

## The production of pronominal clitics: Implications for theories of lexical access

Chiara Finocchiaro

*Scuola Normale Superiore Pisa, Italy*

Alfonso Caramazza

*Harvard University, Cambridge, MA, USA*

In three experiments we investigated the locus of the frequency effect in lexical access and the mechanism of gender feature selection. In Experiment 1, participants were asked to produce gender-marked verb plus pronominal clitic utterances in Italian (e.g., “portalo” (bring it [masculine]) in response to a written verb and pictured object. We found that pronominal clitic production is sensitive to the frequency of the noun it replaces. This result locates the effect of word frequency in lexical access at the level where a word’s grammatical features are represented. In Experiments 2, 3A, and 3B we used a picture-word interference naming task and found that the gender of a distractor word does not affect the production of gender-marked clitics. This result, together with those of Experiments 3A and 3B, which show a semantic interference effect and the absence of a phonological facilitation effect in clitic production, respectively, allows the inference that the retrieval of grammatical gender is an automatic consequence of lexical node selection and not an independent selection process that operates on the principle of selection-by-activation level.

---

Correspondence should be addressed to Alfonso Caramazza, Department of Psychology, Harvard University, William James Hall, 33 Kirkland St., Cambridge, MA 02138, USA. Email: [caram@wjh.harvard.edu](mailto:caram@wjh.harvard.edu)

The research reported here was supported in part by NIH grant DC04542 to Alfonso Caramazza. We would like to thank the members of the Harvard Cognitive Neuropsychology Laboratory and especially Albert Costa for many helpful discussions on various aspects of the research reported here. We also thank Brad Mahon and Robin Peterson for helpful comments on an earlier version of this article.

## INTRODUCTION

The production of pronominal forms and the noun phrases (NPs) they replace share a crucial process. In both cases, the selection/computation of the appropriate phonological forms depends on the prior selection of a lexical node—the head noun—and its associated grammatical features and diacritics. This shared property of NP and pronoun production can be readily appreciated by considering the form and distribution of pronominal clitics<sup>1</sup> in Italian.

Consider the sentence “porta il piatto” (bring the plate). If the NP “il piatto” is contextually identified, it can be pronominalised resulting in the form “portalo” (bring it). Italian clitic pronouns, like NPs, are marked for number and grammatical gender, resulting in four distinct forms: *-lo* [masculine (mas.) singular (sing.)], *-li* [mas. plural (pl.)], *-la* [feminine (fem.) sing.], and *-le* [fem. pl.], as shown in sentences 1a–d.

- |     |  |   |         |
|-----|--|---|---------|
| 1a. | Porta il piatto [mas. sing.] (bring the plate)   | → | Portalo |
| 1b. | Porta i piatti [mas. pl.] (bring the plates)     | → | Portali |
| 1c. | Porta la forchetta [fem. sing.] (bring the fork) | → | Portala |
| 1d. | Porta le forchette [fem. pl.] (bring the forks)  | → | Portale |

As is apparent from these examples, the selection of the appropriate determiner of an NP and the selection of the appropriate form of a pronominal clitic both depend on the gender and number of the noun in the NP. For example, if the gender of the noun is masculine then the determiner would be either *il* or *i* depending on whether the NP is singular or plural, respectively;<sup>2</sup> similarly, when the pronominal clitic replaces a masculine NP its form would be either *-lo* or *-li* depending on whether it is singular or plural, respectively. Thus, the production of pronominal clitics and the production of noun phrases share at least one process: They both involve the selection of the set of grammatical features associated with the head noun of a noun phrase.

Of the several grammatical features that must be retrieved for the production of NPs and pronouns,<sup>3</sup> the selection of grammatical gender is particularly important in the present context. This is because unlike number, which is a diacritic value determined contextually, grammatical gender is an intrinsic property of lexical items, and therefore its retrieval

<sup>1</sup> Romance clitics are phonologically unstressed pronouns that can only be placed adjacent to the verb on which they depend.

<sup>2</sup> When the word following the definite determiner begins with /s+consonant/, /gn/, /z/ or a vowel, the determiner forms are *lo* (singular) and *gli* (plural).

<sup>3</sup> Another grammatical feature that must be selected as part of the information used to determine the form of pronouns is grammatical case. For example, English distinguishes nominative and dative case, as in the contrasts she/her, he/him.

entails the prior selection of the lexical item. In other words, pronoun production requires the selection of the lexical representation where the head noun's grammatical properties are specified, but not the selection of the phonological content of the noun. This property of gender-marked pronouns provides the basis for addressing questions about lexical access at the level where a word's grammatical features are represented independently of the processes implicated in the selection of the word's phonological content (see, e.g., Jescheniak, Schriefers, & Hantsch, 2001; Meyer & Bock, 1999). Here we consider two issues: the locus of the frequency effect in lexical access and the mechanism of grammatical feature selection. Although these two issues are distinct in that one concerns how lexical nodes are selected while the other concerns how the grammatical features associated with a lexical node are selected, they are fundamentally concerned with a common theoretical question: the dynamics of lexical node selection.

The locus of the frequency (and/or age of acquisition) effect in lexical access remains an unresolved issue. The issue is important because it is tied to questions about the dynamics of lexical selection, as well as to questions about the number of levels of representation in the lexical system. Three hypotheses can be distinguished. Two of these rely on a distinction between a level at which a word's grammatical properties are encoded (lemma) and a level at which the morphemes that compose the word are represented (lexeme). Within this architecture of the lexical system, one hypothesis assumes that the frequency effect arises at the lemma level (Dell, 1990) while another assumes that the frequency effect arises at the lexeme level (Jescheniak & Levelt, 1994). The motivations proffered in support of these different positions will be reviewed in the introduction to Experiment 1. For theories that assume that there is only one lexical layer that mediates between semantic and phonological representations, the locus of the frequency effect arises at that level of representation (Caramazza, 1997).

These hypotheses can be addressed by considering the effect of word frequency on the production of pronominal clitics (Experiment 1). In this context, the question becomes: Does the frequency of a word affect the production of the pronoun that stands for it? Since the selection of pronominal forms depends on the retrieval of a word's grammatical features and not its phonological content, a result showing that pronoun production is sensitive to the frequency of the word it replaces would indicate that the locus of the frequency effect is at the lexical level where grammatical feature selection occurs (Caramazza, 1997; Dell, 1990; Griffin & Bock, 1998). A result showing that pronoun production is insensitive to word frequency would suggest that the locus of the frequency effect could be at the lexeme level (Jescheniak & Levelt, 1994).

The other issue addressed here concerns the process of grammatical feature selection. Two proposals have been made. On one view, the selection of a word's grammatical features, such as gender, is an independent (and optional) process that occurs *after* the selection of the word's lexical node (La Heij, Mak, Sander & Willeboordse, 1998; Levelt, Roelofs, & Meyer, 1999; Schriefers, 1993). Grammatical feature selection is seen as a process similar to lexical node selection in that it constitutes an independent decision based on the relative levels of activation of grammatical feature nodes. That is, the assumption is that lexical nodes send activation to their respective grammatical features and the feature node with the highest level of activation is selected. Because this process of selection is based on the activation level of grammatical feature nodes, the selection decision can be "primed"—facilitated (Levelt, 2001) or interfered with (Schriefers, 1993). An alternative hypothesis views the retrieval of grammatical features as an automatic consequence of lexical node selection and therefore as not subject to *independent* facilitation or interference (Caramazza, Miozzo, Costa, Schiller, & Alario, 2001a). These alternative proposals can be evaluated by considering whether the production of a pronominal clitic is subject to facilitation/interference from the gender of a competing lexical item (Experiment 2). A result showing that the gender of a distractor word affects the production of a pronominal clitic would be evidence in support of the hypothesis that gender feature selection is a process that occurs independently of lexical node selection, and is based on the level of activation of feature nodes. A result showing that the gender of a distractor word does not affect pronominal clitic production would be consistent with the assumption that grammatical features are retrieved automatically as a direct consequence of lexical node selection.

To anticipate our results, we found that pronominal clitic production is sensitive to the frequency of the nouns that the pronouns replace, suggesting that the frequency effect has its locus at a level of lexical processing where a word's grammatical features are represented. We also found that the gender of distractor words does not affect the production of pronominal clitics, suggesting that a word's grammatical features become available automatically as part of lexical node selection. However, the interpretation of the results of these experiments rests in part on the assumption that pronominal clitic production helps reveal the effects of processing at stages of lexical access where a word's grammatical features are represented and not at subsequent stages of lexical processing. As already noted, this assumption is widely shared in the literature on pronoun production (e.g., Jescheniak & Schriefers, 2001; Jescheniak et al., 2001; Meyer & Bock, 1999). Nonetheless, it is important to demonstrate empirically the plausibility of this methodological assumption, and we do

this by introducing in the pronominal clitic production task experimental variables that are known to have their effects at distinct levels of lexical access (Experiment 3).

### EXPERIMENT 1. Effects of noun frequency on the production of pronominal clitics

Investigations of the locus of the frequency effect in lexical access have been motivated largely by the assumption that lexical access involves two distinct levels of lexical representation: a level where a word's grammatical properties are specified (lemma) and a level where the word's morphemes are specified (lexeme). Researchers have attempted to determine at which of these levels frequency has its effect in lexical access.<sup>4</sup>

The method of choice in such efforts has been to investigate the production of homophones. The reason for studying homophones is based on a further assumption that is made by some authors (Dell, 1990; Jescheniak & Levelt, 1994) about the architecture of the lexical system; namely, the assumption that homophones share a lexical representation at the lexeme level. Thus, for example, the members of the homophone pair "sun/son" would each have a distinct lemma representation but they would share the single lexeme node /son/. Given this assumption, it follows that if the frequency effect were located at the lexeme level the effective frequency in determining the speed of access of either "sun" or "son" would be the summed (homophone) frequency of "sun" and "son" and not their individual (specific-word) frequencies. Thus, the investigation of the effects of word frequency on homophone production could help determine the locus of the frequency effect in lexical access. However, research on homophone production has led to conflicting results and conclusions.

In an early investigation of the effects of homophone frequency on word production, Dell (1990) used an error-induction paradigm in which participants were asked to prepare for production visually presented phrases containing homophones such as "hymn to sing" or "him to sing", and to produce them as fast as possible upon the presentation of a cue. He found that homophone frequency was a better predictor of the probability of an error in the production of homophones than was specific-word

---

<sup>4</sup> Of course, it is possible that lexical access is frequency insensitive or that both levels of lexical representation are frequency sensitive. It is unlikely that lexical access is frequency insensitive, especially given the overwhelming evidence from aphasia, which shows that patients with intact semantic processing nonetheless show strong frequency effects in naming and other word production tasks (e.g., Caramazza & Hillis, 1990). It is far less clear that frequency sensitivity would be restricted to only one level of lexical representation if, in fact, there were to be two such levels of representation.

frequency, consistent with the assumption that homophones share a lexical representation. Dell interpreted this result within the framework of a highly interactive model of lexical access and concluded, with the help of simulations designed to reproduce the pattern of errors observed in his experiment, that the most plausible locus of the frequency effect is at the lemma level.<sup>5</sup>

Jescheniak and Levelt (1994) used a chronometric paradigm to investigate the locus of the frequency effect in lexical access. They asked Dutch/English bilinguals to translate English words into Dutch as fast as possible. Jescheniak and Levelt reasoned that if the locus of the frequency effect were at the lexeme level, and if homophones shared a lexeme representation, then reaction time (RT) in producing homophones should be the same as that for homophone frequency controls and faster than specific-word frequency controls. This is precisely the result they obtained, and they concluded that the locus of the frequency effect is at the level of lexeme access—the opposite of the conclusion reached by Dell on the basis of the homophone frequency effect he obtained with an error induction task.

This picture is complicated further by the fact that the homophone frequency effect reported by Jescheniak and Levelt (1994) is not obtained reliably. Caramazza, Costa, Miozzo, and Bi (2001a) failed to find a homophone frequency effect in a translation task similar to the one used by Jescheniak and Levelt, except that they used Spanish/English bilinguals instead of Dutch/English bilinguals. In a translation task with German/English bilinguals, Jescheniak, Meyer, and Levelt (2003) found that homophones were named faster than their specific-word frequency controls, but they also found that homophones were produced more slowly than their homophone frequency controls. This latter result is not consistent with the hypothesis that frequency has its effect at the lexeme level where homophones supposedly share a representation (see Caramazza, Bi, Costa, & Miozzo, 2004). Thus, the translation experiments on homophone production have failed to produce consistent results concerning the putative homophone frequency effect.

---

<sup>5</sup> Two points should be noted here. First, it is not obvious that the error-induction paradigm used by Dell reveals effects of lexical processing as opposed to post-lexical processes, given that the task required subjects to prepare and hold in short-term-memory a phonological sequence, which subsequently had to be produced as fast as possible. It is possible that the effects obtained in this task reveal the strength of association among a set of phonemes—determined in part by word frequency—rather than the process of lexical node selection. Second, a strongly interactive model, such as the one proposed by Dell (1990), does not necessarily require that homophones share a lexical representation in order to obtain a homophone frequency effect. It is possible that feedback from the identical sets of phonemes shared by homophones would suffice to obtain a homophone frequency effect.

Finally, simple picture naming experiments that compared the production of homophonic (e.g., nun) and non-homophonic names (e.g., tooth) have consistently found a specific-word and not a homophone frequency effect. In two experiments, one with English and one with Mandarin Chinese speakers, Caramazza et al. (2001a) found that participants produced homophones more slowly than homophone frequency controls, and as slow as specific-word frequency controls. Bonin and Fayol (2002) also found that their French-speaking participants named homophones more slowly than homophone frequency controls. These results cast further doubt on the reliability of the putative homophone frequency effect in speech production.

In light of the complex pattern of results that has emerged from research with homophone production tasks, it is clear that such data are at best inconclusive with respect to the locus of the frequency effect in lexical access. It is best, therefore, to seek converging evidence on the locus of the frequency effect with other experimental paradigms.

As already noted, the focus of this investigation is on those linguistic factors that are causally implicated in the selection of pronouns. We have identified these factors as consisting of the grammatical features associated with the NPs that pronouns replace, including phrasal diacritic features such as number and case as well as word-specific features such as the gender of the head nouns of the NPs. The retrieval of gender necessarily implicates the selection of the lexical representation where this information is specified. It follows, therefore, that the selection of such lexical nodes is causally implicated in pronoun selection. We have also argued that by using a pronoun production task it is possible to investigate the selection of these various types of representations in a context uncontaminated by the phonological properties of the noun that the pronoun replaces. Thus, the investigation of pronoun production may provide a privileged view of the processing factors that are involved in accessing the lexical representation where grammatical information is specified. One such factor is word frequency: Does word frequency affect lexical access at the level where a word's grammatical features are represented? If so, we would expect noun frequency to affect pronoun production since the grammatical features of the noun must be accessed in order to produce the pronoun that replaces it (Caramazza, 1997; Dell, 1990). However, if word frequency were to affect lexical processing at a later stage (Jeschaniak & Levelt, 1994), we would not expect this factor to affect pronoun production. This expectation is based on the assumption that such a stage of noun processing is irrelevant to the process of pronoun selection. The empirical question we address here is: Does noun frequency affect pronoun production latencies?

Consider the following task. A participant is shown a written verb in the infinitive (e.g., “portare”—to bring) followed by a picture of an object (e.g., a helmet, whose name in Italian is the masculine noun “casco”) and is required to produce a simple imperative sentence using the verb and a pronominal form of the name of the object—“portalo” (bring it), in this example. By using pictured objects whose names vary in their frequency of usage, we can assess whether word frequency affects pronoun production latencies. That is, we can ask whether participants produce simple imperative sentences such as “portalo” faster when the pictured object’s name is of high frequency (e.g., *letto* “bed”) than when it is of low frequency (e.g., *casco* “helmet”). In this experiment, we used a slightly more complicated imperative construction involving both direct and indirect objects resulting in simple sentences such as “portamelo” (“bring it to me”; literally: “bring to me it”) and “portacelo” (“bring it to us”; literally: “bring to us it”). The choice of the dative pronominal clitic was indicated by presenting a prepositional phrase (PP), either “a me” (to me) or “a noi” (to us), immediately above the picture. However, the relevant experimental manipulation concerned only the frequency of the direct objects—the names of the pictures—that are implicated in determining the form of the accusative clitics (*-lo* for masculine nouns and *-la* for feminine nouns).

## Method

*Participants.* Twelve native speakers of Italian (age range: 20–37), students and staff at universities in the Cambridge/Boston area took part in the experiment. They were paid for their participation.

*Materials.* A set of 40 pictures whose names varied in frequency (Bortolini, Tagliavini, & Zampolli, 1971) was selected. Half of the pictures had names in the low frequency (LF) range (0–12; mean = 3.4), the other half had names in the high frequency (HF) range (30–757; mean: 114.9. See Appendix A). Masculine and feminine nouns were equally represented in the two frequency groups. Picture names were also controlled for number of syllables (mean, HF = 2.5, LF = 2.5) and length (mean, HF = 6.5, LF = 6.5).

Each picture had a prepositional phrase (PP) printed (Bold caps, Geneva 32) above it. The PP could be “a me” (to me) or “a noi” (to us). Each picture appeared three times, twice with one PP, once with the other PP. “To me” and “to us” trials were counterbalanced across items of different genders (masculine, feminine) and of different frequency value (high, low).

Three blocks of trials were created. In a block, a given picture appeared once. Trials were randomised with the following constraints: (a) Pictures with names of the same gender or with the same PP printed above it were not to appear on more than three consecutive trials; (b) there was no semantic, phonological, or associative relationship between the names of pictures in consecutive trials; (c) the PPs “to me” and “to us” were equally represented in each block. Pictures with HF masculine nouns, HF feminine nouns, LF masculine nouns, LF feminine nouns were nearly equally represented within “to me” and “to us” trials.

We also selected 10 verbs. Each verb was paired with four different pictures (one each from the sets HF masculine, HF feminine, LF masculine, LF feminine). Verbs were always presented in the infinitive form.

Participants had to produce the given verb in the second person of the imperative with the correct clitics corresponding to the pictured object and to the PP. For instance, given the verb “portare” (to bring) and the picture of a cat with the PP “a me” (to me) printed above it, participants were required to say “portamelo” (literally: bring to me it [mas.]).

In order to perform the task, participants had to retrieve the gender and the number of the pictured noun (which was always singular in this experiment) so that they could determine the form of the accusative pronominal clitic. Furthermore, they had to determine the form of the dative pronominal clitic from the PPs “to me” and “to us”. Since the pronominal clitics corresponding to “a me” (to me) and “a noi” (to us) are, respectively, *-me* and *-ce*, and the accusative pronominal clitic forms for masculine and feminine singular nouns are, respectively, *-lo* and *-la*, there were four possible clitic combinations in our experiment: *-melo* (to me it [mas.]), *-mela* (to me it [fem.]), *-celo* (to us it [mas.]), *-cela* (to us it [fem.]). Thus, the produced forms varied as a function of the verb, the gender of the pictured noun and the PP printed above the picture.

Two additional sets of verb-picture pairs were created. The pairs of the first set ( $N = 40$ ) served as practice trials before the experiment proper; the pairs of the second set ( $N = 6$ ) served as warm-up stimuli at the beginning of each block.

*Procedure.* Participants were tested individually in a quiet testing room. They were seated at a distance of about 60 cm from the computer screen. The experiment started with a naming task aimed at familiarising participants with the pictures and their names. When participants produced a name other than the target response they were corrected. Such instances were very rare. Participants then performed a practice block. At the beginning of each trial, a question mark appeared in the centre of the computer screen. The question mark disappeared as soon as

participants pressed the space bar. Then, the verb appeared in the centre of the computer screen for 1000 ms and was immediately replaced by the picture. The picture was removed as soon as participants responded or 3000 ms had elapsed. Stimulus presentation was controlled by the program Pyscope (Cohen, MacWhinney, Flatt, & Provost, 1993). Response latencies were measured by means of a voice key. Instructions emphasised response speed and accuracy. The experimenter recorded the responses manually. At the end of the three experimental blocks, participants were given an additional block of trials in which they were required to name as fast as possible the pictured objects used in the pronominal clitic task.

*Analyses.* Verbal dysfluencies and responses different from the target were scored as errors. Errors, failures to trigger the voice key, equipment malfunction, and outliers—responses exceeding a participant's mean by three standard deviations—were removed from the analysis. These exclusionary criteria were also applied in the other experiments reported here. In the analyses of response latencies, two variables were considered: picture noun frequency (HF vs. LF) and repetition (first vs. second vs. third). These variables were treated as within-subject variables with one exception: In the by-items analyses ( $F_2$ ), frequency was considered a between-subject variable. The same analyses were repeated with participants' error rates as the dependent measure.

## Results

Discarded data accounted for 15.6% of the verb + clitics trials,<sup>6</sup> and for 7.3% of the responses in the simple picture-naming task at the end of the experiment.

In the simple picture-naming task carried out at the end of the experiment, participants were faster at naming HF pictures than LF pictures, HF = 685 ms, LF = 740 ms;  $F_1(1, 11) = 23.6, p = .0005$ ;  $F_2(1, 38) = 6.3, p = .02$ . See Table 1). Thus we can confidently assume that the experimental items are appropriate for revealing a frequency effect in pronoun production if there is one.

ANOVAs carried out on the verb + clitics production latencies revealed a significant effect of frequency,  $F_1(1, 11) = 9.6, p = .01$ ;  $F_2(1, 38) = 8, p = .007$ . Participants were 54 ms faster with HF (867 ms) than LF nouns (921 ms). In addition, there was also a significant effect of repetition,  $F_1(2, 22) =$

---

<sup>6</sup> The high percentage of discarded data may be attributed to the greater difficulty of this experiment with respect to the others reported here (see below). In this experiment, participants were required to produce in the correct order two clitics in response to two different input stimuli: a picture for the object clitic, a written pronoun for the dative clitic. In the other experiments, participants were only required to produce one clitic.

TABLE 1

Results of Experiment 1. Mean reaction times(in ms) and standard deviations (in parentheses) for HF and LF pictures in each repetition of a picture stimulus (pronominal clitic production) and for HF and LF pictures (noun production)

<i>Picture</i>	<i>Pronominal clitic production</i>			<i>Noun production</i>
	<i>Presentation</i>			
	<i>I</i>	<i>II</i>	<i>III</i>	
<i>HF</i>	943 (326)	852 (286)	811 (319)	685 (160)
<i>LF</i>	1012 (360)	893 (318)	859 (348)	740 (164)

17.6,  $p < .0001$ ;  $F_2(2, 76) = 29.9$ ,  $p < .0001$ , but no interaction between the factors Repetition and Frequency,  $F_1(2, 22) = 3.21$ ,  $p = .06$ ;  $F_2(2, 76) = .01$ ,  $p = .9$ . The difference between HF and LF items is 69 ms in the first repetition, 41 ms in the second repetition, and 48 ms in the third repetition (Table 1).

Finally, error rates were significantly lower for HF (10.2%) than LF items (17.7%) in the verb + clitics production task,  $F_1(1, 11) = 11.5$ ,  $p = .001$ ;  $F_2(1, 38) = 7.4$ ,  $p = .01$ , but no difference was found in the simple picture-naming task (errors, HF = 4.2%, LF = 5%,  $p > .1$ ).

## Discussion

The results of this experiment are clear: Pronominal clitic production appears to be sensitive to the frequency of the referent noun. Participants produced simple verb + clitic utterances faster when the noun replaced by the pronominal clitic was of high frequency than when it was of low frequency. Furthermore, the effect was robust across three repetitions of the pronoun production task. We will refer to this effect as the noun-to-pronoun frequency inheritance effect. The demonstration that noun frequency affects the production latencies of the pronoun that replaces it implies that frequency has its effect at a stage of lexical access where the grammatical features of a word are represented (Caramazza, 1997; Dell, 1990).

Before any definitive claims can be made on the basis of the pronominal clitic production task used here, we have to make sure that the observed clitic frequency effect is a true lexical effect as opposed to an effect of picture recognition. That is, it could be argued that the observed frequency effect reflects differential speeds of recognition of the pictures such that low frequency nouns are recognised more slowly than high frequency nouns. This possibility is unlikely since the observed frequency effect persists pretty much unaltered over several presentations of the picture stimuli

both in pronoun production and name production (see Table 1). The persistence of a frequency effect in picture naming tasks has been interpreted as indicating that the locus of this effect is at the level of lexical access as opposed to picture recognition (e.g., Alario, Costa, & Caramazza, 2002; Jescheniak & Levelt, 1994; Levelt, Praamstra, Meyer, & Salmelin, 1998). Nonetheless, we can assess empirically whether variation in picture recognition latencies is correlated with word frequency. We did this by comparing recognition performance for the pictures of the HF and LF nouns in a word-picture matching task. Participants were asked to judge whether a word and a picture matched by pressing one of two buttons. The “yes” button was assigned to the dominant hand. The experimental pictures always matched the word (i.e., they all appeared as “yes” trial), whereas the filler pictures all appeared as “no” trial. Twelve students at the University of Palermo, aged 18–26, volunteered for the experiment. The 40 pictures used in Experiment 1 and 40 additional pictures of concrete objects were used. The latter pictures served only as filler trials and were thus removed from the statistical analysis. An additional set of 20 pictures was used as a practice block before the experiment proper. At the beginning of each trial, a word appeared on the computer screen (Arial 48, small caps) for 1000 ms; then, 200 ms after the word disappeared, a picture appeared on the screen. Pictures were removed as soon as participants responded, or after 2000 ms. The inter-trial interval was fixed at 1500 ms.

Recognition latencies were found to be independent of the frequency of the names of the pictured nouns, HF = 413 ms, LF = 409 ms;  $F_1(1, 11) < 1$ ;  $F_2(1, 38) < 1$ , and there was no difference in error rates (HF = 2%, LF = 2.5%,  $p > .1$ ). Thus, the frequency effect observed in Experiment 1 is unlikely to be caused by differences in recognition latencies between pictures with high and low frequency names, and we can more confidently ascribe the observed frequency effect to some stage of lexical access. It is interesting to note that the frequency effect reported here was obtained even though the critical pronoun appears as the final syllable in long, morphologically complex words. In principle, participants could have prepared the verb stem, and subsequently planned the two clitics while saying the preceding parts of the word. However, our finding is unsurprising given that there is independent evidence showing that the scope of phonological encoding involves at least a phonological word (e.g., Costa & Caramazza 2002; Wheeldon & Lahiri, 2002).

As already pointed out, the demonstration that latencies in producing pronouns are determined, in part, by the frequency of the nouns they replace provides clear evidence that the locus of the frequency effect is at a level of lexical access where a word’s grammatical properties are specified. In models that propose that only one lexical layer mediates between the semantic and phonological contents of words, the results reported here

simply indicate that the process of selecting lexical nodes is frequency-sensitive (Caramazza, 1997). In models that distinguish between lemma and lexeme levels of lexical representation, the results indicate that the locus of the frequency effect is at the level of lemma representations—where grammatical properties are specified (Dell, 1990; Griffin & Bock, 1998)—and not at the level of lexeme representation (Jescheniak & Levelt, 1994; Levelt et al., 1999).

The conclusion reached here is based on the assumption that pronoun selection implies the prior selection of the controlling noun's grammatical features. The assumption is that only the grammatical features must be retrieved during pronoun production (see Jescheniak & Schriefers, 2001; Jescheniak et al., 2001; Meyer & Bock, 1999). Phonological features may be activated as well, but this is not a precondition for pronoun production. There are results which suggest that a noun's phonology is activated in the course of pronoun production (Schmitt, Meyer, & Levelt, 1999) while other results indicate that, whatever activation there might be of the noun's phonology, is very weak (Jescheniak et al., 2001). In our experiment, the noun's phonological content does not play a causal role in pronoun selection and therefore whether or not it was activated in the course of producing pronominal clitics is not consequential. What is important is that the level of lexical representation where a word's grammatical properties are specified seems to be the level where frequency has its effect in lexical access.

Having established that the pronominal clitic production task is sensitive to lexical access processes at the level where a word's grammatical properties are specified, the task can be used to address issues concerning the retrieval of grammatical features. We turn to these issues next.

## EXPERIMENT 2. The retrieval of grammatical gender in the production of pronominal clitics: Evidence from the picture-word interference paradigm

In the Introduction we distinguished between two views of grammatical feature retrieval. In one view, the retrieval of grammatical features is a process that is in some respects similar to the selection of lexical nodes (Levelt, 2001; Schriefers, 1993). The main similarity is that both processes are considered to be “independent” events that are decided strictly on the basis of the activation level of lexical or grammatical feature nodes, as the case may be. Of course, the retrieval of grammatical features is driven by the prior selection of a specific lexical node, but what is important for present purposes is that the selection of a given lexical node only sets the

conditions for the subsequent “independent” selection of its associated grammatical feature nodes. Because on this view grammatical feature nodes are selected on the basis of their levels of activation, the process can be speeded up by priming (Levelt, 2001) or slowed down by competition from other grammatical feature nodes (Schriefers, 1993). If, however, grammatical feature retrieval is an automatic consequence of lexical node selection, it would not constitute an independent event from lexical node selection and thus would not be subject to independent priming (Caramazza et al., 2001a).

There is a growing literature that speaks to the issue of grammatical feature retrieval. Two types of experimental paradigms have been used to address this issue. One approach has focused on demonstrating that the gender of a distractor word in a picture-word interference naming task affects the retrieval of the gender of the target word as revealed through the production of gender-marked words, such as determiners in Dutch (La Heij et al., 1998; Schriefers, 1993; Schriefers & Jescheniak, 1999). The other approach has focused on demonstrating that the gender node of the target word can be primed by its repeated activation and selection (Jescheniak & Levelt, 1994; Vigliocco, Lauer, Damian, & Levelt, 2002).

In a seminal paper, Schriefers (1993) extended the classical picture-word interference naming paradigm to investigate the process of grammatical feature selection. In the picture-word interference naming paradigm, participants are asked to name a picture while ignoring the presentation of a distractor word (e.g., Damian & Martin, 1999; Glaser & Glaser, 1989; La Heij, 1988; Lupker, 1979; Meyer & Schriefers, 1991; Miozzo & Caramazza, 2003; Schriefers, Meyer, & Levelt, 1990). It has been repeatedly shown that picture naming latencies are affected by properties of the distractor word (e.g., Miozzo & Caramazza, 2003) as well as by the type of relationship between the picture name and the distractor word. Thus, for example, it has been shown that a semantic (e.g., Caramazza & Costa, 2000; La Heij, 1988; Lupker, 1979) or phonological relationship (e.g., Damian & Martin, 1999; Meyer & Schriefers, 1991) between picture and distractor word affects the time it takes to name the picture. Schriefers (1993) extended the paradigm to investigate the effects of gender congruency on picture naming—that is, whether the relationship between the gender of the picture name and that of the distractor word would affect picture naming performance.

In Schriefers’ experiments, Dutch speakers were asked to produce noun phrases (NPs) (e.g., “de rode auto”—the red car) in response to coloured pictures with superimposed words that they were required to ignore. Dutch is a gender-marked language in which the form of the determiner and the inflection of the adjective in an NP depend on the gender value of the noun (e.g., common gender: “de rode auto”—the red car; neuter gender: “het

rode boek”—the red book).<sup>7</sup> This property of gender-marked languages allows comparison of NP production latencies for the case where the genders of the picture name and distractor word are the same (gender congruent) versus the case where they are different (gender incongruent). Schriefers (1993) found a clear congruency effect: participants produced NPs faster in the gender congruent than the gender incongruent condition. He interpreted this result as reflecting competition between activated gender nodes in the incongruent condition. This interpretation of the gender congruency effect is based on two assumptions. One assumption is that the selection of gender nodes depends on their level of activation. We will refer to this assumption as the “selection-by-activation-level” assumption. The other assumption is that the selection process is sensitive to the relative activation levels of all activated gender nodes, such that the selection decision is more difficult when the activation levels of the gender nodes are more similar. We will refer to this assumption as the “selection-by-competition” assumption. These two assumptions are exactly the same as those that have been made for lexical node selection in various models of lexical access (e.g., Roelofs, 1992; Starreveld & La Heij, 1996).

The gender congruency effect has been replicated in many experiments in various languages (*Dutch*: La Heij et al., 1998; Schiller & Caramazza, 2003; Schriefers, 1993; *German*: Schiller & Caramazza, 2003; Schriefers & Teruel, 2000; *Croatian*: Costa, Kovacic, Federenko, & Caramazza, 2003). The effect is most robust in tasks involving the production of determiner NPs.<sup>8</sup> The gender congruency effect appears to be language specific: it has been found for Germanic and Slavic languages but not for Romance languages (*Italian*: Miozzo & Caramazza, 1999; Miozzo, Costa, &

---

<sup>7</sup> Actually, the gender marking of adjectives in an NP only surfaces when determiners are absent. Thus compare “de rode auto” and “het rode boek” versus “rode auto” and “rood boek”.

<sup>8</sup> The gender congruency effect is only obtained when participants must produce gender-marked words/morphemes. Thus, for example, La Heij et al. (1998) obtained a gender congruency effect in a picture-word interference production task when Dutch subjects were required to produce det + noun phrases. When participants were required to produce only the noun, no gender congruency effect was observed. This result has been interpreted in two ways. On one interpretation, the absence of a gender congruency effect indicates that gender is not selected when bare nouns must be produced. Since the gender feature is not selected in such cases, there is no opportunity for a congruency effect (Levelt et al., 1999). An alternative interpretation explains the absence of a gender congruency effect in bare noun naming as indicating that the effect is only obtained when gender-marked phonological forms (e.g., determiners) compete for selection. And since participants are not required to produce gender-marked forms there is no opportunity for a gender congruency effect. Miozzo and Caramazza (unpublished raw data) also failed to observe a gender congruency effect when Italian participants were required to produce bare nouns in a picture-word interference naming task.

Caramazza, 2002; *Spanish*: Costa, Sebastian-Galles, Miozzo, & Caramazza, 1999; Miozzo et al., 2002; *Catalan*: Costa et al., 1999; *French*: Alario & Caramazza, 2002). Caramazza et al. (2001b) (see also Costa et al., 1999; Miozzo & Caramazza, 1999) have proposed that the relevant distinction between the languages that show a congruency effect and languages that do not show such an effect is whether determiner selection depends only on the grammatical features of the NP (“early selection” languages) or whether it also depends on the phonological context in which the determiner is to be produced (“late selection” languages). In late selection languages such as Italian, the forms of some determiners depend not only on grammatical properties such as gender and number, but also on the onset of the word that follows the determiner. For example, the singular masculine definite determiner is *il* (as in “il tavolo”—the table) for all contexts except when the following word begins with a vowel, an affricate, a consonant cluster of the form “s + consonant” or “gn”, in which case it is *lo* (e.g., “lo scoiattolo”—the squirrel, “lo strano tavolo”—the strange table, etc.). Since the production of an article must await the availability of local phonological context before its form can be determined, it is possible that any effect of gender congruency occurring at the level of grammatical feature selection would not be visible at the point in time when the form of the determiner is selected for production.

However, there is another interpretation that is possible for the contrasting results between early and late selection languages. Perhaps the reason that late selection languages do not show a gender congruency effect is because gender features do *not* compete for selection—instead they become available automatically as part of lexical node selection. If there is any competition at all in lexical access (see Miozzo and Caramazza, 2003, for arguments against the assumption of “selection-by-competition” of lexical nodes) it is restricted to competition between phonological forms of lexical nodes. But how, then, do we explain the gender congruency effects observed in Dutch and German? One possibility is that the so-called gender congruency effect is really a determiner congruency effect. Thus, for example, when a Dutch speaker is attempting to produce the NP “de auto” (the car) in response to a picture of a car with the word “boek” printed on the picture, the selection of the form *de* occurs in the context of a “competing” form *het* (the determiner associated with the distractor word “boek”), which interferes with the production of the target determiner form. The reason such determiner interference is not found in Romance languages is that by the time a determiner form can be selected the activation of the distractor lexical node would have dissipated along with the activation of its associated gender feature, leaving little opportunity for significant activation of competing determiner forms.

Schiller and Caramazza (2003) reported a set of experiments in Dutch and German that show that the gender congruency effect reflects competition in the selection of the determiner of the target NP rather than competition in the retrieval of the target's gender feature. That is, the interference produced by a distractor word of a different gender stems from the fact that the distractor word and the target call for different determiners, and not from the fact that they have different genders. Participants were asked to produce determiner + noun NPs either in the singular (where nouns of different genders take different determiners: in Dutch “*de tafel*”—the table, “*het boek*”—the book) or in the plural (where the determiner is always the same: “*de tafels*”—the tables, “*de boeken*”—the books), while ignoring gender (in)congruent distractor words. Gender congruency effects were observed for singular but not for plural NPs, supporting the notion that gender congruency effects reveal differences in the availability of determiner forms rather than differences in the availability of the noun's gender feature (for related evidence see Janssen & Caramazza, 2003; Schriefers, Jescheniak, & Hantsch, 2002). The reasoning behind this conclusion is as follows: If the gender congruency effect were at the stage of gender retrieval it should not depend on the form of the determiners for singular and plural NPs. Thus, the gender congruency effects observed in NP production cannot be used to support the notion that grammatical features are accessed through a process of selection-by-activation-level. If anything, the absence of a gender congruency effect for plural NPs observed by Schiller and Caramazza suggests that gender retrieval occurs automatically as a consequence of lexical node selection.

The second set of studies related to gender retrieval explored whether access to a noun's gender facilitates its retrieval in a subsequent trial, the so-called gender recency (or priming) effect (Jescheniak & Levelt, 1994; Vigliocco, Lauer, Damian, Levelt, 2002). Jescheniak and Levelt (1994) asked Dutch participants to perform a gender decision task upon the presentation of a picture: Participants pressed one button if the picture's name was a “de-gender” word and another button if it was a “het-gender” word. Before the gender decision task, participants were asked to name the whole set of pictures. One group of participants was asked to name the pictures using only nouns (e.g., “auto”—car), and another group of participants was asked to name them using gender-marked NPs (e.g., “*de auto*”—the car). The authors argued that while NP naming requires the retrieval of the noun's gender in order to select the gender-marked determiner form, noun naming does not require gender retrieval (see also La Heij et al., 1998; but see general discussion). Thus, if gender features can be primed, one would expect faster decisions when the gender of those words had recently been accessed (NP naming) than when the naming task

did not require gender access (noun naming). The authors found faster gender decision times in the NP condition than in the noun naming condition.

However, follow-up research on gender priming has failed to replicate such an effect. Van Berkum (1996, cited in van Berkum, 1997) failed to replicate the gender priming effect in the same language (Dutch), and with the same experimental task (gender decision). He also failed to find a gender priming effect in a task in which Dutch speakers named coloured pictures by producing gender-marked (colour adjective + noun) NPs (van Berkum, 1997). Experimental trials could be preceded by trials in which participants produced gender-marked NPs or nouns. Naming latencies of gender-marked NPs were independent of whether they were preceded by gender-marked NPs or bare nouns.

Jescheniak and Schriefers (1999) also failed to observe gender priming in a pronoun production task with German speakers. In that study, participants were asked to produce the gender-marked pronoun (e.g., “*er*”) corresponding to the name of a picture (e.g., “*Besen*” –broom). Experimental trials were immediately preceded by trials in which participants named the same picture using a gender-marked (determiner + noun) NP or a bare noun. Jescheniak and Schriefers (1999) found that pronoun naming latencies were independent of whether participants had previously produced an NP with a gender-marked article or a bare noun. These results and those reported by van Berkum (1997) question the reliability of the gender priming effect reported by Jescheniak and Levelt (1994).

Vigliocco et al.’s (2002) used a different paradigm in trying to show that gender selection can be primed. In their study, Dutch participants were asked to translate words from English (prompt words) into Dutch (response words). The English prompt words could appear in lower case or upper case and participants had to produce the correct Dutch noun along with gender-marked adjectives: “Small” if the word was presented in lower case (“*klein*” or “*kleine*”, for neuter and common gender nouns, respectively) or “big” if the word appeared in upper case (“*groot*” or “*grote*”, for neuter and common gender nouns, respectively). The prompt words were presented in two different sets: gender homogeneous sets (where all the response words had the same gender), or gender heterogeneous sets (where half of the words were of one gender and half of the other gender). Since the adjectives “small” and “big” in Dutch are gender-marked, participants re-accessed the same gender feature repeatedly in the homogeneous sets, while no such massive re-access of a gender feature was present in the heterogeneous sets. The authors found that translation times in the homogeneous sets were faster than in the heterogeneous sets, and interpreted the result as indicating that access to the noun’s grammatical gender can be primed.

However, there are two aspects of the experimental design employed by Vigliocco et al. (2002) that complicate the interpretation of the study. In their experiments, the gender homogeneous and gender heterogeneous sets differed systematically not only in the way gender was distributed in each set, but also in other related variables: the type of morphological transformation carried out on the adjectives in each set and the size of the response set. In the homogeneous sets, participants always produced *two* adjective forms—they either produced “klein” and “groot”, for neuter gender homogeneous sets, or “kleine” and “grote”, for common gender homogeneous sets. However, in the heterogeneous sets they were required to produce *four* adjective forms (“klein”, “kleine”, “groot”, and “grote”). Furthermore, in a given homogeneous set the adjectives being produced always took the same inflection (-*e* for common gender sets, and - $\emptyset$  for neuter gender sets), while in a heterogeneous set participants had to make use of two different morphological transformations (-*e* and - $\emptyset$ ). Thus, it is possible that the cause of the difference between homogeneous and heterogeneous sets is not due to facilitation in retrieving gender features, but to differences in response set size and/or in the type of morphological transformations required.<sup>9</sup>

In short, this study cannot be taken as providing evidence for gender priming, and given the repeated failure to observe reliable gender priming in various other experimental paradigms (van Berkum, 1996, 1997; Jescheniak & Schriefers, 1999), it is important to find alternative ways to determine whether grammatical features are retrieved through a process of selection-by-activation-level or whether they become available automatically as a consequence of selecting the lexical node through which those features are accessed. We addressed this issue by adapting the picture-word interference naming paradigm to pronominal clitic production.

In the picture-word interference paradigm, participants are required to name a picture while ignoring a distractor word that is printed on or in the vicinity of the picture. In the task used here, participants do *not* have to name the picture (e.g., “casco”—helmet) but instead must pronominalise it into an accusative pronominal clitic (e.g., *-lo*) attached to a previously given verb (e.g., “portare”—to bring), resulting in a form such as “portalo”—bring it. In order to produce the correct pronominal clitic, participants must access the gender of the pictured noun. The crucial question is: Does the gender of the distractor word affect the retrieval of

---

<sup>9</sup> In two unpublished studies, one with Italian participants (Costa & Caramazza, 2005) and one with Dutch speakers (Perdijk, Spalek, & Schriefers, 2005), that used the same experimental paradigm as Vigliocco et al. (2002) but controlled for response set size, no gender priming effects were observed.

the pictured noun's gender? More specifically, do we find a gender congruency effect in the production of pronominal clitics? If grammatical features are retrieved through a process of selection-by-activation-level we should find a gender congruency effect either because congruent features prime each other, resulting in facilitation (e.g., Levelt, 2001) or because incongruent features compete for selection, resulting in interference relative to the congruent condition (e.g., Schriefers, 1993). In either case, the expectation is that we should find a gender congruency effect. However, if gender feature retrieval follows automatically from lexical node selection we should not find a gender congruency effect in the picture-word interference task with pronominal clitics (e.g., Caramazza et al., 2001b).

Before proceeding with a description of Experiment 2, we must address a potential problem that derives from the use of Italian to investigate issues of grammatical gender selection. Earlier we noted that in Italian the form of definite determiners depends not only on the grammatical features "definiteness", "number", and "gender" but also on the determiner's local phonological context. As a consequence, the selection of a determiner cannot proceed until the segmental content of the word that follows it has been specified (for discussion see Alario & Caramazza, 2002). We also noted that this property of Italian determiners could make an effect of gender congruency invisible even if such an effect were to exist. However, unlike determiners, the selection of pronominal clitics does not depend on phonological context: Once the relevant grammatical features have been "collected", the form of the pronominal clitic is fully specified and can be selected for production. In other words, pronominal clitics can be selected as "early" in Italian as determiners are selected in German and Dutch, thereby making them appropriate to evaluate issues concerning grammatical feature selection.

In this experiment we used a verb + clitic production task in the context of a picture-word interference paradigm. Participants were not required to name the pictured objects but to pronominalise their names as accusative clitics with a given verb (e.g., "portalo"—bring it). Two factors were manipulated: the frequency of the picture names, thereby allowing a replication of Experiment 1, and the genders of the picture names and distractor words.

## Method

*Participants.* Sixteen native speakers of Italian, students at the University of Pisa (age: 20–37) took part in the experiment. They were paid for their participation.

*Materials.* We selected 40 words to serve as distractors (see Appendix B). These distractor words (Arial 32, small caps) were superimposed on the 40 pictures that were previously used in Experiment 1. The distractor words were divided into groups of four. Each group was paired with four pictures (HF fem., HF mas., LF fem., LF mas.). Within each distractor group, distractors were of the same length in letters (mean length: 5.4) and number of syllables (mean number of syllable: 2.1), but differed with respect to grammatical gender (mas. or fem.). Thus, for each group, we had two masculine distractors (mean frequency of mas. distractors: 54.8, range: 0–163), and two feminine distractors (mean frequency of fem. distractors: 54.4, range: 0–300).

Each picture appeared four times, each time with a different distractor (twice with a gender-congruent distractor, twice with a gender-incongruent distractor). Distractors were semantically and phonologically unrelated to the pictures' names. Four blocks were constructed. Each picture appeared only once in each block. Within each block, trial presentation order was pseudo-randomised, such that: (a) pictures with the same gender or within the same frequency range were not to appear for more than three consecutive trials; (b) distractors with the same (in)congruency relation with the target were not to appear for more than three consecutive trials; (c) there was no semantic, phonological, or associative relation between pictures or distractors on consecutive trials.

This experiment differs from Experiment 1 in that participants were only required to produce accusative clitics (e.g., “portalo”), considerably simplifying both the stimuli (no need to have a PP to indicate a dative pronoun) and the response (only one clitic needed to be produced). Two additional sets of verb-picture pairings were used. One set ( $N = 40$ ) was used as practice stimuli before the experiment; the other set ( $N = 8$ ) was used as warm-up stimuli at the beginning of each block.

## Procedure

The procedure was the same as in the previous experiment. Participants first named all the pictures presented with a set of Xs printed on the picture at the location where the distractor word would later appear. The location of the Xs and subsequently the distractors varied in the vicinity of the fixation point. Participants then took part in a long practice block in which they performed the same task as the one used in the experiment proper. After that, the experiment proper began. At the beginning of each trial, a question mark appeared in the centre of the computer screen. The question mark disappeared as soon as participants pressed the space bar. Then, the verb appeared in the centre of the computer screen for 1000 ms and was immediately replaced by the picture with a given distractor.

Pictures were shown for unlimited exposure and were removed as soon as participants responded or 3000 ms had elapsed. Stimulus presentation, RT recording, and error checking were controlled as in the previous experiment.

## Results

In the statistical analyses, the variables Picture Name Frequency and Repetition were treated as within-subject variables with one exception: In  $F_2$  analyses, picture frequency was considered a between-subject variable. The same analyses were repeated with participants' error rates as the dependent measure. The variable Distractor Gender (congruent vs. incongruent) was always treated as within-subject, in both  $F_1$  and  $F_2$ .

The same exclusionary criteria as in Experiment I were used in this experiment. Discarded data accounted for 6.7% of the trials, a considerably lower error rate than in Experiment 1. The effect of frequency on naming latencies across the four repetitions was significant in the subject analysis, HF = 580 vs. LF = 593;  $F_1(1, 15) = 15, p = .001$ , and very close to significance in the item analysis,  $F_2(1, 38) = 3.1, p = .08$ , see Table 2). In addition, we found a significant effect of repetition,  $F_1(3, 45) = 37, p < .0001$ ;  $F_2(3, 114) = 223, p < .0001$ , but no interaction between frequency and repetition,  $F_1(3, 45) = 2.2, p = .1$ ;  $F_2(3, 114) = 1, p = .4$ . RTs were not affected by distractor type. Responses to pictures with same-gender distractors were as fast as responses to pictures with different gender distractors (586 ms vs. 587 ms; both  $F_s < 1$ ).

Error analyses revealed a tendency for LF pictures to be more susceptible to errors than HF pictures (errors, HF = 5.2%, LF = 8.2%), but the difference was significant only in the by-subjects analysis,  $F_1 = (1,$

TABLE 2

Results of Experiment 2. Mean reaction times (in ms) and standard deviations (in parentheses) for each repetition of HF and LF pictures and for congruent and incongruent distractors

		<i>Pronominal clitic production</i>				
		<i>Presentation</i>				
<i>Picture</i>	<i>Distractor</i>	<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>TOT</i>
<i>HF</i>	<i>Gender Congr</i>	652 (200)	597 (196)	554 (170)	522 (153)	577 (181)
	<i>Gender Incong</i>	677 (198)	593 (172)	537 (144)	516 (150)	582 (180)
	<i>TOT</i>	665 (199)	595 (184)	546 (157)	519 (151)	
<i>LF</i>	<i>Gender Congr</i>	679 (204)	613 (187)	570 (161)	522 (145)	595 (188)
	<i>Gender Incong</i>	670 (186)	627 (192)	556 (168)	518 (147)	592 (181)
	<i>TOT</i>	674 (194)	620 (189)	563 (164)	520 (146)	

15) = 6.6,  $p = .02$ ;  $F_2(1, 38) = 1.8$ ,  $p = .2$ ). No effect Distractor Gender emerged from the error analyses (both gender congruent and gender incongruent distractors = 6.7%).

## Discussion

The two main results of this experiment are as follows. First, we replicated the noun-to-pronoun frequency inheritance effect in pronominal clitic production, but this time in the context of a picture-word interference naming task. Participants produced verb + clitic utterances faster in response to pictures with high frequency names than to pictures with low frequency names. The effect was robust over several repetitions of the picture stimuli in the by-subjects analysis and marginally significant in the by-items analysis. Furthermore, the frequency effect was obtained in a simpler production task than was used in Experiment 1. In that experiment participants were required to produce a verb plus dative and accusative clitics (e.g., “portamelo”—bring it [mas.] to me) whereas in this experiment they were required to produce only a verb plus accusative clitic (e.g., “portalo”—bring it [mas.]). And since we have demonstrated that the observed frequency effect is not due to recognition differences for the pictures with high and low frequency names (see Discussion of Experiment 1), the effects obtained here attest to the robustness of the frequency effect in the production of pronominal clitics.

The second result obtained in this experiment is the absence of a gender congruency effect in a pronominal clitic production task. Participants produced verb + clitic utterances equally fast for picture-distractor pairs with same and different genders. This result indicates that gender selection is not a competitive process, and is not subject to priming by same-gender distractors. Instead, the results are consistent with the proposal that grammatical feature retrieval is an automatic consequence of lexical node selection and therefore not subject to independent priming. Although the results of Experiments 1 and 2 show that the pronominal clitic production task used here is sensitive enough to reveal the nature of processes occurring at the level of lexical selection where a word’s grammatical features are represented, the interpretation of null results should proceed cautiously. Thus, before concluding that gender congruency does not affect the selection of gender features, it is important to replicate the result. Furthermore, it is important to demonstrate that distractor words can affect the selection of pronominal clitics in a picture-word interference naming task. In other words, we should document the absence of a gender congruency effect in the context of an experiment where distractor words are shown to have an effect on pronominal clitic production. Experiment 3 was designed for this purpose.

**EXPERIMENT 3. The retrieval of grammatical gender in the production of pronominal clitics: Effects of semantic and phonological relatedness in a picture-word interference paradigm**

The main objective of this experiment is to verify the absence of a gender congruency effect in a picture-word interference naming experiment where we have positive evidence that the distractor word does affect pronominal clitic production. We did this (Experiment 3A) by testing for the presence of a semantic (co-ordinate) interference effect in pronominal clitic production.<sup>10</sup> On the assumption that the semantic interference effect reveals processes occurring at the level where semantic information is used to access lexical nodes, the presence of such an effect in the pronominal clitic production task would guarantee both that the distractor word has been processed and that the task is sensitive enough to reveal interference/facilitation effects. In other words, the assumption is that semantically (categorically) related distractors will delay the selection of the lexical representation of the noun whose gender feature is needed for pronominal clitic selection, thereby delaying the production of the verb + clitic utterance. We also carried out an experiment (Experiment 3B) designed to assess whether phonologically related distractors affect pronominal clitic production.

We noted above that the selection of Italian determiners is possible only after the phonological onset of the following word has already been specified. We also claimed that pronominal clitics are different from determiners in this respect, as they do not depend on the phonological context they appear in. As a consequence, we assumed that pronominal

---

<sup>10</sup> Although the existence of the semantic (co-ordinate) interference effect is beyond dispute (Glaser & Glaser, 1989; Glaser & Döngelhoff, 1984; La Heij, 1988; Lupker, 1979, Rosinski, 1977), the cause of the effect remains controversial. One interpretation (e.g., Roelofs, 1992; Starreveld & La Heij, 1996) assumes that it reflects the competitive nature of the lexical selection process. The claim is that the selection process is sensitive to the *relative* activation level of the target response: the smaller the difference in activation levels between the target and distractor lexical nodes the harder the selection decision. Given this assumption, the semantic interference effect arises because the activation level of the lexical node of a semantically related distractor is assumed to be greater than that of an unrelated distractor. This difference presumably reflects the fact that semantically related distractors receive activation not only from the word stimulus but also from the semantically related target picture. Another interpretation (Costa, Alario, & Caramazza, in press; Miozzo & Caramazza, 2003) is that the semantic (coordinate) interference effect has its locus at the level of deciding which semantic representation—the target picture's or the distractor word's—should be selected for lexicalisation. However, for present purposes, the value of the semantic interference effect is as a signature effect that the distractor word has been processed and its semantic representation has been retrieved.

clitics can be selected as “early” as the determiners of non-Romance languages. On this basis, we interpreted the null effect of gender congruency as an indication of the automatic activation of the gender feature. However, one could reason that the production system of a given language always operates in the same mode, irrespective of whether this mode is actually needed for the specific construction being planned. In the case of Italian, the “late selection mode” would be induced by the fact that the phonology of the following word is needed for the selection of many gender marked elements. On this perspective, Italian speakers would wait for the phonological context also when they have to produce clitics. In other words, clitics would also be selected late. Note, however, that the type of clitics used in our research, which attach at the end of the utterance, exclude the possibility that there is some subsequent phonological context for which to wait. Thus the idea that clitic selection in our experiments is made late by default is not well founded.<sup>11</sup>

Alternatively, one could argue that although the phonological content of the word that follows a clitic is not relevant for clitic selection, the phonological content of the word replaced by the clitic *is* relevant. In other words, pronominal clitic selection only occurs after the selection of the phonological content of its controlling noun. If such were the case, the phonology of the pictured noun would be causally involved in the selection of its clitic. On this account, we should observe a phonological effect in Experiment 3B. However, if the noun’s phonology is not causally implicated in the selection of the pronominal clitic, the expectation is that phonologically related distractor words should not affect the production of verb + clitic utterances.

Finally, note that a test of phonological effects in clitic production is also directly relevant for the localisation of the frequency effect in lexical access. Recall that we interpreted the frequency inheritance effect observed in Experiments 1 and 2 as evidence against the localisation of frequency at the lexeme level. However, we cannot exclude that the frequency effect in pronominal clitic production is due to the fact that participants implicitly generate phonologically spelled out Det + NPs, and use this implicit utterance format to derive the correct clitic pronoun. As a consequence, the frequency inheritance effect could be attributed to this implicit production of a word’s lexeme representation. However, if the implicit production of the controlling noun were responsible for the

---

<sup>11</sup> We are currently exploring the possibility that clitic selection occurs early by investigating the production of proclitic forms such as “lo porta” (he brings it [mas.], literally: it [mas.] brings [second person, singular]). These proclitics are “attached” to the beginning of the verb.

observed frequency effect in clitic production we would then expect a phonological facilitation effect in the retrieval of the noun's phonology and, by extension, in the production of pronominal clitics. On the other hand, if we are able to demonstrate that the controlling noun's phonology does not affect the production of the corresponding clitic, an interpretation along the lines outlined above would be excluded.

To sum up, Experiment 3B is relevant to the issue of whether pronominal clitics are selected early in Italian and to the issue of whether the frequency effect in clitic production might not after all have its locus at the lexeme level.

## Method

*Participants.* Twenty-four students at the University of Pisa, aged 18–30, took part in the experiment—12 each in Experiments 3A and 3B. They were paid for their participation.

*Materials.* Twenty-six new black-and-white pictures were selected. They represented animals, fruit and vegetables, and common artifacts. Half of the picture names were masculine and half were feminine ( $N = 13$ ). The two groups were matched for frequency (fem., mean frequency: 37.5, range: 0–206; mas., mean frequency: 57.2, range: 0–176), length (fem.: 6.3 letters; mas.: 6.2 letters), and number of syllables (fem. and mas.: 2.5; see Appendices D, E). We also selected 13 verbs, mostly the same as those used in the previous experiments. Each verb was paired with two different pictures (1 fem., 1 mas.). The same pictures were used for both Experiments 3A (manipulation of distractor gender and distractor semantics) and 3B (manipulation of distractor gender and distractor phonology).

In each experiment, a given picture was paired with four different distractor words. Relative to the picture name, distractors could be: (a) semantically (Exp. 3A) or phonologically (Exp. 3B) related and gender congruent; (b) semantically (Exp. 3A) or phonologically (Exp. 3B) related and gender incongruent; (c) semantically (Exp. 3A) or phonologically (Exp. 3B) unrelated and gender congruent; (d) semantically (Exp. 3A) or phonologically (Exp. 3B) unrelated and gender incongruent (see Appendices C, D).

Distractors paired with a given picture had comparable mean frequency and frequency range values (Exp. 3A: semantically related distractors, mas. = 17.8, 0–112; semantically unrelated distractors, mas. = 18.2, 0–115; semantically related distractors, fem. = 17; 0–122; semantically unrelated distractors, fem. = 18.8, 0–143. Exp. 3B: phonologically related distractors, mas. = 11.7, 0–32; phonologically unrelated distractors, mas. = 9.7, 0–28;

phonologically related distractors, fem. = 13.1; 0–48; phonologically unrelated distractors, fem. = 14.5, 0–82) and the same length (mean length of semantically related and unrelated distractors: 6.3; mean length of phonologically related and unrelated distractors: 5.5). The distractors of a given picture were also matched for number of syllables (Exp. 3A: semantically related and unrelated, mas. = 2.7; semantically related and unrelated, fem. = 2.5. Exp. 3B: phonologically related, mas. = 2.3; phonologically unrelated, mas. = 2.4; phonologically related and unrelated, fem. = 2.3).

Experiments 3A and 3B were each divided into 4 blocks. A given picture appeared once per block, each time paired with a different distractor. Distractor types as well as pictures with masculine and feminine names were uniformly distributed across blocks. Presentation order was pseudo-randomised, with the constraints that: (a) Pictures with the same gender and/or with the same gender relation (congruent/incongruent) were not to appear for more than three consecutive trials; (b) there was no phonological, semantic, or associative relation between pictures or distractors on consecutive trials.

Two additional sets of verb-picture pairs were selected. They were used respectively in a practice block before the experiment proper ( $N = 40$ ), and as filler trials at the beginning of each block ( $N = 8$ ). These two sets of verb-picture stimuli were identical for Experiments 3A and 3B.

*Procedure.* The procedure was the same as in Experiment 2. As in that experiment, participants were asked to produce the imperative of a given verb with the accusative clitic corresponding to the pictured noun (e.g., “portalo” (bring it) for bring the helmet).

## Results

Discarded data accounted for 9% of the trials, in both Experiments 3A and 3B.

The results of Experiment 3A (see Table 3) confirmed the absence of a gender congruency effect, both in the analysis of RTs (gender-congruent distractors: 550 ms; gender -incongruent distractors: 548 ms; both  $F_1$  and  $F_2 < 1$ ) and errors (gender congruent distractors = 7.2%, gender incongruent distractors = 8.6%,  $p > .1$ ). By contrast, the analysis of RTs showed a clear semantic interference effect. Distractors belonging to the same semantic category as the target picture interfered more than their corresponding controls (557 ms vs. 541 ms;  $F_1(1, 11) = 13.5, p = .004$ ;  $F_2(1, 25) = 6.5, p = .02$ ). There was no interaction between semantic relatedness and gender congruency,  $F_1(1, 11) < 1$ ;  $F_2(1, 25) < 1$ .

TABLE 3A

Results of Experiment 3A. Mean reaction times (in ms), standard deviations (in parentheses), and error rates as a function of semantic relatedness and gender congruency

	<i>Gender congruent</i>		<i>Gender incongruent</i>	
	<i>RTs (SD)</i>	<i>Error rate (%)</i>	<i>RTs (SD)</i>	<i>Error rate (%)</i>
Semantically related	560 (156)	9.9	555 (151)	10.9
Semantically unrelated	541 (139)	8	541 (139)	9

TABLE 3B

Results of Experiment 3B. Mean reaction times (in ms), standard deviations (in parentheses), and error rates as a function of phonological relatedness and gender congruency

	<i>Gender congruent</i>		<i>Gender incongruent</i>	
	<i>RTs (SD)</i>	<i>Error rate (%)</i>	<i>RTs (SD)</i>	<i>Error rate (%)</i>
Phonologically related	590 (181)	10.9	597 (176)	8
Phonologically unrelated	599 (183)	7	592 (178)	9.9

Analysis of the errors in Experiment 3A revealed a statistically not significant tendency for semantically related distractor trials to be more prone to errors than semantically unrelated distractor trials, 9.3% vs. 7.2%;  $F_1(1, 11) = 2$ ;  $F_2(1, 51) = 1.9$ ; in both cases,  $p > .1$ .

The RT results in Experiment 3B (see Table 3) replicated the absence of a gender congruency effect in pronominal clitic production (gender-congruent distractors: 594 ms; gender-incongruent distractors: 595 ms; both  $F_1$  and  $F_2 < 1$ ). The results also showed that phonologically related distractor words did not affect pronominal clitic production (phonologically related distractors: 594 ms; phonologically unrelated distractors: 596 ms; both  $F_1$  and  $F_2 < 1$ ). Furthermore, the two variables did not interact (both  $F_1$  and  $F_2 < 1$ ).

Analyses of error rates across experimental conditions did not reveal any significant effect: phonologically related distractors = 8.3%, phonologically unrelated distractors = 7.2%,  $p > .1$ ; gender congruent distractors = 7.7%, gender incongruent distractors = 7.9%,  $p > .1$ .

## Discussion

The results of Experiments 3A and 3B replicated the clear absence of a gender congruency effect in pronominal clitic production. The replication was obtained with new stimuli and with two new groups of participants.

More importantly, the absence of a congruency effect was obtained in the context of a theoretically plausible pattern of semantic and phonological distractor effects. The results of Experiment 3A showed that the semantic (categorical) relatedness between a distractor word and a picture's name affects latencies in the production of pronominal clitics. This result is consistent with expectations derived from the assumption that selection of the target's lexical representation, at the level where its gender feature is specified, plays a causal role in pronominal clitic production. The latter result is also important in that it demonstrates that the pronominal clitic production task can reveal effects of distractor words when they are present, and as expected on the basis of our analysis of pronominal clitic production. Furthermore, the results of Experiment 3B showed that phonological relatedness between a distractor word and a picture's name does not affect pronominal clitic production, precisely as expected by the assumption that the phonology of the target noun does not play a causal role in pronominal clitic selection.

The pattern of results we have obtained with phonologically and semantically related distractors is in accord with results obtained by Jescheniak and Schriefers (2001) for pronoun production in German. These authors found that semantically related distractors interfere with both noun and pronoun production, but that phonologically related distractors facilitate noun production while having no effect on pronoun production. Working within a theoretical framework that assumes two lexical layers mediating between a word's semantic and phonological content, they interpreted this finding as evidence that pronoun production implicates access of the noun's lemma but not the noun's lexeme.

Although the absence of a phonological facilitation effect in pronominal clitic production is consistent with theoretical expectations and with previous results reported by Jescheniak and Schriefers (2001) in German, we should nonetheless make sure that its absence in our experiment is not merely the consequence of poor stimulus selection. That is, we should make sure that the distractor words used in the phonological condition (Experiment 3B) would have led to phonological facilitation had we required participants to produce the name of the picture. We tested this possibility by having 12 students at the University of Pisa name the picture stimuli with the superimposed same-gender words that we used in Experiments 3B. Participants were 51 ms faster at naming the pictures when distractors were phonologically related (746 ms) to the target noun than when they were unrelated, 796 ms;  $F_1(1, 11) = 17.3, p = .002$ ;  $F_2(1, 25) = 11.8, p = .002$ ). Errors were distributed equally across the two conditions (phonologically related distractors: 3.2%; phonologically unrelated distractors: 2.9%). Thus, the absence of a phonological facilitation effect in the pronominal clitic production task is not the result of poor stimulus

selection but of specific processing characteristics associated with pronoun production—namely, the fact that the noun’s phonology is not selected for output. The latter observation, in turn, definitely excludes that the frequency inheritance effect obtained in Experiments 1 and 2 could be attributed to the fact that participants implicitly generate Det + NPs in order to produce clitics. Were this the case, we should have found a phonological effect in Experiment 3B. Furthermore, the absence of a phonological effect should also be taken as empirical support for the view that clitics are selected early since they do not have to wait for the phonology of the corresponding noun.

In short, the results of Experiments 3A and 3B confirm the absence of a gender congruency effect in pronominal clitic production. The straightforward implication of this result is that the process of grammatical feature retrieval is not based on the principle of selection-by-activation-level but is most likely an automatic consequence of selecting the lexical node that dominates those features.

## GENERAL DISCUSSION

There are two main findings in this study: a noun-to-pronoun frequency inheritance effect and the absence of a gender congruency effect in pronominal clitic production.

In two experiments (Experiments 1 and 2) we found that latencies in producing pronominal clitics in response to pictured objects are a function of the frequencies of the nouns they replace. Pronominal clitics were produced faster when the nouns they replaced were of high frequency than when they were of low frequency. This frequency effect was found to be robust over several repetitions of the picture stimuli and could not be ascribed to recognition differences of pictures with low and high frequency names.

The existence of the frequency inheritance phenomenon allows us to adjudicate between theories of the locus of the frequency effect in speech production. In the Introduction to Experiment 1 we discussed in some detail two alternative hypotheses of the locus of the frequency effect, and we reviewed the state of the evidence on this issue. We pointed out that the noun-to-pronoun frequency inheritance effect is predicted by those theories that locate the frequency effect at that level of lexical access where a word’s grammatical features are represented. This expectation is based on the observation that the selection of pronominal clitics depends strictly and only on the grammatical features—definiteness, case, number, and gender—associated with the NP that the pronoun replaces. Thus, the existence of the noun-to-pronoun frequency inheritance effect can be taken as evidence that the locus of the frequency effect in lexical access is

at the level where the word's grammatical properties are specified. On this reasoning, we would have to locate the frequency effect at the lemma level in those theories that assume a distinction between lemma and lexeme levels of representation, consistent with proposals made by Dell (1990) and Griffin and Bock (1998) but inconsistent with the proposal by Jescheniak and Levelt (1994) and Levelt et al. (1999) that the locus of the frequency effect is at the lexeme level or the so-called out-of-lemma level. In those theories that assume a single lexical layer, the locus of the frequency effect would be at the lexical node selection level since that is where a word's grammatical properties are specified (Caramazza, 1997).

The other main finding of this study is the absence of a gender congruency effect in pronominal clitic production. This result was obtained in three experiments (Experiments 2, 3A, and 3B), with different stimuli (Experiment 2 vs. 3A and 3B) and in contexts in which we found a noun frequency effect (Experiment 2) and a significant semantic interference effect (Experiment 3A). The absence of a gender congruency effect in the context of other effects, which are assumed to occur at the lexical level where a word's grammatical properties are specified, provides the necessary assurance that the task is sensitive enough to reveal effects at the relevant level of lexical access—lexical node selection.<sup>12</sup> Thus, the absence of a gender congruency effect in our experiments provides evidence against the hypothesis that grammatical feature selection is based on a process of selection-by-activation-level.

In the Introduction to Experiment 2 we reviewed extant evidence on the processes involved in gender feature selection and noted that two sets of results have been brought to bear on this issue. One set of results concerns the putative existence of a gender priming effect (Jescheniak & Levelt, 1994; Vigliocco et al., 2002). However, these results either have not been replicated, as in the case of Jescheniak and Levelt's (1994) study (see Jescheniak & Schriefers, 2001; van Berkum, 1996, 1997) or, as in the case of Vigliocco et al. (2002), the design of the experiments cannot exclude alternative explanations. We are left, therefore, with the results on gender congruency, which, although complex, have been replicated across

---

<sup>12</sup> It could be argued that our failure to observe a gender congruency effect reflects the possibility that distractor words do not strongly activate their gender features in Italian. However, this possibility has no plausible motivation since the activation of a noun's gender is assumed to occur automatically as part of the activation of its lexical node. This assumption has received considerable empirical support since the seminal paper by Schriefers (1993), as demonstrated by the reliable gender congruency effects observed in various languages by various researchers. Thus, in the measure to which this assumption is needed to explain the existence of gender congruency effects in some languages it must also apply to Italian. Otherwise, one would have to argue, without any motivation, that Italian does not allow the automatic spread of activation that is a core feature of all current theories of lexical access.

different languages in different laboratories. A consensus is beginning to emerge in the literature that the gender congruency effects that have been observed in German, Dutch, and Croatian may not so much reflect competition/priming at the level of gender feature selection as interference between alternative determiner forms (Caramazza et al., 2001b; Schiller & Caramazza, 2003; Schriefers et al., 2002). Thus, the results we have reported on the absence of a gender congruency effect in pronominal clitic production converge in support of the view that gender feature retrieval is an automatic (i.e., mandatory) process that is not subject to independent priming.

One issue remains to be addressed here. If the gender congruency effects observed in Dutch, German, and Croatian are really determiner congruency effects, why don't we find such effects in Italian and other Romance languages? We have already provided an explanation for the case of determiner production—because determiner selection in Italian occurs late in the process of NP production the activation of the distractor lexical node would have dissipated along with the activation of its associated gender feature, leaving little opportunity for significant activation of competing determiner forms. However, this explanation will not work in the case of pronominal clitics since, as we have argued above, the selection of these forms can occur early—as early as the selection of determiners in Dutch or German. A resolution of this issue requires consideration of the conditions under which one may expect to find competition (or some other type of interference; see Miozzo & Caramazza, 2003) in the selection of phonological forms for production.

In our experiments we used pronominal *enclitics*—that is, clitics that attach at the ends of verbs and that function somewhat like affixes (see Anderson, 1992; Benincà & Cinque, 1993; Monachesi, 1999). The crucial question is whether phonological material that is attached at the ends of words is subject to competition (or can be interfered with) in the same way as would appear to be the case for determiners and other free-standing morphemes. The evidence on this issue is mixed but it is beginning to converge in favour of the view that affixes are not subject to competition. Although Schriefers (1993) observed a gender congruency effect when Dutch participants produced NPs with gender-marked adjectives recent results by Schiller and Caramazza (2003) in German and Dutch and by Costa et al. (2003) in Croatian have failed to obtain a gender congruency effect for utterances in which the noun's gender surfaced only as a bound form. Our results with pronominal enclitics in Italian are in accord with the results obtained by Schiller & Caramazza (2003) and Costa et al. (2003), all showing that the selection of affixal forms is not subject to competition. Thus, the absence of competition in the selection of pronominal enclitics in our experiments does not present a paradox but is consistent with the

empirical generalisation that is emerging from studies with German, Dutch, and Croatian speakers (for further discussion see Costa et al., 2003). Nonetheless, it is important to emphasise that whatever the merits of the arguments presented here, the issue under consideration concerns effects at the level of phonological processing and do not impinge on decisions concerning the processes that are involved in grammatical feature retrieval. The evidence on the retrieval of grammatical features is now relatively clear, at least with respect to grammatical gender: The selection of a noun's gender is not subject to priming or interference by competition.

In conclusion, the main results of this study have clear implications for two issues of current interest in attempts to develop a theory of lexical access: the locus of the frequency effect and the mechanism of grammatical feature selection. The demonstration of a noun-to-pronoun frequency inheritance effect in gender-marked languages establishes that the locus of the frequency effect in lexical access is at the level where a word's grammatical features are represented (Caramazza, 1997; Dell, 1990; Griffin & Bock, 1998) and not at the lexeme level (Levelt et al., 1999). The absence of a gender congruency effect in pronominal clitic production shows that the retrieval of grammatical features is an automatic consequence of lexical node selection (Caramazza et al., 2001b) and not an independent process that is based on selection-by-activation-level (Levelt, 2001; Schriefers, 1993). Taken together these two conclusions illustrate a fundamental contrast between lexical node selection and grammatical feature retrieval: only the former operates on the principle of selection-by-activation level; the latter is a mandatory consequence of lexical node selection.

## REFERENCES

- Alario, F. X., & Caramazza, A. (2002). The production of determiners: Evidence from French. *Cognition*, 82, 179–223.
- Alario, F. X., Costa, A., & Caramazza, A. (2002). Hedging one's bets too much? A reply to Levelt. *Language and Cognitive Processes*, 17, 673–682.
- Anderson, S. (1992). *A-morphus morphology*. Cambridge: Cambridge University Press.
- Benincà, P., & Cinque, G. (1993). Su alcune differenze fra enclisi e proclisi. In *Omaggio a Gianfranco Folena* (Vol. 3, pp. 2313–2326). Padova, Italy: Editoriale Programma.
- Bonin, P., & Fayol, M. (2002). Frequency effects in the written and spoken production of homophonic picture names. *European Journal of Cognitive Psychology*, 14, 289–313.
- Bortolini, U., Tagliavini, C., & Zampolli, A. (1971). *Lessico di frequenza della lingua italiana contemporanea*. Milano: Garzanti.
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, 14, 177–208.
- Caramazza, A., Bi, Y., Costa, A., & Miozzo M. (2004). What determines the speed of lexical access: homophone or specific-word frequency? A reply to Jescheniak et al. 2003. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 278–282.

- Caramazza, A., & Costa, A. (2000). The semantic interference effect in the picture-word interference paradigm: Does the response set matter? *Cognition*, 75, 51–64.
- Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001a). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 1430–1450.
- Caramazza, A., & Hillis, A. (1990). Where do semantic errors come from? *Cortex*, 26, 95–122.
- Caramazza, A., Miozzo, M., Costa, A., Schiller, N., & Alario, F.X., (2001b). Lexical selection: A cross-language investigation of determiner production. In E. Dupoux (Ed.), *Language, brain, and cognitive development: Essays in honor of Jacques Mehler* (pp. 209–226). Cambridge, MA: MIT Press.
- Cohen, J. D., Mac Whinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments and Computers*, 25, 257–271.
- Costa, A., Alario, F. X., & Caramazza, A. (in press). On the categorical nature of the semantic interference effect in the picture-word interference paradigm. *Psychonomic Bulletin and Review*.
- Costa, A., & Caramazza, A. (2002). The production of noun phrases in English and Spanish: Implications for the scope of phonological encoding in speech production. *Journal of Memory and Language*, 46, 178–198.
- Costa, A., & Caramazza, A. (2005). *The retrieval of grammatical gender in speech production: Is gender retrieval sensitive to activation levels?* Unpublished manuscript, Harvard University, Cambridge, MA.
- Costa, A., Kovacic, D., Federenko, E., & Caramazza, A. (2003). The gender congruency effect and the selection of free-standing and bound morphemes: Evidence from Croatian. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1270–1282.
- Costa, A., Sebastian-Galles, N., Miozzo, M., & Caramazza, A. (1999) The gender congruity effect: Evidence from Spanish and Catalan. *Language and Cognitive Processes*, 14, 381–391.
- Damian, M. K., & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 1–18.
- Dell, G. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, 5, 313–349.
- Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, 46, 57–84.
- Glaser, W. R., & Döngelhoff, F. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 640–654.
- Glaser, W. R., & Glaser, M. O. (1989). Context effects on Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, 118, 13–42.
- Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, 38, 313–338.
- Janssen, N., & Caramazza, A. (2003). The selection of closed-class words in noun phrase production: the case of Dutch determiners. *Journal of Memory and Language*, 48, 635–652.
- Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 824–843.
- Jescheniak, J. D., Meyer, A. S., & Levelt, W. J. M. (2003). Specific word frequency is not all that counts in speech production. Reply to Caramazza et al. and new experimental data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 432–437.

- Jescheniak, J. D., & Schriefers, H. (1999). In search of gender recency: evidence from the production of pronouns. *Journal of Psycholinguistic Research*, 28, 739–745.
- Jescheniak, J. D., & Schriefers, H. (2001). Priming effects from phonologically related distractors in picture word interference. *Quarterly Journal of Experimental Psychology*, 54A, 371–382.
- Jescheniak, J. D., Schriefers, H. J., & Hantsch, A. (2001). Semantic and phonological activation in noun and pronoun production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 27(4), 1058–1078.
- La Heij, W. (1988). Components of Stroop-like interference in picture naming. *Memory and Cognition*, 16, 400–410.
- La Heij, W., Mak, P., Sander, J., & Willeboordse, E. (1998). The gender congruency effect in picture-word tasks. *Psychological Research*, 61, 209–219.
- Levelt, W. J. M. (1989). *Speaking. From intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Science*, 3, 223–232.
- Levelt, W. J. M. (2001). Spoken word production: A theory of lexical access. *PNAS. Proceedings of the National Academy of Sciences*, 98, 13464–13471.
- Levelt, W. J. M., Praamstra, P., Meyer, A. S., & Salmelin, R. (1998). An MEG study of picture naming. *Journal of Cognitive Neuroscience*, 10, 553–567.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–37.
- Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory and Cognition*, 7, 485–495.
- Meyer, A. S., & Bock, J. K. (1999). Representations and processes in the production of pronouns: Some perspectives from Dutch. *Journal of Memory and Language*, 41, 281–301.
- Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17, 1146–1160.
- Miozzo, M., & Caramazza, A. (1999). The selection of determiners in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 907–922.
- Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distractor frequency in the picture-word interference paradigm. *Journal of Experimental Psychology: General*, 132, 228–252.
- Miozzo, M., Costa, A., & Caramazza, A. (2002). The absence of a gender congruency effect in Romance languages: a matter of stimulus onset asynchrony? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 388–391.
- Monachesi, P. (1999). *A lexical approach to Italian cliticization*. Stanford, CA: CSLI.
- Perdijk, K., Spalek, K., & Schriefers, H. (2005). *Implicit gender priming in gender production: The role of gender homogeneity, response set, and stimulus type*. Unpublished manuscript, Nijmegen Institute of Cognition and Information, Nijmegen, The Netherlands.
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107–142.
- Rosinski, R. R. (1977). Picture-word interference is semantically based. *Child Development*, 48, 643–647.
- Schiller, N. O., & Caramazza, A. (2003). Mechanisms of determiner selection: Evidence from noun phrase production in German and Dutch. *Journal of Memory and Language*, 48, 169–194.
- Schmitt, B., Meyer, A. S., & Levelt, W. J. M. (1999). Lexical access in the production of pronouns. *Cognition*, 69, 313–335.
- Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 841–850.

- Schriefers, H. J., & Jescheniak, J. D. (1999). Representation and processing of grammatical gender in language production. *Journal of Psycholinguistic Research*, 28, 575–600.
- Schriefers, H. J., Jescheniak, J. D., & Hantsch, A. (2002). Determiner selection in noun phrase production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28, 941–950.
- Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, 29, 86–102.
- Schriefers, H. J., & Teruel, E. (2000). Grammatical gender in noun phrase production: the gender interference effect in German. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 1368–1376.
- Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 896–918.
- van Berkum, J. J. A. (1996). *The psycholinguistics of grammatical gender: Studies in language comprehension and production*. Doctoral Dissertation, Max Planck Institute for Psycholinguistics. Nijmegen, The Netherlands: Nijmegen University Press.
- van Berkum, J. J. A. (1997). Syntactic processes in speech production. The retrieval of grammatical gender. *Cognition*, 64, 115–152.
- Vigliocco, G., Lauer, M., Damian, M. F., & Levelt, W. J. M. (2002). Semantic and syntactic forces in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 46–58.
- Wheeldon, L. R., & Lahiri, A. (2002). The minimal unit of phonological encoding: Prosodic or lexical word. *Cognition*, 82, B31–41.

## APPENDIX A

### Pictured nouns used in Experiments 1 and 2

---

HF: bicchiere “glass”, bottiglia “bottle”, cappello “hat”, casa “house”, chiave “key”, finestra “window”, fiore “flower”, foglia “leaf”, gatto “cat”, letto “bed”, libro “book”, macchina “car”, piatto “plate”, poltrona “armchair”, porta “door”, scarpa “shoe”, tavolo “table”, telefono “telephone”, treno “train”, valigia “luggage”.

LF: antenna “aerial”, bilancia “scales”, botte “barrel”, carriola “wheelbarrow”, casco “helmet”, cerniera “zipper”, chitarra “guitar”, cuoco “cook”, faro “lighthouse”, fiocco “bow”, foca “seal”, freccia “arrow”, guanto “glove”, lucchetto “padlock”, melone “melon”, pipa “pipe”, secchio “bucket”, tovagliolo “napkin”, trattore “tractor”, zucca “pumpkin”.

---

## APPENDIX B

## Pictures and distractors used in Experiment 2

<i>Pictures</i>	<i>Distractors</i>	
	<i>Feminine</i>	<i>Masculine</i>
<i>High Frequency</i>		
FOGLIA "LEAF"	testa "head", benda "bandage"	corpo "body", tonno "tuna"
PIATTO "PLATE"	testa "head", benda "bandage"	corpo "body", tonno "tuna"
PORTA "DOOR"	camicia "shirt", cantina "cellar"	soldato "soldier", tamburo "drum"
GATTO "CAT"	camicia "shirt", cantina "cellar"	soldato "soldier", tamburo "drum"
MACCHINA "CAR"	cipria "powder", pianta "plant"	quadro "picture", tempio "temple"
BICCHIERE "GLASS"	cipria "powder", pianta "plant"	quadro "picture", tempio "temple"
POLTRONA "ARMCHAIR"	gamba "leg", torta "cake"	cielo "sky", tappo "cork"
FIORE "FLOWER"	gamba "leg", torta "cake"	cielo "sky", tappo "cork"
FINESTRA "WINDOW"	carne "meat", volpe "fox"	pesce "fish", germe "germ"
TAVOLO "TABLE"	carne "meat", volpe "fox"	pesce "fish", germe "germ"
BOTTIGLIA "BOTTLE"	nave "ship", fune "rope"	cane "dog", rene "kidney"
TELEFONO "TELEPHONE"	nave "ship", fune "rope"	cane "dog", rene "kidney"
CASA "HOUSE"	terra "earth", fibbia "buckle"	bosco "wood", mosto "must"
CAPPELLO "HAT"	terra "earth", fibbia "buckle"	bosco "wood", mosto "must"
SCARPA "SHOE"	costa "coast", piuma "feather"	campo "field", circo "circus"
TRENO "TRAIN"	costa "coast", piuma "feather"	campo "field", circo "circus"
VALIGIA "LUGGAGE"	guardia "guard", treccia "plait"	braccio "arm", graffio "scratch"
LETTO "BED"	guardia "guard", treccia "plait"	braccio "arm", graffio "scratch"
CHIAVE "KEY"	borsa "bag", renna "reindeer"	fuoco "fire", fungo "mushroom"
LIBRO "BOOK"	borsa "bag", renna "reindeer"	fuoco "fire", fungo "mushroom"

*Continued overleaf.*

<i>Pictures</i>	<i>Distractors</i>	
	<i>Feminine</i>	<i>Masculine</i>
<i>Low frequency</i>		
FRECCIA "ARROW"	testa "head", benda "bandage"	corpo "body", tonno "tuna"
GUANTO "GLOVE"	testa "head", benda "bandage"	corpo "body", tonno "tuna"
PIPA "PIPE"	camicia "shirt", cantina "cellar"	soldato "soldier", tamburo "drum"
MELONE "MELON"	camicia "shirt", cantina "cellar"	soldato "soldier", tamburo "drum"
BOTTE "BARREL"	cipria "powder", pianta "plant"	quadro "picture", tempio "temple"
FARO "LIGHTHOUSE"	cipria "powder", pianta "plant"	quadro "picture", tempio "temple"
FOCA "SEAL"	gamba "leg", torta "cake"	cielo "sky", tappo "cork"
LUCCHETTO "PADLOCK"	gamba "leg", torta "cake"	cielo "sky", tappo "cork"
BILANCIA "SCALES"	carne "meat", volpe "fox"	pesce "fish", germe "germ"
FIOCCO "BOW"	carne "meat", volpe "fox"	pesce "fish", germe "germ"
ZUCCA "PUMPKIN"	nave "ship", fune "rope"	cane "dog", rene "kidney"
TOVAGLIOLO "NAPKIN"	nave "ship", fune "rope"	cane "dog", rene "kidney"
CARRIOLA "WHEELBARROW"	terra "earth", fibbia "buckle"	bosco "wood", mosto "must"
SECCHIO "BUCKET"	terra "earth", fibbia "buckle"	bosco "wood", mosto "must"
ANTENNA "AERIAL"	costa "coast", piuma "feather"	campo "field", circo "circus"
TRATTORE "TRACTOR"	costa "coast", piuma "feather"	campo "field", circo "circus"
CERNIERA "ZIPPER"	guardia "guard", treccia "plait"	braccio "arm", graffio "scratch"
CUOCO "COOK"	guardia "guard", treccia "plait"	braccio "arm", graffio "scratch"
CHITARRA "GUITAR"	borsa "bag", renna "reindeer"	fuoco "fire", fungo "mushroom"
CASCO "HELMET"	borsa "bag", renna "reindeer"	fuoco "fire", fungo "mushroom"

## APPENDIX C

Materials used in Experiment 3A. Pictured nouns are in capitals. Distractors are in lower case letters in the following order: semantically related, masculine—semantically unrelated, masculine—semantically related, feminine—semantically unrelated, feminine

*Feminine pictured nouns*

BANANA “BANANA”—mirtillo “blueberry”—aquilone “kite”—fragola “strawberry”—formina “mould”  
 BOTTIGLIA “BOTTLE”—calice “goblet”—salice “willow”—caraffa “carafe”—rubrica “address book”  
 CAROTA “CARROT”—sedano “celery”—polipo “octopus”—rapa “turnip”—riga “line”  
 CHITARRA “GUITAR”—flauto “flute”—fiasco “flask”—tromba “trumpet”—stalla “stable”  
 FOCA “SEAL”—rospo “toad”—tappo “cork”—renna “reindeer”—barra “bar”  
 FOGLIA “LEAF”—ramo “branch”—viso “face”—erba “grass”—luna “moon”  
 GAMBA “LEG”—polso “wrist”—forno “oven”—coscia “thigh”—crosta “crust”  
 GIRAFFA “GIRAFFE”—canguro “kangaroo”—mattoncino “brick”—scimmia “monkey”—treccia “plait”  
 MACCHINA “CAR”—camion “truck”—canale “canal”—carrozza “carriage”—castagna “chestnut”  
 POLTRONA “ARMCHAIR”—divano “sofa”—mulino “mill”—sedia “chair”—busta “envelope”  
 RANA “FROG”—pesce “fish”—sacco “sack”—mosca “fly”—tasca “pocket”  
 SCARPA “SHOE”—calzino “sock”—lampone “raspberry”—maglietta “T-shirt”—graffetta “paper clip”  
 ZUCCA “PUMPKIN”—melone “melon”—vagone “wagon”—patata “potato”—grotta “cavern”

*Masculine pictured nouns*

CANNONE “CANNON”—fucile “rifle”—manico “handle”—spada “sword”—sfera “sphere”  
 CAPPELLO “HAT”—guanto “glove”—timbro “stamp”—sciarpina “scarf”—formica “ant”  
 CAVALLO “HORSE”—maiale “pig”—bidone “tank”—zebra “zebra”—cetra “lyre”  
 CIGNO “SWAN”—pavone “peacock”—barile “barrel”—quaglia “quail”—quercia “oak”  
 FUNGO “MUSHROOM”—carciofo “artichoke”—battello “boat”—zucchina “courgette”—caviglia “ankle”  
 GATTO “CAT”—leone “lion”—disco “disk”—tigre “tiger”—tenda “tent”  
 LETTO “BED”—armadio “closet”—bullone “bolt”—scrivania “desk”—damigiana “demijohn”  
 PIATTO “PLATE”—cucchiaino “spoon”—fermaglio “clasp”—scodella “bowl”—farfalla “butterfly”  
 PIEDE “FOOT”—braccio “arm”—soldato “soldier”—spalla “shoulder”—chiesa “church”  
 RASTRELLO “RAKE”—badile “shovel”—bavero “collar”—paletta “small shovel”—piscina “swimming-pool”  
 TAMBURO “DRUM”—pianoforte “piano”—carrarmato “tank”—fisarmonica “accordion”—spillatrice “stapler”  
 TAVOLO “TABLE”—banco “desk”—pozzo “well”—panca “bench”—pinza “pliers”  
 TRENO “TRAIN”—aereo “aircraft”—conte “count”—barca “boat”—pizza “pizza”

## APPENDIX D

Materials used in Experiment 3B. Pictured nouns are in capitals. Distractors are in lower case letters in the following order: phonologically related, masculine—phonologically unrelated, masculine—phonologically related, feminine—phonologically unrelated, feminine

---

*Feminine pictured nouns*

BANANA “BANANA”—banco “desk”—disco “disk”—barba “beard”—penna “pen”  
 BOTTIGLIA “BOTTLE”—bottone “button”—gradino “step”—bottega “shop”—capanna “hut”  
 CAROTA “CARROT”—cappotto “coat”—gabbiano “seagull”—cantina “cellar”—rubrica “address book”  
 CHITARRA “GUITAR”—chiodo “nail”—bidone “tank”—chiave “key”—pietra “stone”  
 FOCA “SEAL”—foro “hole”—nido “nest”—foto “photo”—cava “quarry”  
 FOGLIA “LEAF”—forno “oven”—leone “lion”—forca “fork”—punta “point”  
 GAMBA “LEG”—gancio “hook”—polipo “octopus”—garza “gauze”—pizza “pizza”  
 GIRAFFA “GIRAFFE”—gigante “giant”—costume “costume”—giostra “merry-go-round”—sottana “petticoat”  
 MACCHINA “CAR”—mattone “brick”—lampone “raspberry”—maschera “mask”—zucchina “courgette”  
 POLTRONA “ARMCHAIR”—polso “wrist”—gesso “chalk”—posta “mail”—spina “thorn”  
 RANA “FROG”—raso “satin”—fumo “smoke”—rata “instalment”—buca “pit”  
 SCARPA “SHOE”—scudo “shield”—pozzo “well”—scatola “box”—formica “ant”  
 ZUCCA “PUMPKIN”—zoccolo “clog”—armadio “closet”—zuppa “soup”—culla “cradle”

*Masculine pictured nouns*

CANNONE “CANNON”—carciofo “artichoke”—pinguino “penguin”—caviglia “ankle”—farfalla “butterfly”  
 CAPPELLO “HAT”—camino “fireplace”—fiasco “flask”—caverna “cave”—treccia “plait”  
 CAVALLO “HORSE”—cancello “gate”—zucchero “sugar”—campana “bell”—lametta “razor blade”  
 CIGNO “SWAN”—cibo “food”—nodo “knot”—cima “top”—nave “ship”  
 FUNGO “MUSHROOM”—fuso “spindle”—cero “large candle”—funo “rope”—tela “cloth”  
 GATTO “CAT”—gas “gas”—bus “bus”—gara “competition”—luna “moon”  
 LETTO “BED”—lembo “border”—tappo “cork”—lente “lens”—pinza “pliers”  
 PIATTO “PLATE”—piombo “lead”—ditale “thimble”—piatta “jointer plane”—grotta “cavern”  
 PIEDE “FOOT”—pino “pine”—faro “lighthouse”—pila “battery”—rete “net”  
 RASTRELLO “RAKE”—rame “copper”—lupo “wolf”—rapa “turnip”—vite “screw”  
 TAMBURO “DRUM”—tacco “heel”—mento “chin”—talpa “mole”—vasca “basin”  
 TAVOLO “TABLE”—tasto “key”—rospo “toad”—targa “plaque”—serpe “snake”  
 TRENO “TRAIN”—tronco “trunk”—mulino “mill”—tromba “trumpet”—pancia “stomach”

---