Bilingualism: Language and Cognition 6 (3), 2003, 181–200 © 2003 Cambridge University Press DOI: 10.1017/S1366728903001123

On the autonomy of the grammatical gender systems of the two languages of a bilingual*

ALBERT COSTA
Universitat de Barcelona
DAMIR KOVACIC
Scuola Internazionale Superiore di Studi Avanzati (SISSA)
JULIE FRANCK
Université de Genève
ALFONSO CARAMAZZA
Harvard University

In five experiments highly-proficient bilinguals were asked to name two sets of pictures in their L2: a) pictures whose names in the L2 and their corresponding L1 translations have the same grammatical gender value, and b) pictures whose names in the L2 and their corresponding L1 translations have different gender values. In Experiments 1, 2, and 3 Croatian-Italian speakers were asked to name the pictures in Italian by means of NPs in various experimental contexts. In Experiment 4A, Spanish-Catalan and Catalan-Spanish bilinguals were asked to name the pictures in Spanish, and in Experiment 4B, Italian-French bilinguals did so in French. The results of these experiments revealed no differences between same- and different-gender pictures. Furthermore, the performance of Italian, Spanish, and French monolingual speakers parallels that of bilingual speakers. However, a robust frequency effect was observed across experiments. This pattern of results supports the notion that the gender value of the words in the non-response language does not affect processing in the response language, and suggests that the two gender systems of a bilingual are functionally autonomous.

Fluent speech requires, among other things, the selection of the proper lexical items for the intended meaning, and the retrieval of their grammatical properties. The grammatical properties of lexical nodes govern, to some extent, the relationships that words can establish in the sentence. Furthermore, certain grammatical properties also have an effect on the selection of other lexical items in the utterance. Consider, for example, the case in which an Italian speaker wants to produce the noun phrase (NP) *la mela* 'the apple'. In this scenario, the speaker needs to have access to the grammatical gender of the noun (*mela*$_{\text{Fem}}$ 'apple') in order to select the proper determiner form (*la*), since determiners are gender-marked in this language (*la* for singular feminine and *il* (or *lo*) for singular masculine nouns, respectively). Although our knowledge of how grammatical features are accessed in speech production is still limited, some insights have been gathered in recent years (Schriefers, 1993; Van Berkum, 1997; La Heij, Mak, Sander and Willeboordse, 1998; Costa, Sebastián-Gallés, Miozzo and Caramazza, 1999; Miozzo and Caramazza, 1999; Schriefers and Teruel, 2000; Caramazza, Miozzo, Costa, Schiller and Alario, 2001; Alario and Caramazza, 2002; Miozzo, Costa and Caramazza, 2002; Schriefers, Jescheniak and Hantsch, 2002; Janssen and Caramazza, 2003; Schiller and Caramazza, 2003). In this article we extend some of these insights to the context of bilingual speech production. More specifically, we explore the extent to which the gender systems of the two languages of bilingual speakers interact during the production of fluent speech or, in other words, whether the grammatical features of the non-response language (the language in which the experimental task is not being performed) affect the retrieval of the noun's gender feature in the response language (the language in which the experimental task is being performed).

The question of whether the linguistic properties of the non-response language affect production in the response language has been shown to receive a positive answer at several linguistic levels. For example, formal similarity between translation words seems to have an influence in the naming performance. That is, words with phonologically similar translations – cognates, e.g., for English-Italian bilinguals: *organo* 'organ' – are produced faster and more reliably than words with phonologically dissimilar translations – non-cognates, e.g., *sedia* 'chair' (Janssen, 1999; Costa, Caramazza and Sebastián-Gallés,

* This research was supported by NIH grant DC 04542, a grant from the Spanish Government (BSO2001-3492-C04-01), a grant from the Swiss FNS (1114-068250.02), and also by the McDonnell grant 'Bridging Mind Brain & Behavior'. Albert Costa was supported by a Post-doctoral Fellowship from SISSA and by the research program 'Ramon y Cajal' from the Spanish Government. Damir Kovacic was supported by a Pre-doctoral fellowship from Scuola Internazionale Superiore di Studi Avanzati (SISSA). The authors are grateful to Luca Bonati, Marcela Peña, Jacques Mehler and Xavier Alario for their helpful comments on this paper. We also thank Minkha Hoang, Simone Cotti and Olga Gomez for their help in the preparation of the experiments. Requests for reprints may be addressed to Albert Costa.

Address for correspondence
Albert Costa, Departament de Psicologia Basica, Universitat de Barcelona, P. Vall d'Hebron, 171 08035 Barcelona, Spain
E-mail: *acosta@psi.ub.es*

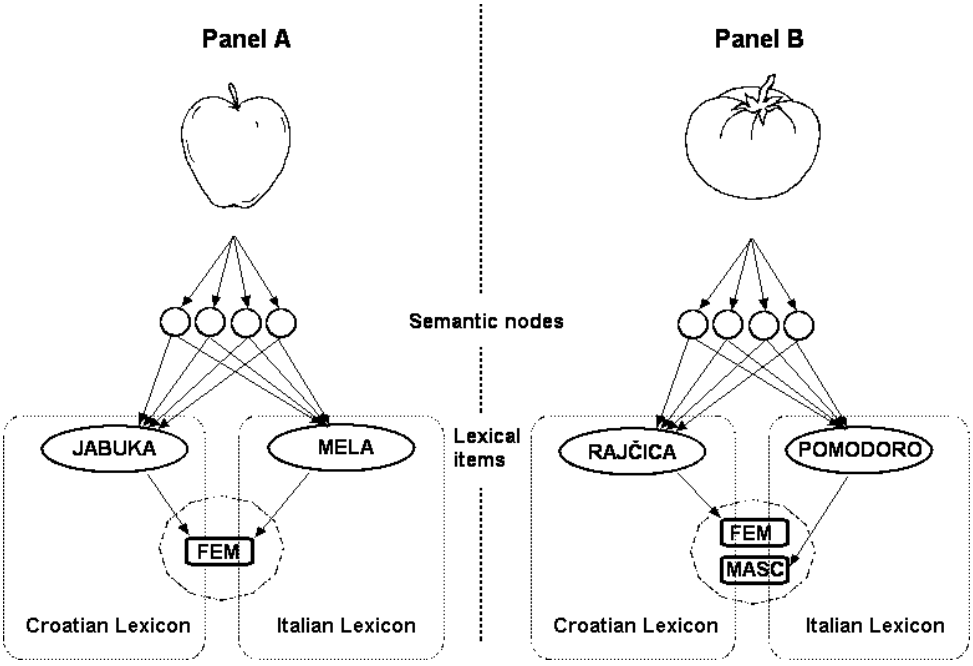

Figure 1. Schematic representation of the gender integrated representation hypothesis. Panel A represents the case of words that have the same gender across languages. Panel B represents the case of words that have different genders across languages. In both cases, there are only two gender features shared by the two lexicons.

2000a; Gollan and Acenas, 2000, in press; see also de Groot and Nas, 1991; Kroll, Dijkstra, Janssen and Schriefers, 2000). According to one account of this phenomenon (Costa et al., 2000a), the cognate effect reveals the interaction between the phonological representations of a target word and its translation at the phonological level. At any rate, this effect reveals that some properties of the words in the non-response language affect the naming process in the response language. In this article we explore whether this interaction across languages is present at the level at which lexical-grammatical properties are represented and accessed. Before addressing this specific issue, we need to put forward some basic assumptions about the functional architecture of bilingual speakers.

There are two main processing assumptions that we adopt here: a) the two languages of a bilingual share the same conceptual system, and b) when producing a word, the semantic system activates the words of the two lexicons. That is, when a Croatian-Italian bilingual wants to produce the concept APPLE in Italian, not only is the corresponding Italian word (*mela*) activated but also its Croatian translation (*jabuka*). These two assumptions have received empirical support (e.g., Hermans, Bongaerts, de Bot and Schreuder, 1998; Costa et al., 2000a; Costa, Colomé and Caramazza, 2000b; Gollan and Acenas, 2000; Colomé, 2001; Lee and Williams, 2001) and have been widely adopted in models of bilingual language processing (Potter, So, von Eckardt and Feldman, 1984; de Bot, 1992; Kroll and Stewart, 1994; La Heij, Hooglander, Kerling and van der Velden, 1996; Poulisse, 1997; Green, 1998; Hermans, Bongaerts, de Bot and Schreuder, 1998; Costa and Caramazza, 1999; Costa, Miozzo and Caramazza, 1999; Lee and Williams, 2001; but see Van Hell and de Groot, 1998).[1]

Regarding the structure of the gender systems of the two languages of a bilingual, there are, at least, two possible hypotheses that one can entertain. For those languages in which the gender systems are relatively similar, it could be argued that the representation of grammatical gender is shared across languages. That is, it is possible that words with the same gender value share their gender feature, regardless of the language they belong to. Consider the case of two gender-marked languages such as Italian and Croatian. Some of the words in these two languages have the same gender value (e.g., *mela*$_{\text{Fem}}$ and *jabuka*$_{\text{Fem}}$ 'apple' both have feminine gender, see Figure 1, panel A), while other words have different gender values (e.g., *pomodoro*$_{\text{Masc}}$, and *rajčica*$_{\text{Fem}}$ 'tomato', see Figure 1, panel B). According to this GENDER-INTEGRATED VIEW,

[1] The assumption of co-activation of response and non-response language begs the question of how the speaker selects the target lexical node in the intended language rather than its translation. Here we assume that this issue has a reasonable solution (see Green, 1998; Costa, in press; La Heij, in press) and we will proceed to consider whether, given the co-activation of the two lexicons of a bilingual, the grammatical properties of the lexical nodes in the non-response language affect processing in the response language.

there is only one single integrated gender system for the two languages of a bilingual, and therefore the first set of words would share their gender feature, while the latter would not.

But how plausible is the idea that the two gender systems of a bilingual are integrated? Why would the bilingual speaker treat two grammatical properties in different languages as being the same grammatical property? After all, these grammatical properties may have different lexical and morphological implications in the two languages. For example, it is possible that in language A (e.g., Italian) the gender value of a given word has consequences for the selection of, say, a determiner while in language B (e.g., Croatian) no such consequences are present. Thus, what are the commonalities between the two grammatical properties, beyond the linguistic term we use to refer to them ('grammatical gender'), such that they might lead bilingual speakers to consider them to be the same linguistic property? In other words, is there any reason for bilinguals to hypothesize that grammatical property X in language A and grammatical property X in language B refer to the same grammatical property? We think that there are at least two reasons for expecting this to be the case.

First, the commonality can be found when exploring the correlation between grammatical and natural (or semantic) gender. In many languages, words referring to concepts that have male semantic properties tend to take the same grammatical gender value (e.g., masculine), and concepts that have female semantic properties tend to take the feminine grammatical gender value. Thus, for those words that refer to natural elements with semantic gender values, the grammatical gender value is systematically the same. Thus, given that such a correlation is present in the two languages of a bilingual, the L2 learner may realize that a certain set of words referring to concepts with semantic gender of type A (man, uncle, male cat, etc.) take always the same grammatical gender value (value Y in language A and value X in language B), and that therefore such a grammatical feature in the two languages may actually refer to the same property, regardless of the specific implications of such a property in the two languages.[2] However, for such an argument to work, semantic information needs to be relevant in the processes of acquiring grammatical gender. And, in fact some researchers have argued that L2 learners tend to use such semantic cues to assign gender to the new L2 words (Andersen, 1984; Carroll, 1999).

There is another domain in which it is possible to find commonality in the grammatical values of L1 and L2, that may help bilinguals to infer that grammatical gender refers to the same property in the two languages. In many languages there is a correlation between the word's phonology and its gender value (in Italian a large percentage of nouns ending with *-o* are masculine, and a large percentage ending with *-a* are feminine). This correlation may exist also in the other language of a bilingual (e.g., in Croatian a large percentage of nouns ending with a consonant are masculine, and a large percentage of nouns ending with a vowel are feminine). This correlation is not only present for individual lexical items, but also extends to morphological gender inflections, independently of the noun's final sound. For example, the masculine inflection in Italian is usually *-o-* (*pomodoro ross-o* 'the red table', *ponte ross-o* 'the red bridge'), while the feminine inflection is *-a-* (*mela ross-a* 'the red apple', *tigre ross-a* 'the red tiger'). Thus, the bilingual speaker may be aware of these correlations and may infer that they are associated with the same grammatical property.

A further suggestion that grammatical gender in L2 is not treated as a completely different feature from L1 comes from the observation that the existence of grammatical gender marking in L1 facilitates the acquisition of the L2 gender agreement system. That is, bilinguals whose L1 lacks grammatical gender agreement have more difficulty in acquiring the L2 agreement system than bilinguals for whom gender agreement is present in L1 (Carroll, 1989; Hawkins and Chan, 1997; Granfeldt, in press). Strictly speaking, this observation only suggests that the acquisition of the mechanisms involved in gender agreement could be facilitated by the existence of a similar mechanism in L1. However, it also opens the question of the extent to which such a facilitation extends not only to the mechanism involved in gender agreement, but also to the structural properties of the gender system. In this respect, it is important to consider the extent to which the similarity between the structural properties of the gender systems of the two languages of a bilingual is a relevant factor for the presence of a gender integrated system. The experiments reported here address this issue by considering the performance of bilinguals of two languages with: a) relatively asymmetrical (structurally speaking) gender systems (Croatian-Italian bilinguals), and b) identical (structurally speaking) gender systems (Spanish-Catalan bilinguals and Italian-French bilinguals).

The second view postulates a complete autonomy of the gender systems of the two languages of a bilingual, THE LANGUAGE-AUTONOMY VIEW. Here, each language would

[2] This argument does not necessarily imply that grammatical gender is a semantic variable. We consider gender as a grammatical variable that does not necessarily relate to the semantic properties of words. However, for some words there is actually a correlation between grammatical gender and semantic properties, and therefore it is possible for the bilingual speaker to notice such a correlation across languages. In fact, semantic variables have been shown to affect gender agreement processes in monolinguals speakers (e.g., Vigliocco and Franck, 1999), which suggests that the semantic properties of some words may be important when computing gender agreement.

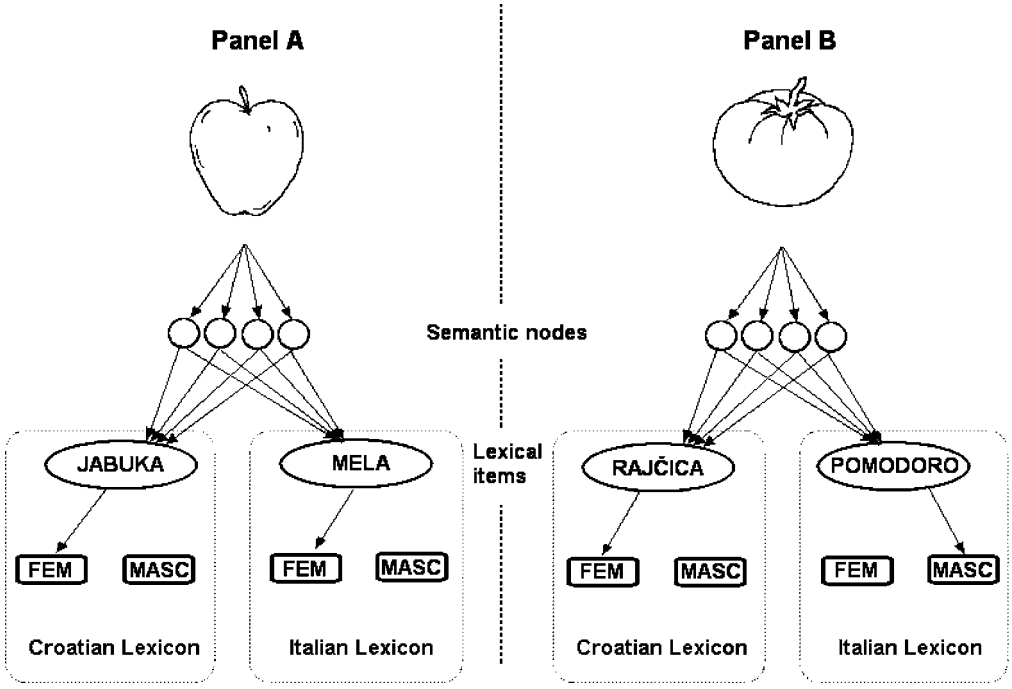

Figure 2. Schematic representation of the gender autonomous representation hypothesis. Panel A represents the case of words that have the same gender across languages. Panel B represents the case of words that have different genders across languages. In both cases, there are two gender features for each language.

have its own specific gender system, and the fact that two translations have different or the same gender values is irrelevant for the organization of the L2 grammatical knowledge. Accordingly, a bilingual speaker would have autonomous gender systems for each of his/her languages. For example, in the case of Croatian-Italian bilinguals, one would postulate independent feminine and masculine gender features for each of the two languages (see Figure 2).

Consider now the situation of a Croatian-Italian bilingual asked to produce an NP in Italian (*la mela*$_{\text{Fem}}$ 'the apple'). Following the assumptions outlined above, upon the recognition of the picture the semantic system activates the lexicons of the two languages of a bilingual, therefore activating the Italian target word *mela*$_{\text{Fem}}$ and its Croatian translation *jabuka*$_{\text{Fem}}$. If we were to apply the same spreading activation principle between the lexical nodes and their corresponding gender features, then all activated words should send some proportional activation to their corresponding grammatical features (see Janssen and Caramazza, 2003). The question then is the extent to which the activation of the gender feature (feminine) of the word in the non-response language (*jabuka*) affects the retrieval of the gender feature (feminine) of the target word in the response language (*mela*).

An answer to this question depends not only on the type of functional architecture of the two gender systems of a bilingual, but also on more general principles regarding gender retrieval in speech production. The two main proposals regarding this issue differ basically in whether or not the retrieval of the gender feature depends on its level of activation. According to one class of models – the activation dependent models (e.g., Schriefers, 1993; Levelt, 2001; Vigliocco, Lauer, Damian and Levelt, 2002) – the speed and efficiency with which the gender of the noun is retrieved depends to some extent on the activation level of that feature at the moment of selection: if the gender feature is highly activated (or if the activation level of other gender features is relatively low) at the moment when it is needed, then its selection is achieved faster than if its activation level is low (or if the activation of other gender features is high). An alternative account – the automatic gender-access model – proposes that the gender value of a given noun becomes automatically available for further processing upon the selection of the noun's lexical node (e.g. Caramazza et al., 2001; Costa, Kovacic, Fedorenko and Caramazza, in press; Schiller and Caramazza, 2003). That is, gender access is a direct (and automatic) consequence of lexical selection, rendering the notion of activation levels irrelevant for gender access.

Considering together these two dimensions, a) the representation of the gender systems of the two languages of a bilingual (autonomous vs. integrated) and b) the principles regarding the retrieval of the gender feature (activation-level dependency vs. automatic access), we can derive the following predictions regarding the effect

		Gender Selection Mechanism	
		Selection by Activation Level	Automatic Selection
Bilingual Gender System	Integrated	Same Gender < Different Gender	Same Gender = Different Gender
	Segregated	Same Gender = Different Gender	Same Gender = Different Gender

Figure 3. Description of the predictions for the naming performance of bilingual speakers for same- and different-gender pictures. The predictions are broken by the different assumptions about the selection of the gender feature and the architecture of the gender system in bilinguals. The only combination of assumptions that predict a difference between the two sets is shaded.

of the gender value of the words in the non-response language.

In the framework of the activation-dependent models of gender retrieval two contrasting predictions can be derived for words that have same- or different-genders across languages. First, if the gender systems of the two languages are, to some extent, shared, the retrieval of the target's gender feature would be faster when the translation word has the same gender (e.g., *mela*$_{\text{Fem}}$, *jabuka*$_{\text{Fem}}$ 'apple') than when it has a different gender (e.g., *pomodoro*$_{\text{Masc}}$, *rajčica*$_{\text{Fem}}$ 'tomato'). This is because, in the former case the target's gender feature (e.g., feminine) would be more activated than in the latter (see Figure 1) since it will receive activation from two sources, the word in the response language (*mela*$_{\text{Fem}}$ 'apple' in Italian) and its translation (*jabuka*$_{\text{Fem}}$ 'apple' in Croatian). Alternatively, if the gender systems of the two languages of a bilingual are relatively autonomous, the selection of the gender feature for same-gender translations and different-gender translations should be similar. This is because in both cases the gender of the target word would not receive any activation from the translation words since the gender features are not shared across languages (see Figure 2). To summarize, in this framework, the integrated view predicts faster gender retrieval for words whose translations have the same-gender value than for words whose translations have a different-gender value, while the autonomous view predicts no difference between the two types of words.

According to the automatic gender-access models, the retrieval of the gender feature does not depend on activation levels and therefore no systematic differences in naming latencies between words whose translations have the same- or different-gender should be observed.

As is shown in Figure 3, there are two dimensions (with two values each) that are relevant to deriving predictions about the naming performance for same- and different-gender pictures. The first dimension is whether or not gender retrieval depends on the level of activation of the gender features, while the second refers to whether the two gender systems of a bilingual are integrated or independent. Importantly, only one combination of these assumptions predicts a difference between same- and different-gender pictures: gender selection is achieved by activation levels and the two gender systems of a bilingual are integrated.

Despite these specific hypotheses, there are other reasons why words with same-gender translations may be easier to produce than words with different-gender translations. It is possible that when acquiring L2, the similarity among the gender values of the new L2 words and their corresponding translations in L1 facilitates the learning of the L2 gender values, therefore resulting in faster (easier) learning for same-gender than for different-gender words. If this learning advantage has a permanent effect in bilinguals' production performance, it is possible that we still find a distinction between the two sets of words later on in life (e.g., see Barry, Morrison and Ellis, 1997; Ellis and Morrison, 1998 for permanent effects of age of acquisition on speech production). There are other linguistic domains in which the effects of already established L1 linguistic properties seem to have a dramatic effect on the acquisition of L2 properties, namely the phonological level (see for example Flege, 1999; Pallier, Colomé and Sebastián-Gallés, 2001).[3]

We report five experiments designed to explore the interaction between the gender systems of the two languages of highly-proficient bilingual speakers during the production of gender-marked NPs. The main manipulation in the experiments refers to whether the target's translation and the response had the same or different gender values.

[3] In this framework, it is assumed that the phonologies of the two languages of a bilingual are in contact and affect each other (see for example Flege, 1999; Yeni-Komshian, Flege and Liu, 2000). In fact, the existence of a foreign accent reveals the inability of some bilingual speakers to establish categories for those L2 sounds that do not have an exact counterpart in L1. In this scenario, it has been argued that the similarity between the phonological systems of L1 and L2 may affect the likelihood of properly acquiring a new L2 sound. That is, when a new L2 sound has to be acquired and the L1 of a bilingual contains a similar sound, the acquisition of that new phoneme becomes difficult, since it gets assimilated to the existing L1 category. The question we address here is whether this kind of interaction between the two languages of a bilingual is also present, to some extent, when learning the gender value of L2 words.

Experiment 1: Does the gender value of L1 words affect L2 picture naming?

In this experiment highly-proficient Croatian-Italian bilinguals were asked to name pictures by means of gender-marked determiner + noun NPs (*la mela* 'the apple'). Italian has two gender values, feminine and masculine, while Croatian has three genders, masculine, feminine and neuter. Despite this structural asymmetry, the distribution of the words across genders is similar in the two languages. In Croatian, masculine words account for 40%, feminine words for 45%, and neuter words for 15% (Corbett, 1991) of the nouns in the language, while in Italian masculine words account for 60% and feminine words for 40% approximately. In both languages, the noun's gender value has implications for the selection of some closed class words and inflectional suffixes. For example, in Italian most adjectives are inflected for gender (*rosso* vs. *rossa* 'red' for masculine and feminine nouns, respectively) and the selection of determiners also depends on the noun's gender value (e.g., the definite determiners *il* or *lo* vs. *la*, for singular masculine and feminine nouns, respectively). Along the same lines, in Croatian, adjectives are also inflected for gender (*moj* vs. *moja* 'my' for masculine and feminine nouns, respectively), and the selection of pronouns depends on the noun's gender value (*ga* vs. *je* 'it' for masculine and feminine nouns respectively, as in 'I see it'). There are also some differences between the two languages regarding the gender agreement process in the context of NPs, such as the fact that Croatian NPs do not include determiners while Italian NPs do. We defer further discussion of this issue to the discussion of Experiment 1.

Pictures were divided into two sets: a) pictures whose names have the same gender in Italian and in Croatian ($mela_{\text{Fem}}$, $jabuka_{\text{Fem}}$ 'apple'), and b) pictures whose names have different genders (e.g., $pomodoro_{\text{Masc}}$, $rajčica_{\text{Fem}}$ 'tomato'). As argued above, if the gender values of the words in the non-response language interact with the gender values of the words in the response language, then a difference between the two sets is expected (see Figure 3). Otherwise, if the gender systems are autonomous, then naming latencies should be independent of the gender value of the translation word. Thus, according to this latter view a null effect is predicted, and therefore it is important to assess the sensitivity of our design to detect lexical effects. At any rate, caution must be exercised when interpreting a null effect.

We also asked a group of monolingual Italian speakers to name the pictures. This is an important control group since we are comparing naming latencies from two different sets of pictures. Thus, if the two sets of pictures are comparable in other respects aside from gender match or mismatch across languages then no difference between the two sets should be observed in this group. That is, to attribute any difference between the two sets of pictures to the gender match or mismatch across languages, that difference should be present only for bilinguals.

Method

Participants

Ten highly-proficient Croatian-Italian bilinguals participated in the experiment. All came from bilingual communities (Istria, Croatia) and began learning Italian in their childhood. They had all been living in Italy for at least 1.5 years at the time that the experiment was conducted (see Appendix A for a description of the sample of participants). Ten college students, native speakers of Italian with no knowledge of Croatian participated in the control group.

Materials

Eighty pictures, half with names of different genders in Italian and Croatian, and half with same-gender names were selected. In each set, half of the words were masculine and half feminine. In order to keep symmetry between the gender values of Croatian and Italian, only Italian words with feminine or masculine Croatian translations were included. The words in the two sets were of comparable frequency (48 vs. 42, for same- and different-gender pictures, respectively, $F < 1$; *Dizionario di Frequenza della lingua Italiana*, CNR) and had comparable number of syllables (2.5 vs. 2.6, same- and different-gender pictures respectively, $F < 1$). With the exception of two words (one in each Picture set), no cognate words were included in the experiment. Care was taken to distribute elements of different semantic categories similarly across the two sets of pictures. The pictures appeared in black and white at the center of the screen and were presented three times (three repetitions) in three blocks of 83 trials (80 experimental trials and three warm-up fillers at the beginning of each block). Each picture appeared once per block. Block trials were randomized with the restriction that trials of the same condition or the same gender appeared in no more than three consecutive trials. Furthermore, a semantic or phonological relationship between two consecutive trials was avoided.

Procedure

Participants were tested individually. They were instructed to name the pictures in Italian as quickly and as accurately as possible by using the definite article *il*, *la* 'the' plus the name of the object. A familiarization phase, in which participants were presented with the entire set of pictures and were asked to name them with simple NPs of the type determiner + noun, preceded the experiment proper. Each trial had the following structure: 1) a question mark

appeared on the screen and stayed on until participants pressed the space bar; 2) a fixation point (+) appeared on the screen for 500 ms; 3) the picture appeared on the screen for 2 seconds or until participants' response; and 4) the next trial started 500 ms after picture offset. Naming latencies were measured from the onset of the picture's presentation by means of a voice-key. The presentation of stimuli was controlled by the program PsyScope (Cohen, MacWhinney, Flatt and Provost, 1993) and was run on a Macintosh computer. The session lasted approximately 30 minutes.

Data analysis

Three types of responses were scored as errors: a) production of names that differed from those designated by the experimenter; b) verbal dysfluencies (stuttering, utterance repairs, production of nonverbal sounds that triggered the voice key); and c) responses longer than 1900 ms and shorter than 300 ms. Recording failures were also scored as errors. We excluded naming latencies for one picture from the different-gender set (*cornice* 'frame') because of the high naming inconsistency observed during the familiarization phase. In order to keep the design balanced, we also excluded naming latencies for its matched word in the same-gender set (*carota* 'carrot'). Naming latencies above and below three standard deviations from each participant's mean were also excluded. Naming latencies and error rates were submitted to an ANOVA with one between-subjects variable ('Group of participants': bilingual vs. monolingual) and two within-subjects variables ('Picture set': same-gender vs. different-gender, and 'Repetition': first, second and third).

Results and discussion

Error rates accounted for 6.6% and 5.0% of the data points for the bilingual and the monolingual groups, respectively (see Table 1). In the error analysis, the main effect of the variable 'Group of participants' was significant only in the item analysis ($F1$ (1, 18) = 1.21, MSE = 10.39, $p > .29$; $F2$ (1, 76) = 4.30, MSE =.76, $p < .04$). The main effect of the variable 'Repetition' was significant ($F1$ (2, 36) = 11.98, MSE = 1.89, $p < .01$; $F2$ (2, 152) = 10.63, MSE = .55, $p < .01$). All the other main effects and relevant interactions were not significant (all ps > .3)

In the analysis of naming latencies the main effects of the variables 'Group of participants' and 'Repetition' were significant ($F1$ (1, 18) = 25.23, MSE = 30172.70, $p < .01$; $F2$ (1, 76) = 1093.84, MSE = 2743.57, $p < .01$, and $F1$ (2, 36) = 13.34, MSE = 1413.59, $p < .01$; $F2$ (2, 152) = 41.53, MSE = 1802.78, $p < .01$, respectively) revealing that monolingual speakers named the pictures faster than bilingual speakers, and that naming latencies speed up with repetitions. Importantly, the main effect of the variable 'Picture set' was not significant ($F1$ (1,18) = 1.65, MSE = 317.30, $p > .21$, $F2 < 1$). Crucial for our purposes here is the lack of an interaction between the variables 'Group of participants' and 'Picture set' (both Fs < 1), revealing that the difference between the two sets of pictures was present for monolingual (3 ms) and bilingual (3 ms) speakers. None of the other relevant interactions were significant (all Fs < 1).

Table 1. *Naming latencies (Mean), error rates (E%) and standard deviations (SD) broken by 'Group of participants' and 'Picture set' in Experiment 1. The three repetitions are collapsed. Naming was performed in Italian.*

	Monolingual speakers			Bilingual speakers		
Picture set	Mean	SD	E%	Mean	SD	E%
Same-gender	632	77	5.4	791	62	6.6
Different-gender	629	71	4.6	788	74	6.9
Difference	3			3		
Low frequency	638	75		800	69	
High frequency	622	75		779	67	
Difference	16			21		

These results reveal that participants' naming latencies are independent of whether the target and its translation had the same gender value, suggesting that gender retrieval in L2 is not affected by the gender values of the corresponding L1 translations. However, this conclusion is based on a null result (the lack of difference between the two sets of pictures for both groups of participants), and therefore it is important to show that our experimental design is powerful enough to reveal significant differences related to the pictures used in the experiment. With this objective, a post-hoc analysis in which we re-arranged the naming latencies according to their word frequencies in Italian (a high-frequency and a low-frequency group) was conducted. Importantly, the distribution of same- and different-gender words in the high- and low-frequency groups was even, equating therefore any possible effect of the gender match or mismatch across languages. However because of this, there was some overlap between the frequencies of the words included in the two groups. Nevertheless, the mean frequency value of the words included in the high-frequency group (average = 81, range 7–348) was significantly different from that of the low-frequency group (average = 9; range 0–41; $p < .05$). Participants named the high-frequency words faster than the low-frequency words ($F1$ (1, 18) = 28.67, MSE = 367.13, $p < .01$; $F2$ (1, 76) = 3.34, MSE = 10700.51, $p < .07$). Furthermore, the magnitude of the frequency effect was

similar for both groups of participants, as revealed by the non-significant interaction between the variables 'Word frequency' and 'Group of participants' (both Fs < 1). Thus, our experiment seems to be sensitive enough to reveal word frequency effects, even though the two sets of words had reduced and overlapping range of frequencies.

So far, the results support the notion that access to the gender feature of one language is independent of the gender value of the translation word in the non-response language. However, this conclusion is based on a null result, and therefore it is important to further explore other experimental situations in which the difference between same- and different-gender pictures may arise. Experiments 2 and 3 aim at exploring this issue.

In Experiment 2, we address whether a difference between the two sets of pictures may have been masked by the slow naming latencies produced by the bilingual speakers. Bilingual participants were, on average, much slower than monolingual participants. Although this is the usual pattern when comparing L1 vs. L2 naming performance (e.g., see for example, Kroll and Stewart, 1994; Meuter and Allport, 1999; Costa et al., 2000a), it is possible that the slow naming latencies mask the detection of a gender mismatch effect.

In Experiment 3, we include two modifications to the design of Experiment 1, in order to maximize the probability of detecting a difference between the two sets of pictures. First, unlike in Experiment 1, in Experiment 3 participants were asked to name the pictures in the two languages, changing from being in the so-called 'monolingual mode' to being in the so-called 'bilingual mode'. It is possible that the probability of detecting effects of the non-response language on the response language increases when the two languages are being used during the experiment (the so-called 'bilingual mode'; see Grosjean, 1998a, b for a discussion of this issue). In the bilingual experimental context, participants cannot ignore the lexical activation of the non-response language (in some trials that language is the response-language), increasing therefore the probability of observing an interaction between the gender systems of the two languages.

The second difference between experiments is the type of NP produced by the participants. In Experiment 1, participants were asked to produce NPs in which the gender of the Italian noun surfaces in the determiner (*il* for masculine and *la* for feminine). However, Croatian has no determiners, and therefore the corresponding grammatical structure in Croatian does not require, in principle, the retrieval of the noun's gender value. It is then possible that an effect of the gender system of the non-response language only arises when the corresponding utterance format in the non-response language also requires the selection of the noun's gender value. If that were to be the case, in the context of determiner + noun NP naming, the gender of the Croatian words would be irrelevant because in the corresponding NP naming situation in Croatian, the retrieval of the gender feature is not necessary.[4] In order to test this possibility in Experiment 3, participants were asked to produce NPs of the type determiner + adjective + noun. This format requires gender selection in the response language (Italian) as well as in the corresponding utterance format in the non-response language.

Experiment 2: Does the gender of the L1 words affect L2 production? Speeded-naming task

The only difference between this experiment and Experiment 1 is that here participants were asked to name the pictures before a sound (a beep) was played. The beep acted as a deadline that participants were encouraged to meet. This procedure should speed up participants' responses, and would increase the sensitivity of the experiment to reveal any systematic difference between same- and different-gender pictures.

Participants

Ten native speakers of Croatian recruited from the same population as in Experiment 1 took part in the experiment. None had participated in Experiment 1 (see Appendix A).

Materials and procedure

The same materials, design and procedure as in the previous experiment were used. The only difference is that an acoustic signal (a short beep) was presented 800 ms after the onset of the picture presentation.

Results and discussion

The same criteria as in Experiment 1 were used here for scoring errors, leading to the exclusion of 9.4% of the data points (see Table 2). In the first analysis we included all the correct responses, even those that did not meet the deadline. The only significant difference in the error analysis was observed for the variable 'Repetition' ($F1$ (2, 18) = 4.32, MSE = 2.91, $p < .03$; $F2$ (2, 152) = 5.27, MSE = .61, $p < .01$).

In the analysis of naming latencies, the main effect of the variable 'Repetition' was significant ($F1$ (2, 18) = 14.91, MSE = 599.59, $p < .01$; $F2$ (2, 152) = 22.04, MSE = 1652.98, $p < .01$). No other significant differences were observed ('Picture set' $F1$ (1, 9) = 2.24, MSE = 341.24, $p > .17$; $F2 < 1$; 'Picture Set' × 'Repetition' both Fs < 1). Following the same post-hoc analysis as that

[4] Note that for this explanation to work one must assume that the lexical nodes of the words in the non-response language only send activation to their corresponding gender features, when the NP format in that language requires the selection of the gender feature.

Table 2. *Naming latencies (Mean), error rates (E%) and standard deviations (SD) and 'Picture set' in Experiment 2. The three repetitions are collapsed. Naming was performed in Italian.*

	Bilingual speakers		
Picture set	Mean	SD	E%
Same-gender	663	85	8.7
Different-gender	670	78	10.3
Difference	−7		
Low frequency	681	90	
High frequency	654	74	
Difference	27		

performed in Experiment 1 to explore word frequency, we observed that naming latencies were slower for low-frequency words than for high-frequency words ($F1$ $(1, 9) = 16.80$, $MSE = 653.58$, $p < .01$; $F2$ $(1, 76) = 5.10$, $MSE = 8144.76$, $p < .02$).

In a further analysis we focused only on those responses that were given before the deadline (77% of responses were given before the 800 ms deadline). The only significant difference was obtained for the variable 'Repetition' ($F1$ $(2, 18) = 17.94$, $MSE = 206.94$, $p < .01$; $F2$ $(2, 152) = 11.34$, $MSE = 901.37$, $p < .01$; all other comparisons Fs < 1).

The results of this experiment replicate those of Experiment 1: while there is a reliable effect of word frequency on naming latencies, they were independent of whether the Croatian translations had the same- or a different-gender from that of the Italian target name. Importantly, however, the average naming latency in this experiment resembles more that of the monolingual than that of the bilingual group of Experiment 1. Thus, even when participants are asked to name the pictures quite fast, no effects of the translation's gender value are observed.

Experiment 3: A further test of the autonomy of the gender systems: Mixed-language naming

In this experiment Croatian-Italian bilingual participants were asked to name the same pictures as those used in the previous experiments, but with two major differences. First, we included filler pictures which participants were asked to name in Croatian rather than in Italian. The language in which a given picture had to be named was indicated by means of its color (black in Croatian, red in Italian). This mixed-language design should increase the probability of detecting any influence of the non-response language on the response language. Second, participants were instructed to name the pictures by means of a noun phrase of the form determiner + adjective + noun (*la mia mela*, literally, 'the my apple'), in which the adjective corresponded always to the gender-marked possessive adjective 'my' (*mio* and *mia*, for masculine and feminine nouns, respectively). In this case, the gender of the noun surfaces both in the production of the determiner (*il* vs. *la* for masculine and feminine nouns, respectively) and in the production of the possessive adjective (*mio* vs. *mia* for masculine and feminine nouns, respectively). The Croatian equivalents of the Italian NPs take the form adjective + noun. Crucially for our purposes here, the possessive adjective 'my' is gender-marked in Croatian (*moj* and *moja*, for masculine and feminine nouns, respectively). Therefore, the production of adjective + noun NPs in Croatian entails the retrieval of the noun's gender.

Method

Participants

Ten native speakers of Croatian from the same population as in the previous experiments participated in this experiment (see Appendix A). None had participated in the previous experiments.

Materials

The same experimental materials as in the previous experiments plus an additional set of 40 filler pictures were selected. These filler pictures were presented in red, and participants were asked to name them in Croatian by means of an adjective + Noun NP. Half of the additional pictures corresponded to masculine nouns and half to feminine nouns in Croatian. Furthermore, half of them had the same gender in Italian and Croatian, and half of them had different genders. The stimuli were presented in 3 blocks of 125 trials (80 experimental trials from the Italian sets, 40 experimental trials from the Croatian set, and five warm-up fillers). Each picture appeared once per block. Block trials were randomized with the following restrictions: a) no more than four pictures from the Italian sets in a row, b) no more than two pictures from the Croatian set in a row, and c) trials of the same gender appeared in no more than three consecutive trials. Furthermore, in those cases in which a picture to be named in Croatian was followed by a picture to be named in Italian, half of the times the later picture belonged to the same-gender set and half of the times to the different-gender set. In other words, when there was a language switch from Croatian to Italian, the probability of encountering a same- or a different-gender picture was the same. In addition, care was taken to avoid any semantic or phonological overlap on consecutive trials. Six different block orders were constructed, and similar numbers of participants were randomly assigned to each block order.

Table 3. *Naming latencies (Mean), error rates (E%) and standard deviations (SD) and 'Picture set' in Experiment 3. The three repetitions are collapsed. Naming was performed in Italian.*

	Bilingual speakers		
Picture set	Mean	SD	E%
Same-gender	884	105	7.1
Different-gender	879	106	7.7
Difference	5		
Low frequency	898	96	
High frequency	865	113	
Difference	33		

Procedure

Participants were instructed to use the color of the picture as language cue. Before the experiment proper, participants were presented with the entire set of pictures and they were asked to name the pictures with the determiner + adjective + noun NP format in Italian and the adjective + noun NP format in Croatian. All other details were identical to Experiment 1.

Results and discussion

Following the same criteria as in Experiment 1, 7.3% of the data-points were excluded (see Table 3). The main effect of the variable 'Repetition' was significant both in the error rates ($F1$ (2, 18) = 14.49, $MSE = 1.01$, $p < .01$; $F2$ (2, 152) = 6.51, $MSE = .57$, $p < .01$) and in the naming latencies ($F1$ (2, 18) = 66.10, $MSE = 703.57$, $p < .01$; $F2$ (2, 152) = 60.57, $MSE = 3020.10$, $p < .01$). All the other comparisons were not significant (both Fs < 1). Following the post-hoc analysis carried out in the previous experiments, low-frequency words appeared to be named significantly more slowly than high-frequency words ($F1$ (1, 9) = 19.09, $MSE = 819.03$, $p < .01$; $F2$ (1, 76) = 4.87, $MSE = 14235.95$, $p < .03$). As in the previous experiments, naming latencies were affected by word frequency, but they were independent of whether the Croatian translation of the Italian picture's name has the same or different gender.

Analyzing the results of the three experiments together, neither the main effect of 'Picture Set' (both Fs < 1) nor its interaction with the factor 'Experiment' ($F1$ (2, 27) = 1.41; $MSE = 510.69$, $p > .26$; $F2 < 1$) were significant, indicating that comparable patterns of results were observed across experiments. In contrast, naming latencies were faster for high-frequency words than for low-frequency words ($F1$ (1, 27) = 61.95, $MSE = 526.69$, $p < .01$; $F2$ (1, 76) = 5.91, $MSE = 24236.67$, $p < .01$), an effect that did not interact with the variable 'Experiment' ($F1$ (2, 27) < 1, $F2$ (2, 152) = 1.10; $MSE = 3740.25$, $p > .33$). This reveals that the magnitude of the frequency effect is comparable across languages. Together these results reveal that when producing Italian gender-marked NPs, the gender values of the Croatian translations do not affect bilinguals naming performance, supporting the notion that the two gender systems of a bilingual speaker are autonomous.

In the following two experiments we further explore the existence of cross-language gender effects in another type of bilingual populations: bilinguals whose two languages are from the same language family and whose gender structure is symmetrical.

Experiments 4A and 4B: Does the similarity between the two gender systems affect the organization of the bilingual's gender systems?

In Experiments 4A and 4B we explore whether cross-language gender interference arises when the two languages of a bilingual have the same structural composition. One could argue that our failure to observe across-language gender interference in the previous experiments stems from the fact that the gender systems of Croatian and Italian are different enough to prevent their integration into a common system. This argument is predicated on the fact that the Croatian gender system has three gender values while the Italian gender system has only two.

The gender systems of the languages used in Experiment 4A (Spanish-Catalan) and in Experiment 4B (Italian-French) are very similar, in the sense that a) they have only two gender values (feminine and masculine) and b) the morphological implications of the noun's gender value in the two languages are similar. For example, in all these languages and in the context of NP production, the noun's gender value determines, among other things, the determiner form and the inflections of adjectives in all four languages.

Experiment 4A: NP naming by Catalan-Spanish bilinguals in Spanish

In this experiment, two groups of highly-proficient Catalan-Spanish bilinguals were asked to name pictures by means of simple determiner + noun NPs in Spanish. The names of the pictures could either have the same or different genders in the two languages. Participants in Group 1 had Spanish as their first and dominant language, and participants in Group 2, Catalan. Thus, participants in Group 1 named the pictures in their L1 (Spanish) and participants in Group 2 named them in their L2 (Spanish). Also, another group, of Spanish monolingual participants, were asked to name the pictures (Group 3). In

order to maximize the possibility of detecting an effect in this experiment, we increased the number of participants included in each group.

The predictions parallel those of Experiment 1. If the gender value of the picture's name translation has an effect on naming latencies, then one should expect an interaction between the factors 'Picture set' and the 'Monolingual/bilingual status of the participants'. This experiment also allows us to explore whether a possible effect of the variable 'Picture set' is modulated by whether the bilingual speakers name the pictures in their L1 or L2.

Method

Participants

Seventy-two participants were included in the experiment (see Appendix B). Participants were evenly distributed between three groups (Spanish-Catalan bilinguals, Catalan-Spanish bilinguals and Spanish monolinguals).

Materials and procedure

Given the extensive similarity between Spanish and Catalan words, it was difficult to find words that had different genders in the two languages and whose referents could be easily depicted. A total of 22 words per picture set, matched in number of syllables (2.5 vs. 2.7 for different- vs. same-gender pictures) and word frequency (190 vs. 196, per five million, respectively; according to the LEXESP corpus, Sebastián-Gallés, Marti, Cuetos and Carreiras, 2000) were selected. Each set contained the same number of masculine and feminine names, and the same number of cognate (7 out of 22) and non-cognate words (15 out of 22) (see Appendix E). All the other details were identical to those of Experiment 1 with, the following exception. After the experiment session we asked all bilingual participants to name the pictures in the language they did not use during the experiment. Thus, since the language used in the experiment session was Spanish, afterwards we asked them to name the pictures in Catalan. We did so to make sure that the names they spontaneously used for the pictures in the non-response language (Catalan) were the same as those intended by the experimenter. In this way, we can be certain that the target's translation preferred by each participant corresponded to the one used to determine the gender match/mismatch. All the data points in which a given bilingual speaker named the pictures in Catalan with a name different than expected were removed from the analysis (7.4% in Group 1 and 2.8% in Group 2.).

Results and discussion

Following the same criteria as in Experiment 1, 2.5%, 3.5% and 8.1% of the responses for Group 1, 2 and 3, respectively, were removed. Error rates revealed a main effect of 'Repetition' ($F1\ (2,\ 138) = 6.82$; $MSE = 21.46$; $p < .01$, $F2\ (2,\ 84) = 6.18$; $MSE = 21.99$; $p < .01$) and 'Group of participants' ($F1\ (2, 69) = 25.04$, $MSE = 50.93$, $p < .01$; $F2\ (2,\ 84) = 42.59$; $MSE = 27.65$; $p < 01$). The main effect of 'Picture set' was not significant (both Fs < 1). None of the interactions was significant (all Fs < 1).

In the analysis of naming latencies the main effects of the factors 'Repetition' and 'Group of participants' were significant ($F1\ (2, 138) = 31.87$, $MSE = 2252.73$, $p < .01$; $F2\ (2,\ 84) = 56.67$, $MSE = 1189.15$, $p < .01$); and $F1\ (2,\ 69) = 7.72$, $MSE = 33737.05$, $p < .02$; $F2\ (2,\ 84) = 57.3$, $MSE = 2060.18$, $p < .01$, respectively). The main effect of the factor 'Picture set' was significant only in the subject analysis ($F1\ (1,\ 69) = 46.70$, $MSE = 1129.08$, $p < .01$; $F2\ (1,\ 42) = 1.83$, $MSE = 25275.16$, $p > .18$), revealing that naming latencies were slightly slower for different- than for same-gender picture names. However and crucially, the interaction between 'Group of participants' and 'Picture set' was far from significant (both Fs < 1), revealing that the difference between same- and different-gender pictures was similar for the three groups of subjects (23 ms, 24 ms and 19 ms, for each group respectively). This result suggests that the difference between the two picture sets is not related to the bilingual status of the participants (see Table 4).

Following the same post-hoc analysis as in Experiment 1, words with high-frequency values were named faster than words with low-frequency values ($F1\ (1, 69) = 98.35$, $MSE = 1736.15$, $p < .01$; $F2\ (1, 42) = 5.55$, $MSE = 28717.70$, $p < .02$). The magnitude of the frequency effect was similar between the three groups of participants ('Group of participants' × 'Word frequency'; both Fs < 1).

The results of this experiment fully replicate those of Experiment 1: monolingual and bilingual speakers show the same pattern of performance when naming the experimental pictures. That is, the fact that the gender value of the target's translation is the same as or different from that of the target word does not seem to affect naming latencies.

Experiment 4B: NP naming by Italian-French bilinguals in French

Method

Participants

Twenty participants took part in the experiment. Half were native speakers of French, and the other half were highly-proficient Italian-French bilinguals (see Appendix C).

Materials and procedure

Sixty pictures were selected, half whose names have the same gender value across languages, and the other half whose names have different gender values across

Table 4. *Naming latencies (Mean), error rates (E%) and standard deviations (SD) broken by 'Group of participants' and 'Picture set' in Experiment 4A. The three repetitions are collapsed. Naming was performed in Spanish.*

	Group 1			Group 2			Group 3		
	Spanish-Catalan bilinguals			Catalan-Spanish bilinguals			Monolinguals		
Picture set	Mean	SD	E%	Mean	SD	E%	Mean	SD	E%
Same-gender	698	76	2.1	746	79	3.4	685	67	8.5
Different-gender	721	80	3.0	770	87	3.9	704	66	7.6
Difference	−23			−24			−19		
Low frequency	728	76	3.3	782	91	4.6	713	71	9.6
High frequency	692	81	1.8	735	78	2.6	677	62	6.6
Difference	36			47			37		

languages (see Appendix F). Only non-cognates were included in the experiment. The two sets were matched in number of syllables (1.7 vs. 1.78, for different- vs. same-gender pictures) and word frequency (19 vs. 17, counts per million, according to the Brulex database, Content, Mousty and Radeau, 1990). Each set contained the same number of masculine and feminine names, and all the picture names started with a consonant. All the other details were identical to those of Experiment 4A.

Results and discussion

The analysis of this experiment follows the same criteria as in Experiment 4A. If after the experiment a bilingual participant used a different name in Italian than that expected by the experimenter, all the data points produced by that speaker for that stimulus were removed (5.8%). In addition, we removed a total of 6.7% data points for the two groups of subjects following the same criteria as in Experiment 1 (see Table 5)

In the analysis of error rates, the main effect of 'Repetition' was significant only in the analysis by items ($F1(2, 36) = 1.6$, $MSE = 67.8$, $p = .2$; $F2$ $(2, 116) = 3.9, MSE = 211.9, p < .05$). None of the other main effects was significant (all Fs < 1). The only significant interaction was than between the 'Repetition' and 'Group of participants' ($F1$ $(2, 36) = 4.4$, $MSE = 183.8$, $p < .05$; $F2\,(2, 116) = 11.6, MSE = 536.9, p < .001$). The variables 'Repetition' and 'Picture set' did not interact ($F1$ $(2, 36) = 3.1$, $MSE = 41.9$, $p = .06$; $F2$ $(2, 116) = 2.3$, $MSE = 125.8$, $p = .1$).

In the analysis of naming latencies, the main effects of the factors 'Repetition' and 'Group of participants' were significant ($F1$ $(2, 36) = 34.2$, $MSE = 1963$, $p < .001$; $F2$ $(2, 116) = 30.9$, $MSE = 7134$, $p < .001$; and $F1$ $(1, 18) = 20.0$, $MSE = 42332$, $p < .001$; $F2$ $(1, 58) = 385.1$, $MSE = 6807$, $p < .001$, respectively). The main effect of the factor 'Picture set' was significant only in the participants analysis ($F1\,(1, 18) = 5.27, MSE = 1321, p < .05$; $F2 < 1$). Importantly, this variable did not interact with the 'Group of participants' (both Fs < 1) indicating that the difference between the two sets of pictures was present both for bilingual and monolingual speakers. The three-way interaction between 'Picture set', 'Repetition' and 'Group of participants' appeared marginally significant in the participants analysis ($F1$ $(2, 36) = 3.01$, $MSE = 1391$, $p = .06$), and significant in the item analysis ($F2\,(2, 116) = 4.8, MSE = 2535, p < .01$). The other interactions did not reach significance (Fs < 1).

Following the same post-hoc analysis as in Experiment 1, high-frequency words were named faster than low-frequency words, which was the effect significant in the participants analysis ($F1$ $(1, 18) = 18.9$, $MSE = 526$, $p < .001$), and marginal in the item analysis ($F2\,(1, 58) = 2.3$, $MSE = 10639$, $p = .09$). The magnitude of the frequency effect was found to be stronger for Italian bilinguals than for French monolinguals, as attested by

Table 5. *Naming latencies (Mean), error rates (E%) and standard deviations (SD) broken by Group of Participants and Picture set in Experiment 4B. The three repetitions are collapsed. Naming was performed in French.*

	Monolinguals			Bilinguals		
Picture set	Mean	SD	E%	Mean	SD	E%
Same-gender	645	101	3.2	809	65	3.7
Different-gender	656	100	3.1	828	69	3.9
Difference	11			19		
Low frequency	657	100		845	62	
High frequency	642	102		797	70	
Difference	15			48		

the significant interaction between 'Group of participants' and 'Word frequency' ($F1$ (1, 18) = 5.28, $MSE = 526$, $p < .05$; $F2$ (1, 58) = 4.47, $MSE = 2141$, $p < .05$).

The results of this experiment are similar to those obtained in the previous experiments. Although in this experiment pictures whose names had different gender values across languages were named slightly more slowly than those that had the same gender values, such a difference was present both for bilingual and monolingual participants. Thus, such a difference cannot be attributed to the different gender values of picture names across languages, since such a variable is not relevant for monolingual speakers.

Together, the results of Experiments 4A and 4B suggest that the retrieval of the target's gender feature is not affected by the gender value of the target's translation, even for those bilinguals whose two languages have a structurally symmetrical gender system.

General discussion

The experiments reported in this article were designed to explore the interaction between the gender systems of the two languages of a bilingual. The question we addressed was the extent to which the gender values of the translation words in the non-response language affects a bilingual's naming performance in the response language. Bilingual speakers of two gender-marked languages were asked to name two sets of pictures in their L2 by means of gender-marked NPs. One set of pictures had L2 names whose L1 translations had the same gender value (either feminine or masculine), while the pictures included in the other set had L2 names whose L1 translations had different gender values. We argued that if the gender values of the translation words in the non-response language were to affect the performance in the response language, then one might expect bilinguals (but not monolinguals) to perform differently with the two sets of pictures.

In Experiment 1, Croatian-Italian speakers were asked to produce NPs in their L2 (Italian). Naming latencies for pictures whose translations had the same gender as the target names and for pictures with different-gender translations were statistically similar. Importantly, the very small difference between the two sets was present both for bilingual and Italian monolingual speakers, indicating that there is no effect whatsoever of the gender properties of the non-response language. In Experiment 2, we asked whether the lack of an effect in Experiment 1 was due to the bilinguals' slow naming latencies. In this experiment, participants were asked to perform a speeded-naming task. As a result they named the pictures on average 200 ms faster than in Experiment 1, but again no significant differences between the two sets of pictures were observed. In Experiment 3, we increased the probability of detecting an interaction between the two gender systems by a) asking participants to perform a mixed-language naming task, and b) using NPs that require gender access in Italian as well as in the corresponding Croatian NPs. The results of Experiment 3 replicated and confirmed our previous observations: no significant differences were observed between same- and different-gender pictures. Experiments 4A and 4B explored whether cross-language gender effects can be found when the gender systems of the two languages of a bilingual are similar. Thus, bilingual speakers of various Romance languages (Spanish-Catalan bilinguals and Italian-French bilinguals) were asked to perform an NP naming task. In both experiments, same-gender picture names were produced slightly faster that different-gender picture names. However, such a difference was also present when Spanish and French monolingual speakers were asked to perform the task, and therefore such a difference cannot be attributed to the gender value of the target's translations. Importantly, and despite the failure to observe any systematic difference between same- and different-gender pictures, our design was sensitive enough to detect lexical effects. In all experiments, the post-hoc analyses revealed that high-frequency words were named faster than low-frequency words.

The results of these experiments challenge the notion that the gender value of the target's translation in the non-response language affects gender access in the response language. The speed and ease with which the NP naming task was performed was independent of whether the target word and its translation had the same or different genders. In the introduction to this paper, we described two different views about how the gender systems of two languages may be represented in the mind of bilingual speakers: the shared (or integrated) view, and the autonomous (or independent) view. According to the first view, the gender systems of the bilingual's two languages are shared (or interconnected) in such a way that when a target word and its translation have the same gender value, the corresponding gender feature would be activated from two sources (e.g., the L2 target word and its translation word in L1), since the two words share the same gender feature (see Figure 1, panel A). Alternatively, when the words of the two languages had different gender values they would point to different gender features (see Figure 1, panel B). According to the autonomous view, the gender systems corresponding to the two languages of a bilingual are independent. On this account, whether or not a given word and its translation have the same or different gender values does not have any representational or functional implications (see Figure 2).

We further argued that, in order to derive predictions from these two views, we had to consider another dimension, namely how gender features are retrieved. We considered two possibilities. On the assumption that the retrieval of the gender feature depends, to some extent,

on its level of activation at the moment of selection, then the integrated view would predict faster naming latencies for words with the same gender values across languages than for words with different gender values. That is, the gender values of the words in the non-response language were expected to affect the processes involved in the response language. Alternatively, in the case of independent gender systems, or in the case of automatic gender selection, no difference between same- and different-gender translations was expected. In other words, only one out of the four possible combinations of assumptions predicted a difference between the two sets of pictures (see Figure 3).

Given that naming latencies are independent of the target's translation gender values, we can reject at least one combination of assumptions – the one in which a difference between the two sets of pictures was expected. That is to say, a model that holds simultaneously the integrated assumption and the selection by activation-levels assumption can be rejected, since it predicts a difference between the two sets of words that was not observed in our experiments. Thus, a model in which gender retrieval is sensitive to activation-levels has to drop the assumption about the shared grammatical gender system. According to several models of speech production, this would seem to be a reasonable step to take. However, whether or not one takes this step crucially depends on the interpretation of the available experimental data on gender priming. This is because, although there are several observations that seem to suggest that the selection of the gender feature is sensitive to activation-levels, the interpretation of such data is also consistent with other views regarding gender selection (see, for example, Caramazza et al., 2001). The other possibility is to drop the selection by activation-levels assumption while keeping the notion of an integrated gender system across languages. On this view, gender access is an automatic process in which the gender values of the words belonging to the non-response language do not affect the eventual selection of the target gender feature, and therefore no differences between same- and different-gender pictures should be observed.

There is a third possibility, which is to give up both assumptions and assume that gender retrieval is an automatic process and that the two gender systems of a bilingual are independent. In fact, if one assumes that gender access is an automatic process that occurs upon the selection of a given lexical node, it seems reasonable to assume that the only gender value that plays a role is that of the selected lexical node, rendering the notion of independent or integrated gender systems irrelevant. At this point, we cannot adjudicate among these possible solutions. At any rate, whichever combination of assumptions turns out to be the correct one, what is important for our purposes here is the fact that the gender values of the words in the non-response language do not affect performance in the response language. Therefore, at this point it seems reasonable to conclude that the gender properties of one language do not affect gender processing in the other language.

The predictions tested above were based on the assumption that the two lexicons of a bilingual are activated during speech production. However, a model in which the semantic system does not activate the non-response language during speech production would also account for the results reported here. If we were to assume that only one lexicon is activated during speech production, it follows that the linguistic properties of the lexical nodes of the non-response language cannot affect processing, since such a language is not activated. Even if this type of models could explain the results reported here, the existence of an interaction between the two languages of a bilingual at other levels of representation (e.g., Hermans et al., 1998; Gollan and Acenas, 2000; Costa, Colomé, Gómez and Sebastián-Gallés, 2003) seems to be inconsistent with this view.

Before concluding, it is important to mention that the results reported in this article come from the naming performance of a population of highly-proficient bilinguals. It is possible that the degree of language autonomy of the two gender systems of a bilingual speaker depends on the degree of L2. It may be the case that the less proficient a bilingual speaker is the greater the interaction between the gender systems. Future research is needed to address the impact of these variables, among others, in the role of the non-response language during speech production.

To conclude, the results of this study reveal that bilingual naming performance is independent of the gender value of the words in the non-response language. The extent to which this language autonomy can also characterize other domains of bilingual lexical access remains to be explored. Certainly, the relative invulnerability of the response language to the properties of the non-response language does not seem to apply at other levels of representation in which the interaction between the two languages of a bilingual seems to be present, as, for example, the phonological level and, maybe, the syntactic. Further research is needed to establish whether other processes involved in speech production (e.g., word order) are also unaffected by the properties of the non-response language.

References

Alario, F. X. & Caramazza, A. (2002). The production of determiners: Evidence from French. *Cognition*, 82, 179–223.

Andersen, R. (1984). What's gender good for anyway? In R. Andersen (ed.), *Second languages: A cross-linguistic perspective*, pp. 77–99. Rowley, MA: Newbury House.

Barry, C., Morrison, C. M. & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency, and name agreement. *Quarterly Journal of Experimental Psychology*, 50, 560–585.

Caramazza, A., Miozzo, M., Costa, A., Schiller, N. & Alario, F. X. (2001). 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.

Carroll, S. (1989). Second language acquisition and the computational paradigm. *Language Learning*, 39, 535–594.

Carroll, S. (1999). Input and SLA: Adult's sensitivity to different sorts of cues for French gender. *Language Learning*, 49, 37–92.

Cohen, J. D., MacWhinney, B., Flatt, M. & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, and Computers*, 25, 257–271.

Colomé, A. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory & Language*, 45, 721–736.

Content, A., Mousty, P. & Radeau, M. (1990). Brulex, une base de données lexicales informatiséepour le français écrit et parlé. *L'Année Psychologique*, 90, 551–566.

Corbett, G. (1991). *Gender*. Cambridge: Cambridge University Press.

Costa, A. (in press). Lexical access in bilingual production. In Kroll & De Groot (eds.).

Costa, A. & Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish-English and English-Spanish bilinguals. *Bilingualism: Language and Cognition*, 2, 231–244.

Costa, A., Caramazza, A. & Sebastián-Gallés, N, (2000a). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1283–1296.

Costa, A., Colomé, A. & Caramazza, A. (2000b). Lexical access in speech production: The bilingual case. *Psicológica*, 3, 403–435.

Costa, A., Colomé, A., Gómez, O. & Sebastián-Gallés, N. (2003). Another look at cross-language competition in bilingual speech production: Lexical and phonological factors. *Bilingualism: Language and Cognition*, 6.3. 167–179.

Costa, A., Kovacic, D., Fedorenko, E. & Caramazza, A. (in press). The gender congruency effect and the selection of free-standing morphemes: The case of adjectives and pronouns in Croatian. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Costa, A., Miozzo, M. & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language*, 41, 365–397.

Costa, A., Sebastián-Gallés, N. Miozzo, M. & Caramazza, A. (1999). The gender congruity effect: Evidence from Spanish and Catalan. *Language and Cognitive Processes*, 14, 381–391.

De Bot, K. (1992). A bilingual production model: Levelt's speaking model adapted. *Applied Linguistics*, 13, 1–24.

De Groot, A. M. & Nas, G. L. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory & Language*, 30, 90–123.

Ellis, A. W. & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 515–523.

Flege, J. E. (1999). Age of learning and second language speech. In D. Birdsong (ed.), *Second language acquisition and the Critical Period Hypothesis* (Second Language Acquisition Research), pp. 101–131. Mahwah, NJ: Lawrence Erlbaum Associates.

Gollan, T. H. & Acenas, L. A. (2000). Tip-of-the-tongue incidence in Spanish-English and Tagalog-English bilinguals. Paper presented at The 3rd International Symposium on Bilingualism, Bristol, England.

Gollan, T. H. & Acenas, L. A. (in press). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Granfeldt, J. (in press). The development of gender attribution and gender agreement in French: A comparison of bilingual first and second languages learners'. In J. Dewaele (ed.), *Focus on French as foreign language: Multidisciplinary approaches*. Clevedon: Multilingual Matters.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67–81.

Grosjean, F. (1998a). Transfer and language mode. *Bilingualism: Language and Cognition*, 1, 175–176.

Grosjean, F. (1998b). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, 1, 131–149.

Hawkins, R. & Chan, C. (1997). The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research*, 13, 187–226.

Hermans, D., Bongaerts, T., de Bot, K. & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language. *Bilingualism: Language and Cognition*, 1, 213–230.

Janssen, N. (1999). Bilingual word production: The time course of lexical activation in a mixed language context. Ms., University of Nijmegen.

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.

Kroll, J. F. & De Groot, A. M. B. (eds.) (in press), *Handbook of bilingualism: Psycholinguistic approaches*. New York: Oxford University Press.

Kroll, J. F., Dijkstra, A., Janssen, N. & Schriefers, H. (2000). Selecting the language in which to speak: Experiments on lexical access in bilingual production. Paper presented at The 41st Annual Meeting of the Psychonomic Society, New Orleans, LA.

Kroll, J. F. & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33, 149–174.

La Heij, W. (in press). Monolingual and bilingual lexical access in speech production: Issues and models. In Kroll & De Groot (eds.).

La Heij, W., Hooglander, A., Kerling, R. & van der Velden, E. (1996). Nonverbal context effects in forward and backward translation: Evidence for concept mediation. *Journal of Memory and Language*, 35, 648–665.

La Heij, W., Mak, P., Sander, J. & Willeboordse, E. (1998). The gender-congruency effect in picture–word tasks. *Psychological Research*, 61, 209–219.

Lee, M. W. & Williams, J. (2001). Lexical access in spoken word production by bilinguals: Evidence from a semantic priming paradigm. *Bilingualism: Language and Cognition*, 4, 233–248.

Levelt, W. J. M. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences*, 98, 13464–13471.

Meuter, R. F. & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory & Language*, 40, 25–40.

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., Costa, A. & Caramazza, A. (2002). The gender congruency effect in Romance languages: Further evidence from its time course in Italian and Spanish. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 388–391.

Pallier, C., Colomé, A. & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Concrete exemplar-based vs. abstract lexical entries. *Psychological Science*, 12, 445–449.

Potter, M. C., So, K.-F., von Eckhardt, B. & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and more proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, 23, 23–38.

Poulisse, N. (1997). Language production in bilinguals. In A. M. B. de Groot & J. F. Kroll (eds.), *Tutorials in bilingualism: Psycholinguistic perspectives*, pp. 201–224. Mahwah, NJ: Lawrence Erlbaum Associates.

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

Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 841–850.

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

Sebastián-Gallés, N., Martí, M. A., Cuetos, F. & Carreiras, M. (2000). *LEXESP: Base de datos informatizada de la lengua española*. Barcelona: Edicions de la Universitat de Barcelona.

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

Van Hell, J. G. & De Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, 1, 193–211.

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

Vigliocco, G., Lauer, M., Damian, M. 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.

Yeni-Komshian, G. H., Flege, J. E. & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Bilingualism: Language and Cognition*, 3, 131–149.

Received February 15, 2002
Revisions received May 6, 2003; August 20, 2003
Accepted August 21, 2003

Appendix A. *Croatian-Italian bilingual sample*

Language history and the self-evaluated proficiency scores of the Croatian-Italian bilinguals are shown. Mean age and the standard deviation (SD) are given in years. Onset of L2 (Italian) acquisition refers to the mean age (in years) at which participants started learning Italian. 'Use of L2' refers to how long (in years) participants had been using Italian regularly. The proficiency scores were obtained by a self-evaluation through a questionnaire filled out by the subjects after the experiment. The scores are on a 10-point scale, in which 10 represents native-speaker level and 1 complete ignorance of the language.

Language history	Experiment 1	Experiment 2	Experiment 3
Age (SD)	26 (3)	23 (3)	24 (4)
Onset of the L2 acquisition	5 (4)	8 (4)	9 (6)
Use of L2 and (SD) (in years)	17 (8)	9 (6)	7 (6)
Living in Italy (in years)	8 (4)	5 (3)	5 (3)

	Experiment 1		Experiment 2		Experiment 3	
Self-evaluated proficiency	Mean	SD	Mean	SD	Mean	SD
Production	9.1	.8	7.8	.9	8.3	1.1
Comprehension	9.7	.5	8.9	.9	9.1	.8
Writing	9.2	.7	7.6	.1	7.6	1.2
Reading	9.5	.7	9.0	.6	9.1	.8

Appendix B. *Spanish-Catalan and Catalan-Spanish bilingual samples*

Language history and the self-evaluated proficiency scores of the Spanish-Catalan and Catalan-Spanish bilinguals are shown. Mean age and the standard deviation (SD) are given in years. Onset of L2 acquisition refers to the mean age (in years) at which participants started learning the L2. 'Use of L2' refers to how long (in years) participants had been using the L2 regularly. The proficiency scores were obtained by a self-evaluation through a questionnaire filled out by the subjects after the experiment. The scores are on a 10-point scale, in which 10 represents native-speaker level and 1 complete ignorance of the language.

	Spanish-Catalan	Catalan-Spanish
Language history	Mean (SD)	Mean (SD)
Age (SD)	21 (2)	21 (2)
Onset of L2 acquisition	5 (2)	5 (2)
Use of L2 (SD)	15 (2)	17 (3)

	Spanish-Catalan		Catalan-Spanish	
Self-evaluated proficiency	Mean	SD	Mean	SD
Production	6.9	2.2	8.3	1.5
Comprehension	9.5	1.0	9.8	.7
Writing	8.1	1.3	9.4	1.1
Reading	9.4	1.1	9.8	.7

Appendix C. *Italian-French bilingual sample*

Language history and the self-evaluated proficiency scores of the Italian-French bilinguals are shown. Mean age and the standard deviation (SD) are given in years. Onset of L2 (French) acquisition refers to the mean age (in years) at which participants started learning French. 'Use of L2' refers to how long (in years) participants had been using French regularly. The proficiency scores were obtained by a self-evaluation through a questionnaire filled out by the participants after the experiment. The scores are on a 10-point scale, in which 10 represents native-speaker level and 1 complete ignorance of the language.

Language history	Mean (SD)
Age (SD)	26 (4)
Onset of L2 acquisition	6 (3)
Use of L2 (SD)	20 (7)
Living in a French-speaking country (in years)	7 (8)

Self-evaluated proficiency	Mean	SD
Production	8.8	1.2
Comprehension	9.3	.6
Writing	8.3	1.3
Reading	9.3	.5

Appendix D. ***Materials employed in Experiments 1, 2 and 3***

Different-gender words			Same-gender words		
Target name	Croatian translation	English translation	Target name	Croatian translation	English translation
Luna	Mjesec	Moon	Chiesa	Crkva	Church
Finestra	Prozor	Window	Camicia	Košulja	Shirt
Lingua	Jezik	Tongue	Gamba	Noga	Leg
Croce	Križ	Cross	Bandiera	Zastava	Flag
Catena	Lanac	Chain	Bottiglia	Boca	Bottle
Chiave	Ključ	Key	Torre	Kula	Tower
Barca	Čamac	Boat	Tazza	Šalica	Cup
Bicicletta	Bicikl	Bicycle	Mela	Jabuka	Apple
Spada	Mač	Spade	Candela	Svijeća	Candle
Cintura	Pojas	Belt	Scarpa	Cipela	Shoe
Foglia	List	Leaf	Gonna	Suknja	Skirt
Ruota	Kotač	Wheel	Pipa	Lula	Pipe
Noce	Orah	Nut	Scopa	Metla	Broom
Tenda	Šator	Tent	Forchetta	Viljuška	Fork
Nuvola	Oblak	Cloud	Sega	Pila	Saw
Cornice	Okvir	Frame	Carota	Mrkva	Carrot
Pentola	Lonac	Pot	Freccia	Strijela	Arrow
Sciarpa	Šal	Scarf	Fragola	Jagoda	Strawberry
Cipolla	Luk	Onion	Ciliegia	Trešnja	Cherry
Brocca	Vrč	Pitcher	Pera	Kruška	Pear
Fuoco	Vatra	Fire	Letto	Krevet	Bed
Braccio	Ruka	Arm	Cane	Pas	Dog
Libro	Knjiga	Book	Treno	Vlak	Train
Vestito	Haljina	Dress	Cavallo	Konj	Horse
Pesce	Riba	Fish	Naso	Nos	Nose
Bicchierre	Čaša	Glass	Tavolo	Stol	Table
Cestino	Košara	Basket	Castello	Dvorac	Castle
Serpente	Zmija	Snake	Ponte	Most	Bridge
Fucile	Puška	Rifle	Pozzo	Bunar	Well
Cucchiaio	Žlica	Spoon	Piatto	Tanjur	Plate
Cammello	Deva	Camel	Dito	Prst	Finger
Maiale	Svinja	Pig	Coltello	Nož	Knife
Pomodoro	Rajčica	Tomato	Cappello	Šešir	Hat
Rubinetto	Slavina	Faucet	Martello	Čekić	Hammer
Guanto	Rukavica	Glove	Pianoforte	Klavir	Piano
Casco	Kaciga	Helmet	Camino	Dimnjak	Chimney
Fungo	Gljiva	Mushroom	Pennello	Kist	Paintbrush
Peperone	Paprika	Pepper	Canguro	Klokan	Kangaroo
Calzino	Čarapa	Sock	Cannone	Top	Cannon
Sommergibile	Podmornica	Submarine	Tamburo	Bubanj	Drum

Appendix E. ***Materials employed in Experiment 4A***

Different-gender words			Same-gender words		
Target name	Catalan translation	English translation	Target name	Catalan translation	English translation
Seta	Bolet	Mushroom	Murcielago	Ratpenat	Bat
Higo	Figa	Fig	Melocoton	Pressec	Peach
Muela	Queixal	Backtooth	Cereza	Cirera	Cherry
Lechuga	Enciam	Lettuce	Calcetin	Mitjo	Sock
Buzon	Bustia	Mailbox	Pimiento	Pebrot	Pepper
Tenedor	Forquilla	Fork	Cuchara	Cullera	Spoon
Peine	Pinta	Comb	Cepillo	Raspall	Brush
Limón	Llimona	Lemon	Zanahoria	Pastanaga	Carrot
Grifo	Aixeta	Faucet	Nuez	Nou	Chestnut
Mantel	Estovalles	Tablecloth	Remo	Trompa	Trunk
Ardilla	Esquirol	Squirrel	Buho	Mussol	Owl
Rodilla	Genoll	Knee	Mariposa	Papallona	Butterfly
Paloma	Colom	Pigeon	Muñeca	Nina	Doll
Tejado	Teulada	Roof	Cebolla	Ceba	Onion
Cuna	Bressol	Cradle	Queso	Formatge	Cheese
Zapato	Sabata	Shoe	Anillo	Anell	Ring
Nube	Nuvol	Cloud	Red	Xarxa	Net
Pendientes	Arrecades	Earrings	Sombrero	Barret	Hat
Bolsillo	Butxaca	Pocket	Caja	Caixa	Box
Nariz	Nas	Nose	Dedo	Dit	Finger
Cama	Llit	Bed	Mesa	Taula	Table
Cabeza	Cap	Head	Puerta	Porta	Door

Appendix F. ***Materials employed in Experiment 4B***

Different-gender words			Same-gender words		
Target name	Italian translation	English translation	Target name	Italian translation	English translation
Bague	Anello	Ring	Bougie	Candela	Candle
Chaussette	Calzino	Sock	Cage	Gabbia	Cage
Cheminée	Camino	Chimney	Cerise	Ciliegia	Cherry
Cuiller	Cucchiaio	Spoon	Chaise	Sedia	Chair
Fusée	Razzo	Rocket	Cloche	Campana	Bell
Glace	Gelato	Ice-cream	Fraise	Fragola	Strawberry
Louche	Mestolo	Ladle	Guitare	Chitarra	Guitar
Montre	Orologio	Watch	Poêle	Pentola	Frying pan
Moustache	Baffo	Moustache	Pomme	Mela	Apple
Passoire	Colino	Colander	Poupée	Bambola	Doll
Poignée	Pugno	Handle	Queue	Coda	Tail
Poubelle	Cesto	Dustbin	Boîte	Scatola	Box
Quille	Birillo	Ninepin	Brosse	Spazzola	Brush
Robe	Vestito	Dress	Citrouille	Zucca	Pumpkin
Tomate	Pomodoro	Tomato	Toupie	Trottola	Top
Bol	Scodella	Bowl	Crabe	Granchio	Crab
Bonbon	Caramella	Sweet	Pinceau	Pennello	Paint brush

Appendix F. ***Continued***

Different-gender words			Same-gender words		
Target name	Italian translation	English translation	Target name	Italian translation	English translation
Crayon	Matita	Pencil	Sifflet	Fischietto	Whistle
Drapeau	Bandiera	Flag	Sparadrap	Cerotto	Plaster
Fauteuil	Poltrona	Armchair	Tabouret	Sgabello	Stool
Fouet	Frusta	Whip	Cadenas	Lucchetto	Lock
Gâteau	Torta	Cake	Céleri	Sedano	Celeriac
Gland	Ghianda	Acorn	Chapeau	Cappello	Hat
Lézard	Lucertola	Lizard	Lapin	Coniglio	Rabbit
Papillon	Farfalla	Butterfly	Poison	Veleno	Poison
Radiateur	Stufa	Radiator	Tiroir	Cassetto	Drawer
Raisin	Uva	Grape	Toboggan	Scivolo	Toboggan
Réveil	Sveglia	Alarm clock	Bouchon	Tappo	Cork
Rideau	Tenda	Curtain	Champignon	Fungo	Mushroom
Vélo	Bicicletta	Bicycle	Concombre	Cetriolo	Cucumber