bates & MacWhinney, 1989; Fodor, 1959; MacWhinney, 1978; MacWhinney, Leinbach, Taraban, & McDonald, 1989; Mills, 1986; Mulford, 1983; Zubin & Köpcke, 1981). Why, then, should languages persist in the use of a costly linguistic device that serves no obvious communicative function?

Although the issue is still controversial, one possible explanation for the pervasiveness and persistence of gender revolves around its role in lexical access. In contrast with other aspects of inflectional morphology (i.e., case, num-

This research was supported by NIDCD Grant PHS DC00216-09 to E.B., and by a grant from the Clinica e Centro di Ricerca Santa Lucia to L.P. During manuscript preparation, E.B. was a visiting scholar at the Istituto di Psicologia, Consiglio Nazionale di Ricerca, in laboratories directed by Virginia Volterra at the Istituto Statale Sordo-Muti, Via Nomentana 56, Rome, Italy. We are grateful to Bob Buffington, Larry Juarez, and Meiti Opie for assistance in stimulus preparation and data collection, to Angela Friederici and Michèle Kail for advice and criticism in the design phases of this study, and to Uli Frauenfelder and William Marslen-Wilson for their advice and criticism during analysis and manuscript preparation. Any errors of design or interpretation that remain are, of course, our own. Please address correspondence to E. Bates, Center for Research in Language 0526, University of California, San Diego, La Jolla, CA 92093-0526.

ber, person, tense, and aspect), gender is an inherent property of nouns that can be retrieved at the moment of lexical access for words presented out of context.² Speakers of a gender-marked language often mark adjectives and pronouns to agree with a noun that is never encoded explicitly in the discourse. For example, a mother may point out a dog to her child and comment, "Guarda che bello!" ("Look how cute!" [masculine]), or point to a car and say, "Hai visto quella?" ("Have you seen that?" [feminine]). Under these and other circumstances, gender can serve as an important cue to the identity of a lexical item and its referent. Bates and MacWhinney (1989) have suggested that gender marking serves a function roughly analogous to the numbers that identify football players on a field in a hectic game (i.e., as a way of uniquely marking and tracking several different referents in a complex discourse).³ Some evidence in favor of this view comes from Kilborn (1987), who showed that German listeners have an advantage over English subjects in a word-monitoring task in which words must be identified in syntactically well-formed but semantically anomalous prose (e.g., "Colorless green IDEAS sleep furiously"). In the same vein, Grosjean, Dommergues, Cornu, Guillelmon, and Besson (1994) have shown that gender marking on the article serves as a powerful cue to recognition of a subsequent noun.

There is still relatively little empirical work on the role of gender in language processing, although a few laboratories have started to investigate the issue (e.g., Brooks, Braine, Catalano, Brody, & Sudhalter, 1993; Burani, 1992; Cassidy & Kelly, 1991; Colé & Ségui, 1994; Deutsch & Wijnen, 1985; Devescovi, D'Amico, Smith, Mimica, & Bates, 1993; Friederici & Schriefers, 1993; Grosjean et al., 1994; Jarema & Friederici, 1994; Kail, 1989; Orsolini, 1993; Radeau, Mousty, & Bertelson, 1989; see also unpublished studies reported in Brown, Senft, & Wheeldon, 1993). The purpose of the present study was to explore the factors that influence gender processing in Italian when words are presented out of context and to explore the influence of gender itself on lexical access in tasks that do or do not require conscious reflection on this particular lexical/grammatical property. We view this study of isolated words as a necessary first step prior to a series of experiments examining the effects of gender on lexical access in Italian, in noun phrases and/or in a sentence context, in normal adults and in patients suffering from focal brain injury.

Two separate tasks were employed for this purpose: repetition of spoken words (hereafter, *word repetition*) and classification of those words according to gender (hereafter, *gender monitoring*). Single-word repetition is a relatively new technique for the study of lexical access, but its utility has been established in several studies (Connine, Mullennix, Shernoff, & Yelen, 1990; Luce, Pisoni, & Goldinger, 1990; Marslen-Wilson, 1990; Radeau & Morais, 1990; Whalen, 1991), including new evidence showing that repetition of auditorally presented words is a sensitive index of semantic priming (Herron, 1994; Liu & Bates, 1993; Slowiaczek, 1994). We chose this method

because it requires very little (if any) metalinguistic reflection; as such, it provides a useful complement to gender monitoring, a task that does require the subject to make a deliberate, conscious decision about the grammatical gender of each target word.

Even fewer studies have looked at gender monitoring. Radeau et al. (1989) have used this procedure to study auditory word recognition in French, and a related procedure called *gender verification* has been used successfully in Dutch (De Ruiter, 1992; Deutsch & Wijnen, 1985). In the present study, the gender-monitoring task provided a means of assessing the interactions between gender (both semantic and grammatical) and other factors that are known to influence lexical access, in a situation that required attention to the gender dimension. Through the joint use of these two relatively new techniques, we could compare the effects of gender and its interaction with other variables with and without explicit awareness of gender per se.

The gender system in Italian is relatively straightforward, compared with many other gender-marked languages, but it has a number of interesting properties that will be explored here. There are only two genders, masculine and feminine (in contrast, for example, with the three genders of German and Russian, or the six genders of Swahili; Grosjean et al., 1994). Gender is an inherent, context-independent property of every Italian noun, and gender agreement must be marked on almost all modifiers (i.e., articles, determiners, and adjectives—numerals are not marked for gender), on all coreferential pronouns (including full pronouns and clitics), and on the past participle of the verb (for verbs using the auxiliary *essere*, the past participle must agree in gender and number with the subject; for verbs using the auxiliary *avere*, the past participle must agree in gender and number with the object, but only when a preverbal clitic pronoun has been used). There are no unmarked or zero noun forms in Italian. Except for a small number of foreign loan words (e.g., *bar*), all Italian nouns end in a vowel, and gender and number are marked together on that final vowel. For the great majority of nouns (and for most agreeing adjectives), masculine forms end in *o* in the singular and *i* in the plural, feminine forms end in *a* in the singular and *e* in the plural. We will refer to these as *phonologically transparent* items. For a minority of both masculine and feminine word types (and some agreeing adjectives), the final vowel is *e* in the singular and *i* in the plural. Because gender cannot be recovered from surface form alone on words within this class, we will refer to them as *phonologically opaque*. In addition to the two largest word classes (transparent and opaque), a very small number of word types in Italian carry contradictory marking. These include idiosyncratic words such as *la mano* (a feminine term meaning *the hand*, with masculine marking on the noun but feminine agreement on all modifiers), and a small class of words derived from Greek, such as *drama* or *telegramma* (masculine words for *drama* and *telegram*, with feminine marking on the noun but masculine agreement on all modifiers). Both transparent and opaque word types were used

in the present study to investigate whether this dimension affects performance by native speakers in either of our tasks. The very small class of aberrant or contradictory forms will not be investigated here.

In Italian (as in all of the gender-marked languages in the world), the relationship between semantic and grammatical gender is indirect at best. For most words, the relationship is completely arbitrary. This includes many words for animals that have semantic gender (e.g., the citation form for *the tiger* is *la tigre*, a feminine term; the citation form for *the lion* is *il leone*, a masculine term). However, there are very few cases in which semantic and grammatical gender are in direct conflict (e.g., there is no equivalent to the German *das Mädchen*, neuter gender for an inherently feminine term meaning *the little girl*). When semantic gender really matters (e.g., *ragazza* vs. *ragazzo*—*girl* vs. *boy*), it usually agrees with grammatical gender. In the present study, we included words with and without inherent semantic gender, but we included no cases in which semantic and grammatical gender conflicted.

The following issues pertaining to gender were addressed in the present study:

(1) If semantic gender behaves like other semantic variables that have been studied in the literature on lexical access (e.g., concreteness, imageability), then we might expect faster reaction times for words that carry semantic gender. Alternatively, in view of the high degree of arbitrariness that holds between semantic and grammatical gender in Italian, we may find that semantic gender has no effect at all on lexical access in this language. There is a third possibility midway between these two extremes: Semantic gender may facilitate response on the gender-monitoring task (where explicit awareness of gender is required) but may have no effect whatsoever on word repetition.

(2) Levelt (1989) has proposed that lexical retrieval involves two distinct and potentially dissociable stages: retrieval of the lemma (the lexical entry with its semantic features and possible syntactic roles) and retrieval of the word form (which provides phonological and morphological specification in a given context). Although Levelt's model is an account of stages in word production, analogous claims have been made about the separation of word form identification and lemma retrieval in word comprehension (for reviews, see Frauenfelder & Tyler, 1987). If grammatical gender is retrieved at the level of the lemma, then there should be no effect of phonological transparency on reaction times (i.e., it should not matter whether the word ends in a transparent *o* or *a* or with a phonologically ambiguous *e*). If we do see effects of phonological transparency (with slower reaction times on opaque items), then we may conclude that there are interactions between these two stages of lexical access. A third alternative is possible: Phonological transparency may have no effect on word repetition, although it does influence the time required to perform on the gender-monitoring task. If this occurs, we may infer that the effects of phonological transparency accrue at a later point in processing, after the lemma has been retrieved.

Although our primary goal in the present study was to investigate the role of gender in lexical access, the fact that both these techniques are so new means that we also needed to determine whether, or to what extent, these methods are sensitive to factors that are known to influence lexical access when other paradigms are used (i.e., such factors as frequency, word duration, number of syllables, age of acquisition, and abstractness). Hence, we will provide some new information about old variables in the study of word recognition, in a language that differs markedly from English along a number of interesting phonological and grammatical dimensions. This information is useful in its own right, but it is also crucial to our evaluation of the role of grammatical and semantic gender in lexical access in Italian. In the results presented below, we will begin by looking at the contribution of such factors as length and frequency to performance on both tasks and end with an assessment of effects associated with semantic and grammatical gender.

METHOD

Subjects

Twenty adult native speakers of Italian participated in this study. All were residents of Rome, attending a major university. All subjects received both the word repetition test and the gender-monitoring test, in a counterbalanced order. The analyses reported below were conducted over items, using mean scores over subjects as dependent variables (for percent correct gender choice, reaction times [RTs] in gender choice, and RTs in word repetition).

Materials

For this exploratory study, we began with a list of 482 Italian words, all common nouns, taken from two sources: De Mauro, Mancini, Vedovelli, and Voghera (1993) and Caselli and Casadio (1992). The former is a compilation of frequencies taken from spoken conversations in several different regions of Italy, and it contains more than 500,000 items. These norms provide three different indices of frequency for each item: raw frequency of the root word (summed across morphological variants), frequency of each morphological variant that appeared in the corpus, and adjusted frequency (a weighted average across the different texts within the corpus). All three frequency measures were considered in the present study. The latter is a parent report questionnaire on language development from 8 to 30 months, containing 690 words that are among the first words spoken by young Italian children. Words were selected from the two lists according to the following criteria: (1) Only singular forms were selected for this particular study. (2) The sample was restricted to words beginning with a consonant. Vowel-initial items were excluded because of the problems we foresaw in splicing such words into an article or adjective context in future experiments (e.g., *l'elefante*, as opposed to *la ragazza* or *il ragazzo*). (3) Foreign loan words, acronyms, slang terms, and proper names were excluded, together with highly abstract, technical, or context-specific terms. (4) We concentrated primarily on two- and three-syllable words with a spoken frequency greater than 3 in the De Mauro norms (except for those words from the MacArthur infant scales that are simply not used in adult conversations; see below). However, we did include a subset of 38 four- and five-syllable words to keep subjects from developing rhythmically based strategies or expectations and to take a preliminary look at the effect of syllable structure on word repetition and gender monitoring (see basic statistics on the data set, below). Hence, words may vary in length from two to five syllables. (5) We included words with and without semantic gender. For the purposes of this experiment, an item was classified as + Sexed if it referred to

an individual capable of sexual reproduction at some point in its life, whether or not the sexual status of that individual is obvious to the speaker/listener (e.g., *butterfly* was classified as +Sexed). (6) We also made a point of including 56 pairs of words (112 words total) that differ only by the final vowel, which is, of course, also the phoneme that guarantees gender identification in Italian. This subset of items will help us to examine the effect of nonphonological factors (especially frequency) when the phonological difference between words is tightly controlled. (7) We were interested in examining the impact of phonological cues to grammatical gender, in particular the contrast between singular masculine words that end in *o* and singular feminine words that end in *a*, compared with masculine and feminine words that end in a phonologically ambiguous *e* in the singular form. We will refer to this dimension (for lack of a better term) as *transparency*. (8) The De Mauro et al. norms indicate that there are far more masculine nouns overall in the Italian language. However, because gender is a major focus of the present study, we obtained a roughly equivalent sample of masculine and feminine words. This means, of course, that we have undersampled masculine words and/or oversampled feminine words, a point that must be taken into consideration in our analysis of apparent confounds.

Although the resulting word list is quite broad (see Table 1 for details), we cannot claim that this list constitutes a balanced or representative sample of nouns in the Italian language. Indeed, one purpose of this study was to uncover those factors that are confounded in the language (and, as we shall see, there are many) and to determine which factors have enough impact on word and gender recognition that they must be counterbalanced carefully in future studies of lexical and morphological processing. In that respect, we have been quite successful—including some surprising findings that have not been reported in other studies of lexical access in Italian, with implications for other languages, including English.

All words were tape-recorded by a female native speaker of Italian, in a neutral intonation with a falling tone on the final syllable. The words were digitized using the Macintosh SoundEdit system and placed in a sound file within the PsyScope experiment prepara-

tion shell developed by MacWhinney, Cohen, and colleagues at Carnegie-Mellon University (Cohen, MacWhinney, Flatt, & Provost, 1993). PsyScope can access and present words in a unique random order for each subject, an important advantage for the present study with a relatively small sample of subjects and a large list of words. For each item, the RT clock began at the onset of the stimulus word and ended with the subject's response (measured from the onset of the vocal response in word repetition and from the buttonpress in gender monitoring). The subjects had to respond within a 300- to 1,500-msec response window measured from the end of the word stimulus (randomly assigned to trials by PsyScope). If the subject failed to respond within that window, a "nonresponse" was registered automatically for that trial. If the subject responded before the onset of the stimulus word (i.e., if a negative RT was registered), that trial was also treated as a nonresponse. The next trial was always presented 500 msec following the subject's response (or the end of the response window on nonresponse trials). Although the 500-msec postresponse component was fixed, the intertrial interval appeared to vary randomly from the subject's point of view, from 800 to 2,000 msec. This was done to maximize attention and discourage the development of expectations or rhythmic strategies of word repetition. Pilot testing had confirmed that this procedure worked better than a fixed and predictable interval and that these particular time parameters put the subjects under a time pressure (and elicited rapid response) but did not result in overlap between a vocal response and the next stimulus.

Fourteen words were eliminated due to experimental error, leaving a total of 468 items. The total duration of each stimulus word in milliseconds was measured using the Macintosh SoundEdit system. We also calculated the uniqueness point for each word, defined as the end of the phoneme that uniquely distinguishes that word from all other alternatives in the Palazzi dictionary for Italian (Palazzi, 1973; for a discussion of uniqueness points and their role in word recognition, see Grosjean, 1980). For example, words such as *banca* and *banco* cannot be uniquely identified until the final vowel. However, the word *frigorifero* (refrigerator) becomes distinct from all

Table 1
Composition of the Word List (N = 468)

Grammatical gender	feminine	226	48.3%
	masculine	242	51.7%
Phonological transparency of gender ending	transparent (-o or -a)	373	79.9%
	opaque (-e)	95	20.1%
Number of syllables	2	293	62.6%
	3	137	29.3%
	4	30	6.4%
	5	8	1.7%
Frication on initial consonant	no initial fricative	308	65.8%
	initial fricative	160	34.2%
Source (age of acquisition)	Adult norms	227	48.5%
	Infant list only	241	51.5%
Semantic gender	without	375	80.1%
	with	93	19.9%
Abstractness	concrete	407	87.0%
	abstract	61	13.0%
Humanness	nonhuman	420	89.7%
	human	48	10.3%
Total length (in milliseconds)	<i>M</i> = 897	<i>SD</i> = 148	range = 569–1401
Length of lead-in segment (word onset to uniqueness point)	<i>M</i> = 771	<i>SD</i> = 163	range = 336–1188
Length of end segment (uniqueness point to word offset)	<i>M</i> = 126	<i>SD</i> = 172	range = 0–819
Word frequency	<i>M</i> = 35.89	<i>SD</i> = 73.4	range = 0–599
Form frequency	<i>M</i> = 25.82	<i>SD</i> = 55.6	range = 0–471
Adjusted frequency	<i>M</i> = 23.96	<i>SD</i> = 60.34	range = 0–551

other words in the Palazzi dictionary by the end of the first two syllables (i.e., *frigo*-).

To locate the uniqueness point physically within each word stimulus, the SoundEdit facility was used to measure manually from the beginning of the word to the end of the disambiguating phoneme (according to the dictionary definition). This is not a simple task, due to the existence of word-internal coarticulation. A combination of visual and acoustic information was used to make this judgment. First, the rater inspected the waveform visually to locate a probable phonemic transition from the end of the disambiguating phoneme (according to the dictionary) and the next phoneme. Second, the cursor was set at the supposed transition point, establishing a lead-in segment (defined as the length in milliseconds from word onset to this uniqueness point) and an end segment (defined arithmetically as total word duration minus length of the lead-in segment). The lead-in segment and the end segment were played back acoustically to determine whether an optimal transition had been found. This process was reiterated until an optimal match was reached between the visual and the auditory judgment. It must be acknowledged that the concept of a uniqueness point assumes a clean transition that is often absent in the overlapping segments of fluent and coarticulated speech.

Since pluralization is always marked on the final vowel in Italian, the uniqueness point for singular versus plural forms of a word is necessarily word-final. Hence, the uniqueness point in the present study refers to the point at which the subjects could uniquely identify the word root (assuming that the word will be singular, which was always the case in this study). It is also worth noting that stress plays an important role in word identification in Italian. For example, the root noun *bagno* (bath) can be uniquely identified by the end of the middle consonant (as in *bagn*-) because it takes stress on the first syllable. That is, there is no other word in the Italian dictionary that starts with initial stress on *bagn*-. But if stress is not taken into account, there are a number of derived and diminutive forms that could follow after *bagn*- (e.g., *bagnino*, which means bath attendant or lifeguard; *bagnarola*, a word for bathtub; *bagnasciuga*, which means waterline or shoreline. All of these items are stressed on the penultimate syllable). Ignoring stress, the uniqueness point for *bagno* would have to be the final vowel. In this study, uniqueness points were always calculated with stress taken into account, which provides an earlier estimate in many cases.

A total of 14 predictor variables were derived for each word: grammatical gender (feminine = 1; masculine = 2); three dichotomous semantic measures, including semantic gender (-Sexed = 0; +Sexed = 1), humanness (-Human = 0; +Human = 1), and abstractness (because there are no systematic norms for concreteness or imageability in Italian, this was operationalized as a simple dichotomy between words that have concrete referents that can be seen or felt directly and words that do not [concrete = 0, abstract = 1]); phonological transparency of gender marking (transparent words ending in *o/a* = 1; opaque words ending in *e* = 2); four measures of length, including number of syllables, duration in milliseconds, length in milliseconds prior to the uniqueness point (referred to below as the lead-in), and length in milliseconds after the uniqueness point (referred to below as the end segment); presence/absence of a fricative or affricate in the initial consonant (0 = no fricative or affricate; 1 = fricative or affricate); age of acquisition, where 1 indicates that the item was taken from the De Mauro adult language norms, and 2 indicates that the item was taken from the list of first words acquired by children (and was not available in the De Mauro norms); three measures of frequency (root word frequency, form frequency, and adjusted frequency).

Table 1 provides information on the characteristics of the final 468 word samples along these dimensions.

Procedure

The subjects were tested individually in a quiet room. For the word repetition task, they were told that they would hear a series of

words. They were asked to repeat each word as quickly as possible without making a mistake and to speak clearly into the microphone. RTs for word repetition were collected by a voice key contained within the Carnegie-Mellon "button box," an auxiliary to the Macintosh workstation that contains an independent timing crystal with 1-msec accuracy. RTs were measured from the onset of the target word to the onset of the subject's repetition of that word and were fed directly into the PsyScope file. For the gender-monitoring task, the subjects were asked to place the index finger of the preferred hand on a spot between two plastic buttons. For each item, they were asked to indicate the gender of that word by pressing one of the buttons for *feminine* and the other for *masculine* (indicated by a symbol above each button). Assignment of genders to button position was counterbalanced over subjects to minimize possible effects of a right-going bias (i.e., half the subjects had to indicate *feminine* on the left and *masculine* on the right, whereas the other half had to indicate *masculine* on the left and *feminine* on the right). The subjects were asked to return the index finger to the central position between the two buttons after each response. Order of presentation for the two tasks was counterbalanced, and a different unique random order was used within each task for each subject.

Three dependent variables were derived from performance in these two measures: gender accuracy (percent correct classification), gender-monitoring reaction time (GMRT; in milliseconds), and word repetition reaction time (WRRT; in milliseconds, calculated from the onset of the stimulus to the onset of the subject's repetition). Means were calculated over subjects and entered into analyses over items.

RESULTS AND DISCUSSION

Overall Performance on the Two Tasks

Not surprisingly, our subjects were very good at the gender-monitoring task, with an average of 95% accuracy over items (range = 67%–100%, *SD* = 6%). However, errors did occur, and they were not randomly distributed over the data set (see below). Errors were rare on the word repetition task (under 2%) and will not be discussed further.

GMRTs averaged 1,134 msec (range = 866–1,457 msec, *SD* = 108 msec), while reaction times for word repetition averaged 957 msec (range = 451–1,241 msec, *SD* = 115 msec). Both these averages may appear to be rather slow, relative to results from studies of lexical access in English. However, it is important to remember that Italian words are considerably longer, on average, than English words (e.g., a mean of 2.6 syllables in the present study, compared with the monosyllabic words that predominate in most English-language studies of lexical access; for some exceptions, see Goodman & Huttenlocher, 1988, and Taft & Hambly, 1986).

In fact, when RTs are calculated from the end of the word, the average RT was 238 msec for gender monitoring in the present study (range = -131–534 msec, *SD* = 101 msec). This means that Italian subjects are making their gender classifications less than 250 msec after the final vowel, which carries definitive gender information. When RTs are calculated from the uniqueness point, the average RT for gender monitoring was 363 msec (range = 6–904 msec, *SD* = 149 msec). Assuming that a certain amount of time is required to make a binary decision and to generate and execute a motor response, we may take this finding as evidence for the idea that gender information is accessed and classified close to the point at which the word itself is identified.

Since the word repetition task does not require a metalinguistic decision, it may provide an even more faithful view of the moment at which the lexical item is retrieved. In fact, when word repetition times are calculated from the end of the word, the mean RT in this experiment was only 61 msec (range = -528-311 msec, $SD = 118$ msec). Repetitions were initiated before the end of the word in 131 of 468 cases (28% of all items).

When RTs are calculated from the uniqueness point, the mean for word repetition was 186 msec (range = -405-661 msec, $SD = 162$ msec). At first glance, this appears to provide compelling evidence in favor of the proposal that words are identified at the uniqueness point when they are presented out of context. However, it is also true that repetitions were initiated *before* the uniqueness point on 50 items (10.7% of the stimuli). Apparently, Italian speakers/listeners can use other sources of information to identify single-word targets out of context. What might these sources be? A closer examination of the data showed that all of these exceptional items were words that cannot be distinguished from others in the Palazzi dictionary until the final vowel. It seems reasonable to assume that a great deal of information about the shape of the word is already available on such items, including aspects of intonation and coarticulatory cues that might help to identify the final vowel. We will return to this point later on in our discussion of correlations between predictor and outcome variables.

Figure 1 compares the RTs that were obtained for gender monitoring and word repetition, measured from the beginning of the word, the end of the word, and the uniqueness point. (For each of the three RT measures, the difference between GMRT and WRRT is a mathematical constant, so the parallel lines in Figure 1 are no surprise.) We quantified the difference between tasks by subtracting the gender-monitoring and word repetition scores for each word, yielding a mean difference score of 177 msec (range = -30-639 msec, $SD = 96$ msec). There were only six items (1.3% of all cases) on which WRRTs were longer than the mean GMRT. Interestingly, all six are words that begin with a fricative consonant (see effects of frication, below). But for the vast majority of words, it seems that gender classification requires some extra steps beyond the recognition processes that are required for repetition. There are several possible explanations for this robust finding. On the one hand, we may speculate that gender is indeed retrieved at the moment of lexical access, but the classification decision requires some kind of postaccess check—an extra step may be sensitive to factors that have little or no effect on word repetition. On the other hand, the responses required for these two tasks are different in a number of other respects (e.g., the gender-monitoring task requires a binary decision followed by a manual response, while the word repetition task requires no metalinguistic decision of any kind, with a purely vocal response). There is no way that we can decide among these possibilities in the present study, although it will be useful to determine whether gender monitoring and word repetition are sensitive to different predictor variables (especially phonological transparency and semantic gender).

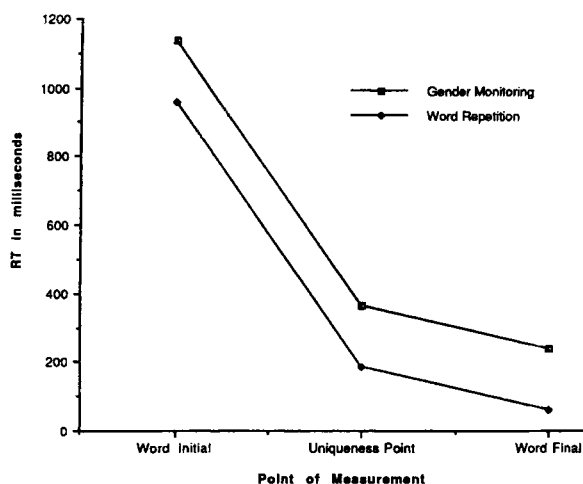


Figure 1. Reaction times for gender-monitoring and word repetition measures from word onset, the uniqueness point, and word-final position.

Relationships Among the Predictor Variables

Pearson product-moment correlations were calculated across all the predictor variables. It was clear from these analyses (summarized in Table 2) that many of the predictors are confounded. Some of these are natural confounds in every language. For example, there are a number of significant relationships between frequency and length, in accord with Zipf's law (which states that the most frequent words in any language tend to be relatively short). There are also confounds that appear to be unique to Italian. For example, words that end in *e* tend to be longer than words that end in *o* or *a*, and words with semantic gender are longer on average than words without. Confounds like this greatly complicate all the analyses that follow, but they are potentially important for the study of lexical access in Italian, and we would not have discovered them without undertaking an exploratory study of this kind. We will deal with these confounds in the analyses that follow by relying on a combination of regression analysis (examining the independent contribution of each predictor when all the others are controlled) and analysis of variance on selected items that are matched for length.

Correlations of the various predictor variables with gender accuracy, GMRT and WRRT are reported in Table 3. We will begin a more detailed examination of these predictors by considering phonological effects on the three major dependent variables.

Frication

We had been alerted by colleagues who work within the auditory modality to look carefully at word-initial consonants, with particular emphasis on fricatives and affricates. It is difficult for the computerized voice key to detect the onset of an initial fricative or affricate, a factor that might influence WRRTs. It may also be difficult for human perceivers to detect initial fricatives and affricates, a factor that could influence both GMRTs and WRRTs.

Table 2
Correlations Among the Predictor Variables

	Gender	Transparency	Sexed	Abstract	Human	Total Length	Lead-in Length	End Length	No. of Syllables	Frication	Word Frequency	Form Frequency	Adjusted Frequency
Gender	—												
Transparency	+ .06	—											
Sexed	+ .01	-.01	—										
Abstract	-.05	+ .03	-.18†	—									
Human	+ .04	-.01	+ .68†	-.11†	—								
Total length	+ .07	+ .13†	+ .13†	-.14†	+ .17†	—							
Lead-in length	+ .05	-.07	+ .22†	-.02	+ .29†	+ .39†	—						
End length	+ .02	+ .17†	-.09*	-.10*	-.13†	+ .49†	-.61†	—					
No. of syllables	+ .07	+ .15†	+ .14†	-.13†	+ .18†	+ .73†	+ .12†	+ .52†	—				
Frication	-.01	-.04	-.01	-.07	+ .04	+ .23†	+ .17†	+ .04	-.05	—			
Word frequency	+ .02	+ .02	+ .02	+ .29†	+ .12†	-.06	+ .02	-.08*	-.07	+ .01	—		
Form frequency	-.00	+ .00	-.04	+ .31†	+ .04	-.10*	-.01	-.08*	-.10*	.00	+ .94†	—	
Adjusted frequency	+ .02	+ .00	-.01	+ .30†	+ .09*	-.07	-.01	-.07	-.08*	+ .01	+ .98†	+ .93†	—
Source (age of acquisition)	-.06	-.14†	+ .08*	-.36†	-.04	+ .21†	+ .00	+ .18†	+ .25†	-.06	-.10*	-.13†	-.10*

* $p < .05$. † $p < .01$. ‡ $p < .001$.

Table 3
Correlations Between Predictor and Outcome Variables

	% Correct	GMRT	WRRT
Frication on initial consonant	-.02	+.32‡	+.54‡
Total length (in milliseconds)	-.05	+.73‡	+.63‡
Lead-in length	-.07	+.46‡	+.36‡
End length	+.03	+.20‡	+.20‡
No. of syllables	-.03	+.42‡	+.27‡
Word frequency	+.04	-.09*	-.07
Form frequency	+.01	-.12‡	-.09‡
Adjusted frequency	+.05	-.10*	-.07
Source (age of acquisition)	+.07	-.08*	+.08*
Semantic gender (+Sexed)	+.02	+.15‡	+.05
Abstract	+.01	-.03	-.09*
Human	+.08*	+.21‡	+.12‡
Grammatical gender	+.16‡	+.05	-.02
Transparency of gender marking	-.24‡	+.21‡	+.06
% Correct	—		
GMRT	-.21‡	—	
WRRT	-.06	+.63‡	—

* $p < .05$. † $p < .01$. ‡ $p < .001$.

All items were classified dichotomously for presence or absence of a fricative or affricate on the initial consonant (0 = no fricative or affricate; 1 = any fricative or affricate in initial position, including single consonants followed by a vowel and/or consonant clusters).⁴ It is clear from the correlations in Table 3 that this variable bears no relationship to gender-monitoring accuracy ($r = -.02$, n.s.), but does influence both GMRT ($r = +.32$, $p < .001$) and WRRT ($r = +.54$, $p < .001$) when they are measured from the beginning of the word. Taken together, these results suggest that frication is involved in two separate effects: a true perceptual effect (which slows word-initial RTs for both tasks), and an artifact of measurement with a computer voice key (which only influences WRRT in this study).

Before we conclude that frication has an independent effect on RT, we must consider the confounds that are evident in Table 2. In particular, the frication scale is significantly correlated with total length in milliseconds measured from the word onset point ($r = +.23$, $p < .001$) and with length of the lead-in segment (i.e., length in milliseconds measured from the beginning of the word to the uniqueness point; $r = +.17$, $p < .001$). By contrast, there is no relationship between frication and length of the end segment after the uniqueness point ($r = +.04$, n.s.), which means that the frication effect is restricted to the early portions of the word. Aside from the fact that fricative consonants themselves tend to be relatively long, it may also be the case that speakers unconsciously compensate for the perceptual difficulty associated with frication by lengthening the next vowel. Obviously, this confound could have an effect on the RT correlations described above.

To control for this and other potential confounds, we carried out a series of regression analyses on the relationship between frication and three dependent variables (accuracy, GMRT, and WRRT). As we have noted, there were 14 predictor variables available in this study (see Tables 1 and 2). However, total length in milliseconds is redundant with the combination of lead-in length and length of the end segment. Therefore, for this regression analysis and all of those that follow, we used a total of 13 predictors (excluding total length in milliseconds). On each analysis, frication was entered into the equation on the last step. All of the remaining predictor variables (a total of 12) were entered into the equation on the first step (lead-in length, end length, number of syllables, all three frequency measures, age of acquisition, the three semantic measures, gender, and transparency), permitting us to assess the unique contribution made by the frication measure. For these analyses, we also concentrated only on word-initial RT scores, since the use of lead-in length and end length as predictor variables effectively controls for differences among the three different measurement points. We will follow a similar stepwise procedure for all the regression analyses discussed in this paper (i.e., 12 variables controlled on the first step, with the predictor in question entered on the last step to assess its unique contribution). Results of all these regression analyses are summarized in Table 4.

Results suggest that frication does have an independent effect on RTs when other factors are controlled. Frication makes a large contribution to word repetition latencies when it is entered on the last step, increasing the total variance accounted for by an additional 11.5% ($p < .001$). Examination of the partial correlations shows (as expected) that this is a positive contribution (i.e., RTs are longer for items that begin with a fricative or affricate). The frication variable makes no independent contribution to gender accuracy but does make a small but reliable contribution to GMRT (1.0%; $p < .001$), again in the positive direction (i.e., increasing RTs). These results suggest that words with an initial fricative consonant should be avoided or carefully controlled in studies of auditory lexical access in Italian, particularly when those studies require a vocal response.

Length in Milliseconds

Table 3 shows that the various measures of length were strongly correlated with performance on both of our tasks. Indeed, these were the largest effects in the study, ranging as high as $r = +.73$ ($p < .001$) for GMRT, and $r = +.63$ ($p < .001$) for WRRT. In addition, the results in Table 3 suggest that lead-in length and length of the end segment may have somewhat different effects on performance (with larger contributions from lead-in length). However, the correlations in Table 2 also indicate that length is confounded with a number of other predictor variables, including frication, transparency, semantic gender, abstractness, humanness, frequency, and age of acquisition.

To sort out these confounds, we conducted a series of stepwise regression analyses analogous to the ones re-

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

	% Correct	GMRT (%)	WRRT (%)
Total variance accounted for	12.3 [§]	66.1 [§]	59.5 [§]
Frication on initial consonant	0.0	1.0 ^{§(+)}	11.5 ^{§(+)}
Lead-in length	0.4	29.5 ^{§(+)}	19.5 ^{§(+)}
End length	0.0	17.8 ^{§(+)}	13.8 ^{§(+)}
Number of syllables	0.0	0.6 ^{‡(-)}	1.7 ^{§(-)}
All phonological indices	1.5 ⁽⁻⁾	54.7 [§]	55.4 [§]
Word frequency	0.0 ^a	0.8 ^{‡a(-)}	0.3 ^{*a(-)}
Form frequency	0.0 ^a	0.7 ^{‡a(-)}	0.2 ^a
Adjusted frequency	0.0 ^a	0.7 ^{‡a(-)}	0.3 ^{*a(-)}
All frequency indices	0.5	0.8 [†]	0.3
Source (adults only=1; infant=2)	0.3	2.4 ^{§(-)}	0.1
Semantic gender (- Sex=0; + Sex=1)	0.0	0.0	0.2
Abstractness (- Ab=1; + Ab=2)	0.2	0.1 ⁽⁺⁾	0.0 ⁽⁺⁾
Human (- Hum=0; + Hum=1)	0.7 ^{*(+)}	0.1	0.2 ^{*(+)}
Grammatical gender (fem.=1; masc.=2)	3.3 ^{§(+)}	0.0	0.3 ^{*(-)}
Transparency of ending (o/a=1; e=2)	5.9 ^{§(-)}	1.5 ^{§(+)}	0.1

* $p < .10$. † $p < .05$. ‡ $p < .01$. § $p < .001$. ^aAnalysis conducted with one frequency predictor only, no others entered on first step. ⁽⁺⁾Positive partial correlation for significant individual effects. ⁽⁻⁾Negative partial correlation for significant individual effects.

ported above for frication, with lead-in length and end length each given an opportunity to enter into the equation last when the other 12 predictors are controlled on the first step (see Table 4). Results indicate that lead-in length and end length do indeed make robust and independent contributions to variance in RT on both these tasks. On analyses of GMRT, lead-in length contributes 29.5% to the total variance accounted for when it is entered on the last step, whereas end length contributes an independent 17.8%. On analyses of WRRT, lead-in length adds 19.5% on the last step, whereas end length adds 13.8%. In all these cases (significant at $p < .001$), partial correlations indicate that the contribution is positive (i.e., greater length results in longer RTs).

These results are similar to findings by Radeau et al. (1989) for French. We agree with Radeau et al. that such findings are problematic for theories such as the Cohort Model, which assume that words can be uniquely identified (i.e., all competitors are eliminated) at the uniqueness point. If a strong version of this model were correct, then length after the uniqueness point should not be a significant contributor to WRRT. Instead, it appears from this study that lead-in length and end length both make unique contributions to RT, on both the word repetition and the gender-monitoring tasks. This result is quite compatible with earlier reports by Grosjean (1980), who showed that many words are not identified with confidence until after the uniqueness point. Indeed, when word recognition is

studied in context, Grosjean reports that some items are not uniquely identified until the next word is underway. At the same time, it is clear from our study that some words are identified before the uniqueness point—a finding that is also compatible with Grosjean's results. This must mean that additional cues are available to assist subjects in word recognition—which brings us to an analysis of length in syllables, another phonological factor which could have an effect on word recognition out of context.

Number of Syllables

Table 3 shows that number of syllables is significantly correlated with RT on both tasks. But as we might expect, Table 2 shows that number of syllables is also correlated with many of the other predictor variables, especially the length measures discussed above. Indeed, we might simply assume that syllable number is redundant with the above analyses of length. However, there are good reasons to believe that syllable structure adds additional information all its own. For example, the same syllable tends to be shorter in duration when it occurs within a multisyllabic word (e.g., the contrast between *typewriter* and *type* in English). In addition to these contrasts in duration, multisyllabic words may contain more coarticulatory and intonational cues, giving them an advantage in word recognition when absolute length is controlled.

To determine whether syllable structure makes an independent contribution to performance on our tasks, we

carried out stepwise regression analyses similar to those reported above (summarized in Table 4). When number of syllables was entered into the equation on the last step (controlling for the other 12 predictors), it had no significant effect on gender-monitoring accuracy. However, it explained an additional 0.6% of the variance in GMRT ($p < .01$) and an additional 1.7% of the variance in WRRT ($p < .001$). In both these cases, the partial correlations were negative—which means that words with more syllables lead to faster RTs. We suggest that this is true because subjects have more information about word identity at their disposal on multisyllabic words, when absolute length and other confounding factors are controlled.

To summarize so far, phonological factors play a major role in determining speed of lexical access, although they have no effect on accuracy in the gender-monitoring task. Taken together, the four phonological measures (frication, lead-in length, end length, and number of syllables) account for 55.4% of the variance in WRRT when other factors are controlled ($p < .001$), compared with 54.7% for GMRT ($p < .001$), and only 1.5% of the variance in gender accuracy scores ($p < .10$). For the two RT measures, these phonological contributions are so large that they do not leave much room for other factors to have an effect. And, indeed, the raw correlations in Table 3 suggest that contributions from the other predictor variables are relatively small. Nevertheless, it is useful to determine whether these factors have any significant effect at all when phonological variables are controlled.

Frequency

Three different indices of frequency were extracted from the De Mauro norms for our purposes here: absolute frequency of the root word across the corpus (word frequency, including all singulars and plurals), frequency for the particular morphological form of that word (form frequency), and a weighted index of root word frequency that takes into account the word distribution across different conversational contexts (adjusted frequency). As Table 2 shows, these three indices were highly correlated across the 468 nouns employed in our study: $r = +.94$ for word frequency and form frequency, $+ .98$ for word frequency and adjusted frequency, and $+ .93$ for form frequency and adjusted frequency. Given the magnitude of these correlations, we should not expect the three indices to differ markedly in their contributions to performance on gender monitoring and word repetition.

Indeed, the three frequency indices did not differ very much in their pattern of correlations. We were more surprised, however, to find that frequency plays a very small role across the board on these two tasks. Table 3 summarizes raw correlations between the three frequency measures and our three measures of performance. Gender-monitoring accuracy was not correlated with any of the frequency measures. Of the six correlations between frequency and RT on either task, four were significant beyond the $p < .05$ level: GMRT correlates at $r = -.09$ with word frequency, $r = -.12$ with form frequency, and $r = -.10$ with adjusted frequency; WRRT correlates at $r =$

$-.09$ with form frequency. These effects are all in the predicted direction (faster performance on more frequent items), but, on average, they account for less than 1% of the RT variance.

Because frequency is confounded with several of the other predictor variables in this data set (see Table 2), it is possible that frequency effects were masked or suppressed by other associated factors. To investigate this possibility, we carried out a series of stepwise regression analyses, entering each of the frequency measures into the equation last. Because these measures are so highly correlated (between $+ .93$ and $+ .98$), we did not use the other two frequency measures as controls on Step 1, since this would have effectively wiped out any contribution that an individual index could have made on Step 2. Hence, these regressions involve 10 predictors on Step 1, instead of 12. In addition, we looked at the contribution of the three frequency indices when they are added together as a block on the last step.

Results (summarized in Table 4) suggest that frequency still has relatively little effect on performance when other factors are controlled. Entered together on the last step, the frequency variables added a nonsignificant 0.5% to the variance in gender-monitoring accuracy, and none of the contributions from individual frequency measures reached significance. The three frequency indices taken together added a significant but minuscule 0.8% to the variance in GMRT ($p < .05$). The contribution of each individual predictor on the last step varied between 0.7% and 0.8%, also statistically reliable ($p < .01$) but very small. Partial correlations showed that these small effects were all in the predicted direction (i.e., faster RTs on higher frequency items). Frequency made no significant contribution to WRRTs, individually or as a block.

It may be that frequency effects have been swamped by other, more powerful factors in this study. To determine whether frequency effects would emerge more clearly in words that do not vary along all these dimensions, we conducted some additional analyses on the 56 pairs of words that differ only on the final vowel.⁵ Within each pair, we used the frequency norms to determine which of the two forms was dominant (i.e., which was most frequent in word and/or form frequency). Simple F tests were conducted to compare the dominant and nondominant items on each of our dependent variables. None of the comparisons even approached significance [$F(1,111) < 1.00$, in all three cases]. We then broke the minimal pairs down into other subsets, based on the intuition that frequency might play a more important role when other semantic factors were controlled. The various subsets that we tried included concrete versus abstract words, words with and without semantic gender, and words with human versus nonhuman referents. We also divided the pairs into words that share a semantic base (e.g., *ragazzo/ragazza*, as in *boy/girl*) and words that are unrelated (e.g., *caso/casa*, as in *case/house*).⁶ No matter how we looked at the data, there did not appear to be a significant advantage for higher frequency items within these minimal pairs (i.e., the difference between dominant and nondominant items did not reach significance, with $F < 1.00$, in every analysis).

In one last effort to see if larger frequency effects were lurking somewhere in the data, we went back to the full sample of 468 words and ran a series of correlations of frequency (all three indices) with RT (for both tasks) within different subcategories of items. For example, we checked to see whether larger effects of frequency would emerge for words that are phonologically opaque (i.e., words for which the subjects could not use superficial surface cues to assign gender). If anything, results were in the opposite direction: correlations for the 373 transparent words ranged from $-.08$ ($p < .06$) to $-.16$ ($p < .001$), and correlations for the 95 opaque words ranged from $+.00$ (n.s.) to $+.10$ (n.s.). Another possibility could lie in the fact that Italian words tend to be longer than the English words that have been used in most studies showing frequency effects in lexical access. We therefore conducted separate correlational analyses of RT and frequency for words with two syllables, three syllables, and four to five syllables. The prediction was that shorter words would show larger frequency effects, because they are more similar to the words that have been used in most English-language studies. In fact, when words were broken down into subsets of this kind, none of the correlations were reliable, no doubt due to the decrease in sample size. Another analysis was conducted to see whether frequency effects would prove larger for words without an initial fricative (on the argument that fricatives skew the variance so much that they mask other, more lexically based effects). For the 308 words without an initial fricative, correlations with RT were still disappointingly low, ranging from $-.07$ (n.s.) to $-.15$ ($p < .01$). In short, no matter how we look at these data, we are left with the conclusion that spoken-word frequency contributes very little to these measures of auditory lexical access in Italian.

Given all the attention that has been paid to frequency in the literature on lexical access, why are all these frequency effects so small? The role of frequency in recognition of visually presented words is well established (see Balota, 1992, for a review), but its effect on recognition of auditorially presented words is still controversial (see Luce et al., 1990; Marslen-Wilson, 1990; Liu & Bates, 1993). When such effects are found, they tend to be very small—as are the small correlations presented here. In this respect, our findings for Italian are compatible with other reports for recognition of spoken words in English (but see DeRuiter, 1992, and Taft & Hambly, 1986).

Age of Acquisition

In our list of stimuli, we had included a combination of words from the De Mauro et al. frequency norms for spoken conversations by Italian adults and words from the MacArthur Communicative Development Inventory for the first stages of lexical development in Italian infants. We included words from both sources for two reasons. First, as developmental psycholinguists, we are interested in the factors that influence the order in which words are acquired (e.g., the putative relationship between acquisition and frequency in the adult language). Second, a number of studies in the psycholinguistic literature have im-

plicated age of acquisition as a factor in lexical access by adults (e.g., Morrison, Ellis, & Quinlan, 1992).

In selecting items for this exploratory study, we were immediately struck by the lack of overlap between items in the De Mauro et al. adult norms and the words that are first acquired by Italian children. Adults spend very little time talking about animals, toys, and breakfast. Infants spend no time at all talking about real estate and office politics. In the adult norms, the most common substantives are relatively abstract words, such as *persona*, *tipo*, and *caso* (i.e., *person*, *type*, and *case*). In the language acquired by small children, first words tend to be common household items, toys, food, animals, and places to go. In part because of the striking lack of overlap between these two sources, there are significant negative correlations between source and the various indices of frequency (see Table 2). Indeed, many items on the MacArthur list have a frequency of zero on the De Mauro list.

Not surprisingly, there was a significant negative correlation between abstractness and age of acquisition ($r = -.36$, $p < .001$), and there was a tendency for children to acquire words with semantic gender (no doubt reflecting a general bias toward animals and people). Words that are acquired in infancy tend to end with an unambiguous gender marking ($r = -.14$, $p < .001$). In addition, there are confounds between age of acquisition and word length. Specifically, words from the MacArthur list tend to be slightly longer (2.29 syllables for the adult source, 2.64 syllables for the infant source; $t = 5.79$, $p < .001$). Although it may seem quite surprising that infants begin with slightly longer words (on average) than those used most often by adults in informal speech, this finding is quite consonant with the frequency effects reported above (remember Zipf's Law regarding the relationship between frequency and word length in informal speech). That is, the nouns that are used most often by adults tend to be rather short but also relatively abstract (designed to fit a large array of situations). It is also interesting to note that the age of acquisition variable was correlated with total length ($r = +.21$, $p < .001$) and with length of the end segment ($r = +.18$, $p < .001$), but it was not related to length of the lead-in segment ($r = +.00$). In other words, children may pick up slightly longer words, compared with the high-frequency abstract words used by adults, but this does not mean that they favor words that take a longer time to recognize!

Turning to the correlations in Table 3 with our dependent variables, it appears that subjects are slightly but not significantly more accurate in judging the gender of words that are acquired early by children ($r = +.07$, $p < .10$). They are also slightly faster to judge words that are acquired early (for GMRT, $r = -.08$, $p < .05$) but slightly slower to repeat them (for WRRT, $r = +.08$, $p < .05$). To control for all the confounds between source and the other predictor variables, we also carried out stepwise regression analyses on accuracy, GMRT, and WRRT, with age of acquisition entered into the equation on the last step after variance due to the other 12 predictors was removed (see Table 4). These analyses show that age of acquisition does

make a reliable contribution on the last step to GMRT (2.4%, $p < .001$), but it made no significant contribution to gender-monitoring accuracy (0.3%, n.s.) or to WRRT (0.1%, n.s.). Examination of the partial correlations showed that the contribution to GMRT is negative (i.e., words that are acquired earlier by children elicit faster RTs), in line with previous studies of estimated age of acquisition on performance in lexical tasks (see Morrison et al., 1992).

Semantic Variables

Three dichotomous semantic variables were available for our purposes here: semantic gender (+/-Sexed), abstractness (+/-Abstract), and humanness (+/-Human). As Balota (1992) has shown, semantic factors, such as imageability and concreteness, do increase speed of lexical access in the visual modality when words are presented out of context. It is less clear whether these effects hold in the auditory modality, and it is not at all clear whether these effects pertain to gender classification. However, by analogy to the effects that have been reported for visually presented words, we might predict that words will be repeated faster and gender classifications will take less time when the word has semantic as well as grammatical gender.

The correlations between predictor and performance variables in Table 3 suggest that semantic factors do affect RTs, but they do not necessarily make things easier. For example, semantic gender and humanness are positively correlated with GMRT, exactly the opposite of what we would predict if semantic gender were a facilitating factor. However, the correlations among predictor variables in Table 2 show that these semantic measures are confounded with the strong phonological effects that we described above. That is, words with the classification +Sexed tend to be longer in milliseconds, and they have more syllables. The same is true for words classified as +Human.

To control for confounds among the various predictor variables, we again carried out stepwise regression analyses controlling for all of the other predictors on the first step (see Table 4). Results suggest that semantic variables make no independent contribution to performance on these tasks. None of the unique contributions of semantic gender or abstractness even approached significance. The variable +/-Human added 0.7% to the variance in gender-monitoring accuracy on the last step ($p < .06$), with partial correlations showing that judgments were slightly more accurate for words with human referents. However, the contribution to GMRT did not even come close (0.1%, n.s.), and the contribution to WRRT approached significance (0.2%, $p < .10$), but in the wrong direction (i.e., slightly longer WRRTs for words with human referents).

We must conclude that the effects of semantic factors in this experiment are minimal. In particular, semantic gender does not have a significant effect on gender classification. This brings us to our final set of factors: grammatical gender and phonological transparency of the gender cue.

Grammatical Gender and Transparency

Table 3 shows that grammatical gender is significant and positively correlated with gender-monitoring accu-

racy ($r = +.16, p < .001$), which means that subjects are somewhat more accurate in judging masculine words. There were no significant correlations between grammatical gender and either of the RT measures.

Table 3 also shows that the phonological transparency of gender marking has a robust effect on both accuracy and RT in the gender-monitoring task. Specifically, words that end in the ambiguous vowel *e* lead to lower accuracy ($r = -.24, p < .001$) and slower RTs ($r = +.21, p < .001$). However, transparency has no reliable effect on WRRT ($r = +.06, n.s.$).

Although Table 2 shows that grammatical gender is happily independent of all the other predictor variables, transparency is confounded with length (*e*-final words are longer, especially in the end segment, and have more syllables) and with age of acquisition (children prefer transparent words that end in *o* or *a*). We therefore carried out another series of stepwise regression analyses on the dependent measures, examining the independent contributions of gender and transparency when all of the remaining predictor variables are controlled on the first step (see Table 4).

For the most part, these regressions mirror the above results with raw correlations. When grammatical gender is entered on the last step, it adds an additional 3.3% to gender-monitoring accuracy ($p < .001$), but the contribution to GMRT is not reliable (0.0%, n.s.). The contribution to WRRT on the last step approaches significance (0.3%, $p < .07$), with partial correlations showing that word repetition is slightly faster for masculine words.

When transparency is entered on the last step, it has reliable effects on both aspects of gender-monitoring performance. Transparency increases the variance in accuracy by 5.9% ($p < .001$), with a negative partial correlation indicating phonologically opaque words elicit more errors. It also increases the variance in gender monitoring by 1.5% ($p < .001$), with a positive partial correlation indicating that opaque words tend to slow things down. However, transparency has no detectable effect on WRRT when other factors are controlled (0.1%, n.s.).

We have established that grammatical gender and phonological transparency of gender marking have a significant effect on accuracy and/or RT in the gender-monitoring task, while these factors have little or no influence on WRRT. Because the form of the final syllable plays an important role in extraction of gender information, it seemed appropriate to explore the interaction between grammatical gender and transparency in more detail. Toward this end, we carried out a series of 2×2 analyses of variance for each of our three dependent variables.

In the analysis of word repetition, there were no significant main effects or interactions. This is not surprising, since the contribution of both these factors to word repetition was relatively small in our regression analyses. In the analysis of GMRT, there was a significant main effect of phonological transparency [$F(1,457) = 21.24, p < .0001$], reflecting faster classifications on transparent words ($M = 1,122$ msec) than on opaque words ($M = 1,179$ msec). The main effect of grammatical gender and

the gender \times transparency interaction were not reliable. In the analysis of gender-monitoring accuracy, there was a significant main effect of gender [$F(1,467) = 17.89, p < .0001$], reflecting more errors on feminine words, and a main effect of transparency [$F(1,467) = 36.19, p < .0001$], reflecting more errors on words that end in the ambiguous vowel *e*. The interaction also reached significance [$F(1,467) = 38.93, p < .0001$]. As illustrated in Figure 2, the largest number of errors in this experiment were committed on feminine words that end in *e*. This category resulted in a mean accuracy of 87.9% (12.1% error), compared with a mean of 95.9% (4.1% error) for masculine words ending in *e*. In fact, the accuracy level for opaque masculine words was quite comparable to the rates for transparent feminine words ending in *a* (mean accuracy = 95.7%, 4.3% error) and for transparent masculine words ending in *o* (mean accuracy = 96.2%, 3.8% error).

It has been known for some time that phonologically opaque words present problems for novice first- and second-language learners of Italian and for individuals with language disorders (see Caselli, Leonard, Volterra, & Campagnoli, 1993; Leonard, Bortolini, Caselli, & Sabbadini, 1993). However, to our knowledge, this is the first demonstration that phonological transparency affects performance by healthy adult native speakers. If the same effects had been observed on both of our tasks (i.e., on word repetition and gender monitoring), then we might conclude that there are interactions between the level of lemma retrieval (where grammatical gender should be retrieved) and the level of word form assignment (where phonological effects should apply). However, because no effects of phonological transparency were observed in word repetition, it seems more likely that the feedback between phonological marking and lexical/grammatical gender occurs after word access, during the stage at which a binary decision about gender is made. We will return to this point later.

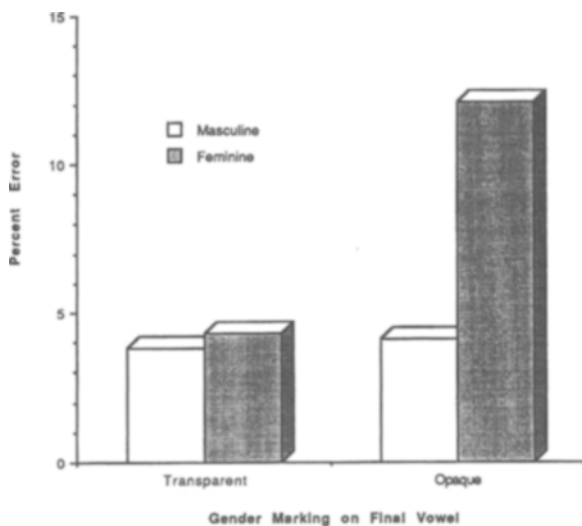


Figure 2. Percent errors in gender monitoring as a function of grammatical gender and phonological transparency of the final vowel.

The interaction between gender and transparency in gender-monitoring accuracy is more surprising, and the explanation is less obvious. However, opaque feminine words have also proven especially problematic in tests of Italian grammatical morphology in normal and language-disordered children (Leonard, Bortolini, Caselli, McGregor, & Sabbadini, 1992) and in deaf adults (Caselli, Rampelli, Volterra, & Maragna, 1994). To account for this fact, Caselli et al. have invoked the problem of homophony. Consider the following singular/plural declensions:

Feminine transparent singular:	<i>ragazza</i>	(girl)
Feminine transparent plural:	<i>ragazze</i>	(girls)
Feminine opaque singular:	<i>tigre</i>	(tiger)
Feminine opaque plural:	<i>tigri</i>	(tigers)
Masculine transparent singular:	<i>ragazzo</i>	(boy)
Masculine transparent plural:	<i>ragazzi</i>	(boys)
Masculine opaque singular:	<i>leone</i>	(lion)
Masculine opaque plural:	<i>leoni</i>	(lions)

Note that masculine forms always end with *i* in the plural, regardless of their singular forms. For feminine nouns, things are much more complicated. Transparent feminine forms end in *a* in the singular and *e* in the plural; opaque feminine forms end in *e* in the singular and *i* in the plural. Hence, the final vowel *e* is homophonous (i.e., the same morpheme can signal feminine singular, masculine singular, or feminine plural—a fact which may place such items at a considerable disadvantage in a gender-monitoring task).

CONCLUSION

The primary goal of this study was to examine the contributions of semantic and grammatical gender to lexical access in Italian. A secondary goal (and a necessary step to the first) was to investigate the range of factors that contribute to lexical access in word repetition and gender monitoring, in the auditory modality, in the Italian language. Let us consider what progress has been made toward each of these goals and then consider some implications of the present study for theories of lexical access (with emphasis on directions for future research).

Starting with the secondary goal, our results suggest that word repetition and gender monitoring (two relatively new procedures) are both sensitive measures of auditory word recognition in Italian. For example, RT on both tasks was tied to the uniqueness point (i.e., the point at which a word separates from all its closest competitors). And when length is taken into account (keeping in mind that Italian words are often 2–3 times longer than their English translation equivalents), performance on both tasks is remarkably fast and efficient. Indeed, our Italian subjects were able to initiate word repetitions about 180 msec after the uniqueness point, on average, with gender classifications occurring approximately 180 msec later.

We also explored the contribution of several factors that are known to play a role in auditory and/or visual word recognition in other languages, including frequency, age of acquisition, semantic factors (e.g., abstractness), and

several phonological variables (e.g., number of syllables, length in milliseconds, initial frication). In this study of isolated Italian words, phonological factors appear to be especially important, including length and frication (which slow things down) and number of syllables (which speeds things up when other factors are controlled). These findings underscore the close relationship between phonology and morphology in real-time processing. By contrast, variables that have proven important in other studies of lexical access (usually in English, and usually in the visual modality) appear to play very little role in this study of auditory word access in Italian. Age of acquisition did make an independent contribution to GMRTs, with faster RTs on words acquired in infancy. However, semantic factors had no detectable effect on performance in either task, and the effects of frequency were minute to nonexistent when other confounds were controlled.

This brings us to the primary goal of the present study, on the role of gender in lexical access. Three findings are relevant to this issue: (1) the effects (or noneffects) of semantic gender, (2) the effects of grammatical gender (with emphasis on phonological cues to gender on the final vowel), and (3) differences in performance profiles for gender monitoring (which requires explicit attention to the gender variable) and word repetition (which does not require attention to gender).

With regard to the first point, it now seems clear that semantic gender has little effect on lexical access—at least for normal Italian adults, in the tasks that we have used here. It seems that we have found a domain of human behavior in which sexuality is irrelevant.

With regard to the second point, we did find effects of phonological markings for gender. Specifically, Italian subjects take less time to classify a word as masculine or feminine if that word ends in a phonologically transparent vowel (*o* for masculine, and *a* for feminine), compared with words that end in a phonologically ambiguous vowel (*e* for both masculine and feminine). In addition, they made more errors on *e*-final words. A selective disadvantage for phonologically ambiguous words has been observed in children acquiring Italian, in deaf adults, and in language-disordered children (Caselli et al., 1993; Caselli et al., 1994; Leonard et al., 1992; Leonard et al., 1993). However, to the best of our knowledge, the present study constitutes the first demonstration that normal adults are affected by these cues. Errors were particularly frequent for *e*-final feminine words, a finding that may have something to do with homophony (e.g., the fact that transparent feminine words end in *e* in the plural while opaque feminine words end in *e* in the singular).

This brings us to a crucial difference between tasks, with implications for the locus of gender effects. Specifically, effects of phonological transparency were observed only on the gender-monitoring task. This contrast between tasks suggests that effects of surface gender marking may be restricted to a postlexical stage in processing and/or to tasks in which conscious attention to the gender dimension is required. We may speculate that subjects use the phonological cue in a final “check” before responding in

the gender-monitoring task, a process that does not occur in word repetition.

We did not set out to test competing theories of lexical access in the present study, but our results are relevant to a number of issues in this field.

Our findings regarding word recognition before and after the uniqueness point are relevant to models of lexical access in comprehension. It is clear that Italian subjects make use of multiple cues to word recognition, permitting them to identify many words before their uniqueness point. At the same time, we also find that RT is influenced by the length of the end segment (i.e., the remainder of the word after the uniqueness point). These findings (like those of Radeau et al. for French) are difficult to reconcile with early versions of the Cohort Model, where word candidates are accessed and eliminated in a strict left-to-right order. However, they are compatible with more recent versions of the Cohort Model, in which word access begins as early as possible, based on early and incomplete phonetic information (including coarticulatory cues that sometimes precede the onset of the target word). These results are also compatible with gating studies by Grosjean and colleagues, who have shown that many words in English and French are recognized before the uniqueness point, whereas others are not identified with confidence until hundreds of milliseconds after the word is complete.

Our results also provide information about the locus of gender effects in both comprehension and production. In Levelt's model of word production (Levelt, 1989), grammatical gender is retrieved at the moment of lexical access, while phonological cues to gender belong to an independent level of word form. Bidirectional interactions between these levels are precluded by the model, at least for word production. By contrast, Dell and O'Sheaghda (1992) have proposed a model of word production in which information from the word form level can feed back onto the stage of lemma retrieval, in a continuous bidirectional flow. Assuming that the word repetition task constitutes a form of word production (an assumption that may be controversial), results of the present study are largely compatible with Levelt's view.

This conclusion may not hold when lexical access is studied in a phrase or sentence context. In ongoing studies within our laboratory (e.g., Bates, Devescovi, Hernandez, & Pizzamiglio, 1994), we are now using the word repetition and gender-monitoring tasks to investigate gender priming, operationalized in terms of a match or mismatch in gender between an adjective prime and a noun target (e.g., *BRUTTO-TORRE* or *beautiful* [masc.]—*tower* [masc.] vs. *BRUTTA-TORRE* or *beautiful* [fem.]—*tower* [masc.]). This includes the use of an ecologically valid neutral baseline, with gender-ambiguous adjectives that end in *e* (e.g., *GRANDE-TORRE* or *great* [amb.]—*tower* [masc.]). The new studies demonstrate clear-cut evidence for gender priming (i.e., faster RTs when adjectives and nouns match), in line with findings by Grosjean et al. (1994) for French. For word repetition, significant facilitation and inhibition are both observed relative to the neutral baseline, which suggests that gender priming may in-

volve a combination of prelexical as well as postlexical effects. If one assumes that word repetition is a form of word production (a controversial assumption, as noted above), this could be viewed as evidence for interaction between the level of word form and the level of lemma retrieval, a potential problem for the Levelt model.

At the very least, it is clear that the lexical and phonological factors uncovered in this study need to be controlled in future research on gender processing—to the extent that such controls are permitted by the language. In fact, as we set about the business of unconfounding all these factors to create new experimental stimuli, we are constantly reminded that many confounds are built into the language. If these correlations are so well entrenched, they are probably there for a reason. These reasons will hopefully become clear as we start to collect more information about gender processing in a phrase or sentence context—that is, the circumstances in which the pervasive and mysterious phenomenon of grammatical gender first evolved.

REFERENCES

- BALOTA, D. A. (1992). Visual word recognition: The journey from features to meaning. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 303-358). San Diego, CA: Academic Press.
- BATES, E., & DEVESCOVI, A. (1989). A cross-linguistic approach to sentence production. In B. MacWhinney & E. Bates (Eds.), *The cross-linguistic study of sentence processing* (pp. 225-253). New York: Cambridge University Press.
- BATES, E., DEVESCOVI, A., HERNANDEZ, A., & PIZZAMIGLIO, L. (1994). *Gender priming in Italian*. Unpublished manuscript, University of California, San Diego.
- BATES, E., & MACWHINNEY, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The cross-linguistic study of sentence processing* (pp. 3-73). New York: Cambridge University Press.
- BROOKS, J., BRAINE, M. D. S., CATALANO, L., BRODY, R. E., & SUDHALTER, V. (1993). Acquisition of gender-like noun subclasses in an artificial language: The contribution of phonological markers to learning. *Journal of Memory & Language*, **32**, 76-95.
- BROWN, P., SENFT, G., & WHEELDON, L. (Eds.) (1993). *Max-Planck Institute for Psycholinguistics: Annual Report 1992*. Nijmegen: Max-Planck Institute for Psycholinguistics.
- BURANI, C. (1992). Patterns of inflectional errors with reference to the Italian adjectival system. *Rivista di Linguistica*, **4**, 255-272.
- CASELLI, M. C., & CASADIO, P. (1992). *Fondazione MacArthur: Lo sviluppo comunicativo nella prima infanzia* [MacArthur Foundation: Communicative development in early infancy]. Rome: Istituto di Psicologia, Consiglio Nazionale delle Ricerche.
- CASELLI, M. C., LEONARD, L., VOLTERRA, V., & CAMPAGNOLI, M. (1993). Toward mastery of Italian morphology: A cross-sectional study. *Journal of Child Language*, **20**, 377-393.
- CASELLI, M. C., RAMPPELLI, L., VOLTERRA, V., & MARAGNA, S. (1994). *Linguaggio e sordità* [Language and deafness]. Florence: La Nuova Italia.
- CASSIDY, K., & KELLY, M. (1991). Phonological information for grammatical category assignment. *Journal of Memory & Language*, **30**, 348-369.
- COHEN, J., MACWHINNEY, B., FLATT, M., & PROVOST, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments, & Computers*, **25**, 257-271.
- COLÉ, P., & SÉGUI, J. (1994). Grammatical incongruity and vocabulary types. *Memory & Cognition*, **22**, 387-394.
- CONNINE, C. M., MULLENIX, J., SHERNOFF, E., & YELEN, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 1084-1096.
- DELL, G., & O'SHEAGHDA, P. (1992). Stages of lexical access in language production. *Cognition*, **42**, 287-314.
- DE MAURO, T., MANCINI, F., VEDOVELLI, M., & VOGHERA, M. (1993). *Lessico di frequenza dell'italiano parlato* [Frequency dictionary of spoken Italian]. Rome: Fondazione IBM Italia.
- DE RUITER, J. P. (1992). *Predicting and measuring the moment of auditory word recognition*. Unpublished doctoral dissertation, University of Nijmegen, Department of Cognitive Science.
- DEUTSCH, W., & WIJNEN, F. (1985). The article's noun and the noun's article: Explorations into the representation and access of linguistic gender in Dutch. *Linguistics*, **23**, 793-810.
- DEVESCOVI, A., D'AMICO, S., SMITH, S., MIMICA, I., & BATES, E. (1993). *The development of sentence comprehension in Italian and Serbo-Croatian: Local versus distributed cues*. Unpublished manuscript, University of Rome "La Sapienza," Department of Psychology.
- FODOR, J. (1959). The origin of grammatical gender. *Lingua*, **8**, 1-41.
- FRAUENFELDER, U., & TYLER, L. (Eds.) (1987). *Spoken word recognition*. Cambridge, MA: MIT Press/Bradford Books.
- FRIEDERICI, F., & SCHRIEFERS, H. (1993). *Semantic and morphosyntactic context effects on word recognition: The influence of peripheral and central processing variables*. Unpublished manuscript, Cognitive Science Lab, Institut für Psychologie, Berlin.
- GOODMAN, J., & HUTTENLOCHER, J. (1988). Do we know how people identify spoken words? *Journal of Memory & Language*, **27**, 684-698.
- GROSJEAN, F. (1980). Spoken-word recognition processes and the gating paradigm. *Perception & Psychophysics*, **28**, 267-283.
- 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.
- HERRON, D. (1994). *Grammatical and acoustic contributions to the processing of open- and closed-class words*. Unpublished master's thesis, University of California, San Diego, Department of Cognitive Science.
- International dictionary: English/Spanish, Spanish/English*. (1973). New York: Simon & Schuster.
- JAREMA, G., & FRIEDERICI, A. D. (1994). Processing articles and pronouns in agrammatic aphasia: Evidence from French. *Brain & Language*, **46**, 683-694.
- KAIL, M. (1989). Cue validity, cue cost, and processing types in sentence comprehension in French and Spanish. In B. MacWhinney & E. Bates (Eds.), *The cross-linguistic study of sentence processing* (pp. 77-117). New York: Cambridge University Press.
- KILBORN, K. (1987). *Sentence processing in a second language: Seeking a performance definition of fluency*. Unpublished doctoral dissertation, University of California, San Diego, Department of Psychology.
- LEONARD, L., BORTOLINI, U., CASELLI, M. C., MCGREGOR, K., & SABBADINI, L. (1992). Morphological deficits in children with specific language impairment: The status of features in the underlying grammar. *Language Acquisition*, **2**, 151-179.
- LEONARD, L. B., BORTOLINI, U., CASELLI, M. C., & SABBADINI, L. (1993). The use of articles by Italian-speaking children with specific language impairment. *Clinical Linguistics & Phonetics*, **7**, 19-27.
- LEVELT, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- LIU, H., & BATES, E. (1993). *Single-word shadowing and the study of lexical access* (Tech. Rep. No. 9306). La Jolla: University of California, San Diego, Center for Research in Language.
- LUCE, P., PISONI, D., & GOLDINGER, S. (1990). Similarity neighborhoods of spoken words. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing* (pp. 122-147). Cambridge, MA: MIT Press.
- MACWHINNEY, B. (1978). The acquisition of morphophonology. *Monographs of the Society for Research in Child Development*, **43** (2, Serial No. 174).
- MACWHINNEY, B., LEINBACH, J., TARABAN, R., & McDONALD, J. (1989). Language learning: Cues or rules? *Journal of Memory & Language*, **28**, 255-277.
- MARATOS, M. (1981). Problems in categorial evolution: Can formal categories arise from semantic ones? In W. Deutsch (Ed.), *The child's construction of language*. London/New York: Academic Press.
- MARSLÉN-WILSON, W. (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing* (pp. 148-172). Cambridge, MA: MIT Press.

- MILLS, A. (1986). *The acquisition of gender: A study of English and German*. Berlin: Springer-Verlag.
- MORRISON, C., ELLIS, A., & QUINLAN, P. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory & Cognition*, **20**, 705-714.
- MULFORD, R. (1983). Comprehension of Icelandic pronouns: Semantic vs. formal factors. *Journal of Child Language*, **12**, 443-453.
- ORSOLINI, M. (1993, July). *Children's semantic processing of gender cues*. Paper presented at the Sixth International Congress for the Study of Child Language, Trieste, Italy.
- PALAZZI, F. (1973). *Novissimo dizionario della lingua italiana* [New dictionary of the Italian language]. Milan: Ceschino.
- RADEAU, M., & MORAIS, J. (1990). The uniqueness point effect in the shadowing of spoken words. *Speech Communication*, **9**, 155-164.
- RADEAU, M., MOUSTY, P., & BERTELSON, P. (1989). The effect of the uniqueness point in spoken-word recognition. *Psychological Research*, **51**, 123-128.
- SLOWIACZEK, L. M. (1994). Semantic priming in a single-word shadowing task. *American Journal of Psychology*, **107**, 245-260.
- TAFT, M., & HAMBLY, G. (1986). Exploring the Cohort Model of spoken word recognition. *Cognition*, **22**, 259-282.
- WHALEN, D. (1991). Subcategorical phonetic mismatches and lexical access. *Perception & Psychophysics*, **50**, 351-360.
- ZUBIN, D., & KÖPCKE, K. (1981). Gender: A less than arbitrary grammatical category. In C. M. R. Hendrick & M. Miller (Eds.), *Papers from the Seventeenth Regional Meeting*. Chicago: Chicago Linguistic Society.
2. This is of course not true for adjectives and other modifiers, since gender assignment for these items depends upon the noun that they modify.
3. One might ask why gender is so often used for this referential function, compared with all the other semantic dimensions that could have been drafted to mark and track lexical items. For example, in technical discourse, one often finds arbitrary use of numbers or letters to create new lexical subtypes (e.g., Type A, Type B, Type C). So why gender? Bates and MacWhinney (1989) suggest that gender may have been a readily available and easily recognized marker for arbitrary subclassification in preliterate cultures that had no access to numbers or letters for this purpose.
4. In this paper, we used a simple presence/absence measure to evaluate the contribution of initial frication. However, the phonetic space that defines initial consonants could be partitioned in many different ways. For example, some exploratory analyses of these data show that RTs are significantly longer for words with an initial [s-] (including single consonants and/or consonant clusters) than they are for other fricatives or affricates, although the remaining fricatives and affricates still result in significantly longer RTs than do other initial consonant types. On the other hand, factors such as presence/absence of voicing on an initial fricative or affricate have no detectable effect on RT. Further exploration of phonetic contributions to lexical access would take us far outside the scope of the present study, but our preliminary findings suggest that such factors merit a detailed investigation in their own right.
5. Our thanks to Uli Frauenfelder for suggesting this analysis.
6. We also looked to see whether these two sets of words (i.e., pairs with and without a semantic contrast) differed in any other way (e.g., whether the words that carry a semantic contrast are easier to classify or faster to repeat), but a series of correlational analyses using the feature "shared semantics" yielded no significant relations to any of our dependent variables.

NOTES

1. In Spanish, the word *mar* can be construed as masculine (*el mar*) or feminine (*la mar*), depending on the context (*International Dictionary: English/Spanish, Spanish/English*, 1973). This peculiarity of Spanish appears to be an exception to the claim that gender is retrieved for nouns at the moment of lexical access, although this apparent exception could be resolved by postulating two homophonous lexical entries for *mar*, one feminine and one masculine.

(Manuscript received March 3, 1994;
revision accepted for publication January 31, 1995.)