THE SEMANTIC ROLE OF GENDER: GRAMMATICAL AND BIOLOGICAL GENDER MATCH EFFECTS IN ENGLISH AND SPANISH

by

Tamar Degani

B.A., University of Haifa, Israel, 2003

Submitted to the Graduate Faculty of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2007

UNIVERSITY OF PITTSBURGH

ARTS AND SCIENCES

This thesis was presented

by

Tamar Degani

It was defended on

April 5, 2007

and approved by

Erik D. Reichle, Associate Professor, Psychology

Tessa Warren, Assistant Professor, Psychology

Thesis Director: Natasha Tokowicz, Assistant Professor, Psychology

THE SEMANTIC ROLE OF GENDER: GRAMMATICAL AND BIOLOGICAL GENDER MATCH EFFECTS IN ENGLISH AND SPANISH

Tamar Degani, M.S.

University of Pittsburgh, 2007

How does language affect thought? Do the grammatical structures of the language we speak influence the way we think about objects and ideas? The linguistic relativity hypothesis (Whorf, 1956) proposes that the specific language we speak affects the way we think about reality. Predictions made under this hypothesis (e.g., Boroditsky, Schmidt, & Phillips, 2003) posit that grammatical gender is an example of a linguistic structure that affects other aspects of thought. Specifically, because speakers of languages like Spanish denote a grammatical gender to every noun, including those with inanimate referents, this systematic distinction is thought to become part of the meaning representation of objects. Under this hypothesis, pairs of words that match in grammatical gender would be considered as more similar in meaning than pairs that do not share a gender. In four experiments we examined the role of grammatical gender, as well as biological gender, as an organizing dimension of the semantic representation of speakers of Spanish and English. With respect to biological gender, as denoted by English, we found that native English speakers consider pairs of words that share a biological gender (e.g., queen-cow) to be more similar in meaning than pairs that do not share a gender (e.g., king-waitress) (Experiment 1). However, match in biological gender was not sufficient to produce a priming effect in a lexical decision task (Experiment 4). With respect to grammatical gender, as denoted by Spanish, we found that in contrast to the predictions made under the linguistic relativity hypothesis, pairs that match in grammatical gender (e.g., ‘camisa’ (f) – ‘mesa’ (f), shirt-table respectively) did not elicit higher semantic similarity ratings by native Spanish speakers compared to unmatched pairs (Experiment 2), and furthermore these pairs were not processed more quickly or accurately in a primed naming task (Experiments 3A and 3B). We discuss the theoretical and practical considerations that may underlie these effects.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

Languages differ in many grammatical and lexical aspects, which can influence the way speakers of different languages describe the world. However, the degree to which those different descriptions reflect or affect the way people perceive and think about the world is an open question. The linguistic relativity hypothesis, also known as the 'Sapir-Whorf hypothesis' (Whorf, 1956; see Lucy, 1997 for discussion of its origins; for a recent book see Gentner & Goldin-Meadow, 2003), suggests that the particular language we speak affects the way we think about reality. Speakers of different languages are therefore hypothesized to vary in the way they think about the world in accordance with those linguistic aspects that differ across languages.

One aspect in which languages differ is the way gender is implemented. Some languages, like English, employ a 'natural gender' system, in which semantic gender is denoted by pronouns or other grammatical constructs. In those languages only biological gender is marked. Other languages, like Spanish, employ a formal or 'grammatical gender' system, in which all nouns, including those with inanimate referents, are assigned a gender. The current project seeks to explore how the implementation of gender in the language (natural vs. grammatical) influences the semantic representations of speakers of English and Spanish. Specifically, we examined whether biological gender and grammatical gender serve as organizing dimensions of the semantic space, by examining whether gender match increases the perceived meaning similarity of pairs of words and facilitates processing of such pairs in a priming task.

Recent research exploring the linguistic relativity hypothesis has raised the possibility that grammatical gender is an example of a grammatical structure that has implications for other aspects of thinking, specifically semantic representation. To understand why, we need to consider that speakers of languages that employ a grammatical gender system, like Spanish, Hebrew, Russian etc., are not only encouraged, but indeed required to mark objects' gender using definite articles or pronouns, and sometimes to modify adjectives and verbs to agree with the gender of the noun. It is plausible that this systematic grammatical distinction is perceived as reflecting meaning in the world, and is therefore internalized to be part of the semantic representation of objects during the course of development (Konishi, 1993), particularly in languages in which grammatical gender assignment for animate nouns follows biological gender. We can imagine that children learning such a language would have no *a priori* reason to think grammatical gender is different from any other meaningful grammatical distinction like the plural inflection (Boroditsky, Schmidt, & Phillips, 2003; see Bowers, Vigliocco, Stadthagen-Gonzalez, & Vinson, 1999, for a comparison of grammatical gender and the mass/count distinction). We can then ask if, for instance, a native speaker of Spanish, who systematically refers to a table (i.e., 'mesa') with adjectives or articles in the feminine form, thinks about tables as having feminine attributes.

Previous research has suggested that people's thinking about objects may indeed be influenced by the grammatical gender their native language assigns to objects' names. Konishi (1993) found that native Spanish speakers and native German speakers rated concrete nouns that are marked as masculine in their native language as higher on a potency scale, which is considered to be a masculine trait, when this scale was administered along with other semantic differential scales. Interestingly, a subset of the stimuli included translation equivalents in

Spanish and German which are assigned the opposite grammatical gender by these two languages. Items that were masculine in German were rated higher on the potency scale by the native German speakers, but not by the Spanish speakers, and likewise items that were masculine in Spanish (but not in German) were rated higher on potency by the Spanish speakers but not by the German speakers. The significant effect of grammatical gender on ratings of potency found for these specific stimuli suggests that the language's grammatical gender system, and not some inherent property of the specific nouns chosen, underlies this effect.

Similarly, Sera, Berge, and del Castillo-Pintado (1994) compared native Spanish speakers to native English speakers in their classification of pictures of inanimate objects as masculine or feminine (Experiment 1) or in assigning a male or a female voice to those pictures (Experiment 2). Their results suggest that the native Spanish speakers performed these tasks in accordance with the grammatical gender their language assigns to these objects. They further examined the age at which children start to make these classifications based on grammatical gender, and found that starting from the second grade, Spanish speaking children assign voices to objects based on grammatical gender, but these effects were not evident for Spanish children in kindergarten (for similar results with attribute or typical name assignment, see Flaherty, 2001)

Using the same paradigm, Sera, Elieff, Forbes, Burch, Rodríguez, and Dubios (2002) compared monolingual speakers of English, Spanish, French, and German in assigning either a male or a female voice to pictured objects. As in the previous study, French and Spanish speakers assigned voices to pictured objects according to the grammatical gender in their language starting in the second grade. Interestingly, the German speakers' judgments were not influenced in the same way by the grammatical gender in German. Specifically, feminine items were not assigned female voices by the German speakers at above chance level. Sera et al.

created a computational model to try simulate these differences in performance of speakers of different languages, and concluded that a grammatical gender system with two genders and a high correlation between grammatical and natural gender can lead to overgeneralization of masculine and feminine traits to inanimate objects. The grammatical gender system in German does not follow these constraints, because it makes use of 3 genders (feminine, masculine, and neuter) and the correlation between natural and grammatical gender is weakened by some animate nouns that are assigned a neuter gender.

Vigliocco, Vinson, Paganelli, and Dworzynski (2005) discuss two alternative accounts of why and when semantic effects of grammatical gender are to be expected. The first account, termed the *Similarity and Gender Hypothesis* posits that nouns that are similar in grammatical gender (or in any other syntactic or morphophonological property) are perceived as similar in meaning, by virtue of the similar linguistic contexts in which these nouns occur. That is, syntactic agreement provides similar linguistic contexts to nouns that share grammatical gender (i.e., similar determiners, adjectives, and pronouns in sentences), and this similarity in context is interpreted as similarity in meaning. This hypothesis does not require an association between the grammatical gender of nouns and the sex (or biological gender) of the referents.

The alternative hypothesis offered by Vigliocco et al. (2005), the *Sex and Gender Hypothesis*, posits that semantic effects of gender are dependent on an association between the gender of nouns referring to humans and the sex (biological gender) of the referents. Once this co-occurrence of linguistic features (gender: masculine or feminine) and conceptual features (sex: male or female) is established, generalization to other nouns is possible. A constrained version of this hypothesis posits that the principle extend only to words referring to sexuated entities (animals). A less constrained version would propose that the strong association between the gender of nouns and the sex of human referents could lead to the assignment of male- or female-like conceptual properties even to those entities for which sex is not a relevant dimension (i.e., inanimate nouns). Under this hypothesis, semantic effects of grammatical gender will only be expected for speakers of languages that allow this association between gender and sex to develop (i.e., in which grammatical gender is dependent on sex for humans), and will be weaker or absent in words for which sex is not a semantically relevant property.

The lack of grammatical gender effects for speakers of German in the Sera et al. (2002) study can therefore be explained by the Sex and Gender Hypothesis, because in German the association between linguistic and conceptual gender features is weakened by the existence of neuter animate nouns. Further support for this hypothesis is provided by the experiments conducted by Vigliocco et al. (2005), in which participants were presented with triads of words (1140 triads with animal referents; 2024 triads with artifact words). Monolingual speakers of English, which does not employ a grammatical gender system, Italian, which employs a two-gender system, and German, which employs a three-gender system, were asked to judge which two of the three words in each triad were most similar in meaning. The authors found that the Italian speakers, but not the English or the German speakers, chose animate pairs that matched in grammatical gender more often than would be expected by chance, when analyzing the triads that included opposite genders. The authors suggested that the difference between the Italian and German speakers is a result of the less consistent mapping between sex of referents and the gender assigned to the noun in German, especially because in German all nouns in the diminutive form are marked as neuter.

In contrast to these conclusions for speakers of German, Boroditsky and Schmidt (2000) provided some support for the Similarity and Gender Hypothesis, when they found that German speakers' memory performance, as well as that of Spanish speakers, was influenced by grammatical gender (Experiment 2). In their study, participants were asked to learn proper names of objects (in English) chosen to have opposite grammatical genders in Spanish and German. As they predicted, participants were more likely to remember the gender of the proper name when it matched the grammatical gender of the object in their native language. For example, the gender of the proper name 'Patricia', when learned coupled with 'apple', was better remembered by the Spanish speakers compared to the German speakers, because the word for apple is feminine in Spanish but masculine in German.

Boroditsky, Schmidt, and Phillips (2003) reported a series of experiments testing a similar population of German and Spanish speakers in English. In one experiment, they presented participants with a list of objects' names, chosen to have opposite grammatical gender in the two languages, and asked the participants to describe these objects with the first three adjectives that came to mind. Spanish and German speakers were found to generate adjectives to objects in English (their second language) based on the gender assigned to the objects in their native language. For instance, the word "key", which takes the masculine form in German but the feminine form in Spanish, was described as hard and heavy by the German speakers, but as shiny, tiny, and golden by the Spanish speakers. These adjectives were rated as stereotypically masculine and feminine, respectively, by a group of English speakers.

In a second experiment (also reported in Phillips & Boroditsky, 2003), participants rated the similarity of pairs of unlabeled pictures depicting objects and people on a scale of 1 (not similar) to 9 (very similar). Spanish as well as German Speakers rated objects as more similar to a person when the grammatical gender of the object matched the biological gender of the person than when the gender did not match. Using an artificial language, the authors also presented native English speakers with 8 pictures of males and females, along with 12 pictures of inanimate objects in two groups, and taught them which would be considered “soupative” and which “oosative”. This distinction always corresponded to the biological gender of the male and female pictures, but also generalized to inanimate objects. After this distinction was perfectly learned, participants rated the similarity of object-person pairs. Similar to the result for the Spanish and German speakers, object-person pairs that shared a gender (i.e., both belonged to the same category in the artificial language) were rated as more similar than pairs that did not share a gender (i.e., one soupative and one oosative). The authors concluded that peoples’ thinking about objects can be influenced by grammatical aspects that differ across languages.

Eberhard, Scheutz, and Heilman (2005) replicated these effects, and further showed that they generalize to similarity ratings of inanimate objects paired with ‘new humans’ (i.e., humans referred to by nouns not presented during learning). They also simulated these results with a connectionist model and concluded that these findings could be due to indirect associative links between lexical representations and grammatical features and between grammatical features and conceptual properties of sex, and that direct associative connections between the inanimate objects and the conceptual properties of sex are not required.

Although the results reviewed thus far seem to suggest that grammatical gender exerts a semantic influence, a few caveats should be considered. First, because all of the paradigms described above were off-line, it is possible that these findings depend on a strategy employed by participants. Second, these results may not be extendable to all nouns; it is possible that grammatical gender is internalized into the semantic representations for some nouns, but not for all of them. For example, grammatical gender may be internalized for words that are very animate-like or interactive in nature (e.g., car), but not for others. Furthermore, the idea that grammatical gender becomes part of the semantic representation faces several theoretical challenges, such as the arbitrariness of gender assignment across languages (Foundalis, 2002), and the existence of concepts with two labels with opposite grammatical gender within the same language. For instance, in Spanish, computer is referred to both as computadora (f) and as ordenador (m); it is not clear how grammatical gender would be internalized for such concepts. Finally, as evident by the Sera et al. (2002) and Vigliocco et al. (2005) studies, it is not clear how grammatical gender exerts its influence for speakers of languages that employ three genders, like German, in which grammatical gender assignments does not always follow the biological gender.

Indeed, other studies have failed to replicate these semantic effects of grammatical gender. For example, in an elegant study, Vigliocco, Vinson, Indefrey, Levelt, and Hellwig (2004) induced semantic substitution errors using a continuous picture naming paradigm with speakers of German. Semantically-related lexical retrieval errors are thought to reflect co-activation of semantically-related lexical candidates that on some proportion of trials give rise to semantic substitution. Crucially, these semantic substitutions tend to preserve the grammatical class and grammatical gender of the target noun (Marx, 1999). Vigliocco et al. explored whether this gender preservation is because words that share syntactic properties tend to be semantically more similar, as would be predicted by the linguistic relativity hypothesis, or whether gender is preserved as a result of the interplay between retrieving lexico-semantic representations and building sentential frames (at either the syntactic or the morphophonological level). Their results show that although gender preservation effects were present when participants produced phrases with determiners marked for gender (Experiment 2), the effect was not found when participants produced bare nouns or phrases with determiners that were not marked for gender (Experiments 1 and 3, respectively). They therefore concluded that gender preservation does not reflect greater semantic similarity between words that share the same gender, but rather is a result of the use of the gender property at the morphophonological level when building sentential frames during production. These results point to syntactic rather than semantic effects of grammatical gender. The authors nonetheless point out that their results may not generalize to abstract nouns, or to conceptually grounded gender (i.e., biological gender).

One other element of Vigliocco et al.'s experiments is worth noting. Grammatical gender effects, under their account, are limited to effects that are language-specific and reflect increased similarity among words that happen to share the same grammatical gender. These are to be distinguished from connections between syntactic properties to semantics which are common across languages. In accordance with their conceptualization they looked at the gender contribution to semantic substitution only after partialling out form similarity and semantic similarity measures taken from English speakers. These semantic similarity measures reflect language independent connections between syntactic properties and semantics common across languages, and were operationalized as item-by-item semantic distance measures derived from speaker-generated feature norms for the English translations. Therefore, although grammatical gender was preserved in 46.3% of the substitutions when participants produced bare nouns (more than chance because German has three genders), the gender factor was not significant once English semantic similarity (reflecting syntactic-semantic connections common across languages) was taken into account. Also, the specific semantic field of the stimuli (i.e., animals and body parts) may have contributed to the magnitude of the effect of the syntactic-semantic connections common across languages, because it is possible that for these specific semantic domains higher language-independent syntactic-semantic connections exist. Thus this null effect may be limited to these semantic domains, and might not be extendable to other semantic areas.

Furthermore, these studies were conducted in a three-gendered language (German) and might not yield the same results in a two-gendered language like Spanish.

The results concerning the semantic effects of grammatical gender are therefore mixed. The extent to which such effects exist and their generalizability to other stimuli, languages and paradigms are in need of further investigation. The current project focuses on Spanish, which employs a two-gender system, in which grammatical gender follows biological gender, and examines if grammatical gender influences semantic processing in an ‘off line’ as well as in an ‘on line’ paradigm that is less open to strategic control. However, before we examine if such semantic effects exist, we first sought to establish that *biological* gender, as manifested in English, indeed influences performance in a semantic task. If biological gender, which clearly represents meaning in the world, does not guide people’s performance in semantic tasks, we would not expect grammatical gender to do so.

We hypothesized that if biological and grammatical gender are semantic dimensions that contribute to the organization of the semantic space, then word pairs that share a gender would be judged as being more semantically related compared to word pairs that do not share a gender. In Experiment 1 we examined this question with respect to biological gender match, and tested native English speakers in a semantic similarity rating task. In Experiment 2 we examined the question with respect to grammatical gender match, and tested native Spanish speakers in the same semantic similarity rating task.

To foreshadow, the results of these two experiments suggested that biological, but not grammatical gender, serve as a dimension that guides people’s perceptions of semantic similarity. However, these experiments do not speak to the possible effects of such a dimension in an on-line, less strategic task. To examine the extent to which gender match guides on-line performance, we used a semantic priming paradigm in Experiments 3 and 4.

We should emphasize that the linguistic relativity hypothesis does not predict that all aspects of language should affect thought to the same extent. Therefore by examining the semantic effects of grammatical gender we can not speak to the validity of the linguistic relativity hypothesis in general. We focus on the specific prediction made in the spirit of the linguistic relativity hypothesis (e.g., Boroditsky, Schmidt, & Phillips, 2003) with respect to grammatical gender, according to which semantic effects of grammatical gender are an example of the effects of language on thought. For simplicity we will refer to this specific prediction as the linguistic relativity hypothesis, yet keeping in mind that our results do not speak to the validity of the hypothesis for other grammatical constructs.

Across all these experiments, the linguistic relativity hypothesis would predict effects of gender match, for both biological and grammatical gender. A more modular perspective, positing functional independence between syntactic and semantic representations, would predict effects only for biological gender, but not for grammatical gender. The results are discussed with respect to these contrasting approaches, and to the possible mechanisms that would give rise to these effects.

2.0 EXPERIMENT 1: SEMANTIC SIMILARITY RATING OF BIOLOGICAL GENDER MATCH IN ENGLISH

In this experiment we explored whether biological gender serves as an organizing dimension of semantic representations. Natural gender is marked by pronouns in English. Aside from this pronominal system, only a few English nouns are inherently marked for gender, some by morphological markings (e.g., actress, waitress) (Scheutz & Eberhard, 2004), but not others (e.g., nun, priest). Scheutz and Eberhard posit that because sex is a salient property of humans in general, it is likely to be a feature that is included in the conceptual representation of any noun denoting humans. For nouns that obligatorily denote a female or a male category (e.g., queen, prince) a semantic feature would be specified as female or male, respectively. For other nouns denoting humans, the relative frequency of male and female exemplars in the conceptual category in the world (e.g., the proportion of secretaries who are female) would determine the strength of the specification of a female vs. male value. Therefore, biological gender can be thought of as a semantic feature that is shared between referents. For instance ‘bull’ and ‘nephew’ may both share the semantic feature ‘masculine’, whereas ‘cow’ and ‘queen’ would both be ‘feminine’. Alternatively, gender may correspond to a *set* of features that are shared between members of the same biological gender category. These features may reflect core characteristics or may be stereotypical or prototypical of members of the category (e.g., all the masculine nouns are strong). Although differentiating these two alternatives would be of interest, it is beyond the scope of the current project. Most relevant for the purpose of the current study, if two nouns share biological gender, we can expect them to share either a single feature or a set of features, which may make them more similar in meaning than two nouns that do not share biological gender. Following Andonova, D'Amico, Devescovi, and Bates (2004), we define biological gender as a characteristic of nouns referring to living beings (humans and animals) that are perceived as different on the basis of biological sex. It is not used to refer to words representing animate creatures whose sexuality is not entirely obvious to human language users (e.g., insect or fly).

Based on these considerations we hypothesized that pairs of words that share biological gender (e.g., cow - waitress) would be perceived as more similar in meaning than pairs of words that differ in their biological gender (e.g., nephew - queen). To test our hypothesis we employed a semantic similarity rating task, in which participants were asked to rate how similar in meaning a pair of words was on a scale of 1 to 7, on which 1 corresponded to "completely different" and 7 corresponded to "exactly the same". To examine these effects in the context of other semantic relations, we also included a subset of stimuli that were either related or unrelated in meaning, but had no biological gender (e.g., honey – sweet).

2.1 METHOD

2.1.1 Participants

No participant took part in more than one of the following experiments. Forty native English speakers (23 males; mean age 18.6 years), who were not raised learning another language in addition to English, participated in this experiment toward Introductory Psychology class credit. Five additional participants were tested and replaced because their language history questionnaire indicated that they had been raised learning another language in addition to English. The language history questionnaire data for the final set of participants in this experiment, and the experiments that follow, are shown in Table 1.

Table 1. Language History Questionnaire Data by Experiment

	Experiment				
Measure	1	2	3A	3B	4
Number of Participants	40	27	24	15	96
Number of male participants	23	13	7	7	38
First Language (L1)	English	Spanish	Spanish	Spanish	English
Age (years)	18.6 (.93)	35.7 (12.4)	33.2 (10.9)	31.7 (7.9)	19.4 (3.2)
Age Began L2 (years)	12.5 (3.6)	10.5(4.5)	11.5 (8.5)	11.6 (6.2)	13.8 (2.6)
Time Studied L2 (years)	N/A	N/A	15.0 (11.7)	10.9 (8.3)	N/A
L2 Immersion Experience (years)	0	10.4 (12.1)	4.9 (5.4)	5.0 (6.0)	N/A
L1 Reading Ability	9.6	9.7	9.1	9.7	N/A
L2 Reading Ability	4.8	8.8	8.2	8.5	N/A
L1 Writing Ability	9.6	9.1	9.1	9.6	N/A
L2 Writing Ability	3.9	7.9	7.7	7.5	N/A
L1 Conversation Ability	9.8	9.3	9.4	9.8	N/A
L2 Conversation Ability	3.4	7.9	7.6	7.8	N/A
L1 Speech Comprehension Ability	9.8	9.5	9.4	9.8	N/A
L2 Speech Comprehension Ability	3.9	8.5	8.2	8.5	N/A

2.1.2 Design

A 2 relatedness (related vs. unrelated) by 2 relation type (biological gender vs. pure semantic) within-participants design was used.

2.1.3 Stimuli

The words in this experiment were presented to participants in pairs. We therefore refer to the first word in each pair as a prime, and to the second word as a target. In the biological gender condition 72 words with animate referents and a clear biological sex (e.g., mother, bull, salesman) were used; half of these were feminine. Feminine items were matched to masculine items in length, frequency, and mean RT to make a lexical decision (taken from Elexicon, Balota et al., 2002). The characteristics of these words are shown in Table 2. These words were used to create 16 matched pairs (8 feminine-feminine and 8 masculine-masculine) and 16 unmatched pairs (8 feminine-masculine and 8 masculine-feminine). Eight different versions of these pairings were created and were counterbalanced across participants[1].

---

1 Across the different versions, five gender words appeared only as primes (father, female, princess, son, women) and three words appeared only as targets (nun, monk, mare). These 8 words served in the filler nonword-word pairs in Experiment 4. Importantly, each version of the experiment consisted of 16 matched gender pairs and 16 unmatched pairs.

Table 2. Properties of the 'biological gender' condition stimuli used in Experiments 1 and 4

	Stimulus Type	
	Feminine	Masculine
English Frequency (HAL)	11378.17 (16573.9)	21031.17 (38371.59)
English Log Frequency (HAL)	7.81 (.35)	8.60 (.35)
Length (number of letters)	6.47 (2.65)	5.81 (2.15)
Mean RT to make a lexical decision	681.65 (110.4)	655.63 (77.19)

Note: None of the differences between feminine and masculine stimuli are significant ($p > .50$). Standard deviations are shown in parentheses. Hal frequencies are based on the HAL corpus (Lund & Burgess, 1996), which consists of approximately 131 million words. Log frequency refers to log-transformed HAL frequency norms. All characteristics were taken from Elexicon (Balota et al., 2002).

The pure semantic condition consisted of a total of 38 pairs of words taken from previous studies in which they had been shown to create significant priming effects. We refer to this condition as 'pure semantic' to emphasize that the relation between items in each pair is along semantic dimensions and are not associative in nature (for a discussion of associative vs. "pure" semantic priming see Lucas, 2000; McNamara, 2005; Neely, 1991). Sixteen target words combined with semantically related prime words or with 16 unrelated primes were taken from McRae and Boisvert (1998). All participants saw the same 16 target words, half primed by a semantically related prime and half primed by an unrelated prime. Similarly, 22 target words, combined with 22 semantically related primes and 22 unrelated primes were taken from Plaut and Booth (2000). As with the McRae and Boisvert stimuli, participants all saw the same 22 targets, half primed by a semantically related prime and half primed by an unrelated word.

Across the eight versions of the experiment, each target word was primed by a related word on half of the trials, and by an unrelated prime word on the other half of the trials.

Related primes were matched to the unrelated primes on length, frequency, and time to make a lexical decision (taken from Elexicon, Balota et al., 2002) (see Table 3). Furthermore, in each of the 8 versions, related pairs were matched to unrelated pairs in terms of length and frequency of the primes and targets.

Table 3. Properties of the 'pure semantic' condition stimuli used in Experiments 1 and 4

	Stimulus Type		
	Targets	Related Primes	Unrelated Primes
Number of items	38	38	38
English Frequency (HAL)	34045.58 (67933.99)	20342.92 (37650.29)	38554.03 (74607.47)
English Log Frequency (HAL)	9.05 (1.82)	8.68 (1.74)	9.28 (1.66)
Length (number of letters)	5.29 (1.10)	5.29 (1.10)	5.08 (1.02)
Mean RT to make a lexical decision	622.13 (73.30)	626.83 (58.81)	633.12 (84.36)

Note: None of the differences between the types of stimuli are significant (all p_s > .20). Standard deviations are shown in parentheses. The same targets were used with both the related and unrelated primes. Hal frequencies are based on the HAL corpus (Lund & Burgess, 1996), which consists of approximately 131 million words. Log frequency refers to log-transformed HAL frequency norms. All characteristics were taken from Elexicon (Balota et al., 2002).

In summary, the stimuli consisted of 196 English words, presented to participants as 80 pairs. Of these, 32 pairs formed the gender condition, 38 pairs formed the pure semantic condition, and 10 pairs served as fillers. Examples of stimuli in each condition are shown in Table 4.

Table 4. Example stimuli from Experiments 1 and 4

	Relation Type		
	Biological Gender		Pure Semantic
Related Pair	bull - nephew	cow - queen	honey - sweet
Unrelated Pair	waitress - king	father - lady	screw - clock

2.1.4 Procedure

Participants were tested on a computer. Instructions were given in English, on the computer screen, and were emphasized by the experimenter. Participants were told they would be presented with a list of English word pairs, and would be asked to rate how similar in meaning each pair is on a scale of 1 to 7, with 1 indicating "completely different", and 7 indicating "exactly the same"; similar semantic similarity rating tasks have been used by Tokowicz, Kroll, De Groot, and Van Hell (2002, but for translation), Schweppe and Rummer (2007) using a 1 to 10 scale, and Boroditsky et al.(2003), and Eberhard et al. (2005) for pairs of pictures on a 1-9 scale. Participants were to complete this rating task at their own pace, but were encouraged to complete the entire experiment in one sitting. Each participant was presented with 80 pairs to be rated, preceded by two examples. Upon completion of the rating task, participants were presented with a language history questionnaire on the computer, in which they rated their proficiency in first and second language reading, writing, conversation, and speech comprehension on a scale of one to ten, and indicated the age at which L2 learning began and the types of L2 exposure (modified from Tokowicz, Michael, & Kroll, 2004).

2.2 RESULTS

Data by participants were analyzed using a repeated measures Analysis of Variance (ANOVA) with relatedness and relation type as within participants variables, and are reported as F_1. Data by items were analyzed using a repeated measures ANOVA with relatedness as a within item variable and relation type as between items variable, and are reported as F_2[2]. The mean semantic similarity ratings for items from each condition are presented in Table 5.

Table 5. Mean similarity rating (Experiments 1 & 2) and mean reaction time and percent correct (in parentheses) (Experiments 3A, 3B, & 4) for related and unrelated pairs

	Experiment									
	1		2		3A		3B		4	
Condition	Gender	Pure	Gram	B&G	Gram	B&G	Synt	B&G	Gender	Pure
Related	2.74	4.31	2.19	2.04	529.09 (97.0)	533.77 (97.9)	513.88 (98.0)	519.14 (96.7)	722.27 (96.7)	697.79 (97.7)
Unrelated	1.66	1.48	2.24	2.16	527.53 (98.3)	533.33 (97.1)	520.03 (97.0)	524.44 (97.7)	726.82 (96.6)	709.43 (97.2)
Relatedness advantage	1.08**	2.83**	-.05	-.12	-1.56 (-1.3)	-.44 (0.8)	6.15 (1.0)	5.3 (-1.0)	4.55 (0.1)	11.64 (0.5)

Note: ** $p < .001$. Values for Experiments 1 and 2 are mean similarity rating on a scale of 1-7, '1' representing "completely different", and '7' representing "exactly the same". Values for Experiments 3A, 3B and 4 are mean reaction time and accuracy in parentheses. Relatedness advantage is computed as Related-Unrelated for the similarity ratings (Experiments 1&2) and for accuracy data (Experiments 3A, 3B, &4), but as Unrelated-Related for the RT data (Experiments 3A, 3B, & 4).

2 F_2 was computed based on the target items, which were primed by related or unrelated items. A total of 105 targets are included in these analyses, 38 in the pure semantic condition and 67 in the gender condition (as explained in Footnote 1). The 10 targets of the filler pairs were not included in the analysis.

The main hypothesis, that related pairs (M = 3.53) would be rated as more similar in meaning compared to unrelated pairs (M = 1.57) was supported by the rating data, F_1 (1, 39) = 507.01, MSE = .30, $p < .001$, F_2 (1, 103) = 272.76, MSE = .71, $p < .001$. Furthermore, pure semantic condition pairs were rated as more similar in meaning (M = 2.89) than gender condition pairs (M = 2.20), F_1 (1, 39) = 90.34, MSE = .21, $p < .001$, F_2 (1, 103) = 33.00, MSE = .71, $p < .001$. However, this effect should be interpreted with caution, because stimuli across the two relation types were not matched for important linguistic characteristics[3]. Furthermore, these effects were qualified by a significant interaction between relatedness and relation type, F_1 (1, 39) = 193.25, MSE = .16, $p < .001$, F_2 (1, 103) = 47.89, MSE = .71, $p < .001$, such that relatedness had a stronger influence in the pure semantic condition compared to the gender condition (see Figure 1).

---

3 We conducted a repeated measures ANCOVA by items, with length and log frequency as covariates. The effect of relation type was still significant even after length and log frequency of the stimuli were covaried, F_2 (1, 101) = 28.38, MSE = .72, $p < .001$. Therefore it does appear that pure semantic condition pairs were perceived as more similar in meaning overall than gender condition pairs.

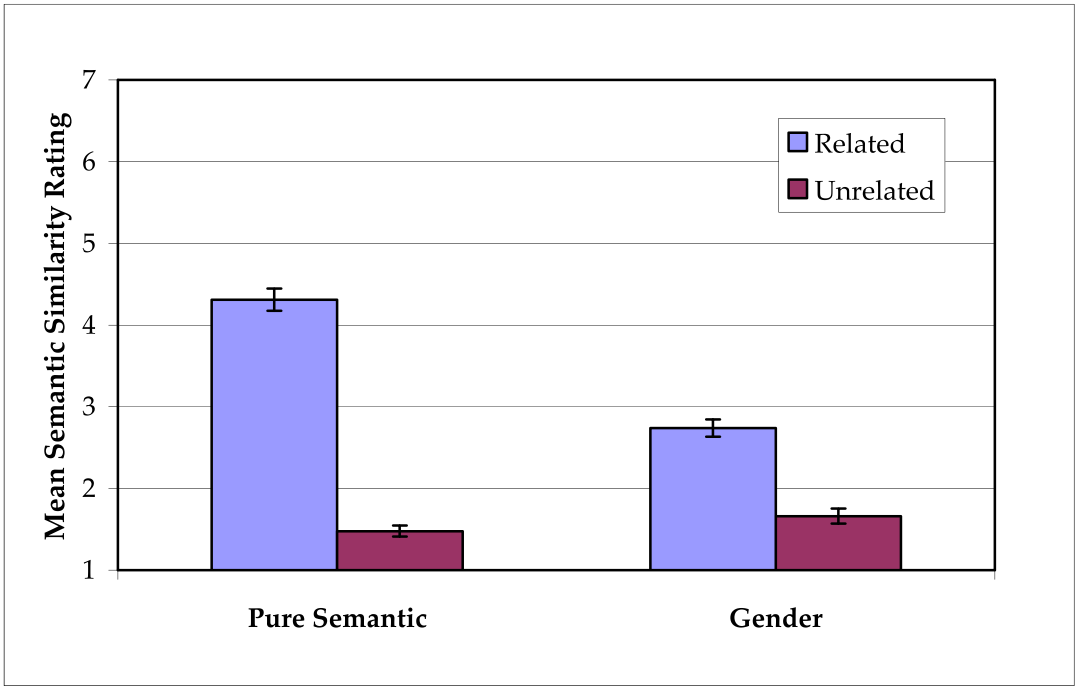

Figure 1. Mean semantic similarity rating in Experiment 1

Our main question of interest was whether pairs matched in biological gender are treated by participants as more similar in meaning than pairs that were not matched in gender. To address this question directly we conducted a one-way ANOVA including the gender condition stimuli only, with relatedness as a within participants/item variable. As we predicted, the effect of gender match was significant in both the analysis by participants, F_1 (1, 39) = 168.78, MSE = .14, $p < .001$, and in the analysis by items, F_2 (1, 66) = 68.54, MSE = .66, $p < .001$, indicating that pairs that share a biological gender are perceived as more similar in meaning (M = 2.74) than words that do not share a gender (M = 1.66).

2.3 DISCUSSION

These results therefore indicate that pairs that share a biological gender are treated as more similar in meaning than pairs that do not share a gender. Thus, biological gender may be one of the dimensions that guide conceptual organization. The co-existence of the semantic effects of biological gender match and the pure semantic relatedness in the same task strengthen our confidence that participants were indeed performing the task as instructed, and rated the similarity of the pairs in terms of their meaning. Pure semantic condition pairs were rated as more similar in meaning overall compared to the gender condition pairs, even after length and frequency were taken into account. It is possible that these pairs share more semantic features than do the gender condition pairs, even when one assumes gender match takes the form of multiple shared features. This may suggest that sharing a biological gender does not overshadow other semantic connections in the lexicon. Semantic relatedness of the pairs that do or do not share a gender was not manipulated, making it difficult to estimate the additive effect of gender match to other types of semantic relatedness.

To examine if grammatical gender serves a similar role for speakers of languages that employ a grammatical gender system, such as Spanish, we conducted Experiment 2, in which native Spanish speakers performed the same semantic similarity rating task in Spanish, but importantly rated the similarity in meaning of pairs of Spanish words that either matched in grammatical (rather than biological) gender (e.g., madre-mesa) or did not (e.g., chico-camisa).

3.0 EXPERIMENT 2: SEMANTIC SIMILARITY RATING OF GRAMMATICAL GENDER IN SPANISH

The main purpose of this experiment was to examine if grammatical gender serves as an organizing dimension of the conceptual representation of speakers of languages that employ a grammatical gender system, similar to the role served by biological gender for speakers of English. We tested native Spanish speakers with the same meaning similarity rating task as that used in Experiment 1. However, because in Spanish inanimate nouns are assigned a grammatical gender, it was not possible to include a control set of stimuli that are related or not in meaning and do not convey a gender. Furthermore, biological and grammatical gender are coupled in Spanish, such that words that have a biological gender (e.g., madre, which means mother - female) have a matching grammatical gender as well (Sera et al., 1994). We therefore compared pairs that do or do not share a grammatical gender only (i.e., inanimate noun pairs) to pairs in which one of the items conveys a biological as well as grammatical gender (i.e., animate-inanimate pairs). This allowed us to examine if the presence of a biological gender contributes to the effect of grammatical gender match.

Examining the effect of grammatical gender separately for animate and for inanimate nouns was also motivated by the studies reported by Vigliocco et al. (2005). As described earlier, Vigliocco et al. examined the effect of grammatical gender match in Italian (a two-gender language like Spanish) compared to English and German. Interestingly, they found that the

Italian speakers tended to select word pairs sharing a gender (more often than the English or German speakers), when asked to judge which two of three words are most similar in meaning, but the effect was present only for words referring to animals but not for words referring to artifacts. A similar pattern of results was found in their second experiment, in which semantic substitution errors were induced in a continuous naming paradigm (similar to Vigliocco et al., 2004). We therefore decided to explore the effect of grammatical gender in Spanish for animate and inanimate nouns separately. However, because our main focus was grammatical, rather than biological gender, all targets were inanimate nouns and animacy was manipulated for primes only.

We expected the effect of grammatical gender match, if present, to be stronger in the animate-inanimate condition in which biological gender in addition to grammatical gender primes the target. We therefore predicted targets in gender matched pairs to be processed more quickly and/or accurately than targets in pairs that did not match in gender. We further predicted that gender match would have a stronger influence in the animate-inanimate condition.

3.1 METHOD

3.1.1 Participants

Twenty-seven native Spanish speakers (13 men; mean age 35.7 years) participated in this experiment. Participants were volunteers, recruited through email and advertisements. Because the rating task was completed on a computer through a web-based interface, participants were able to complete the experiment from any location in the world, on their own time. These considerations contributed to the variability in the background of our participants. We will return to this issue in the general discussion. Importantly, however, as indicated in Table 1, the final set of participants were all native Spanish speakers who did not grow up learning another language in addition to Spanish. Three additional participants were replaced because their language history questionnaire indicated that they had learned another language together with Spanish or learned Spanish as a second language. Eight additional participants were excluded from the analysis because they failed to use the full range of the rating scale. It is interesting to note that these eight participants rated all (or almost all – more than 90%) of the pairs as completely different (i.e., 1). All the participants received the same written instructions in Spanish, however it is possible that the meaning of the instructions was interpreted differently by speakers of different countries of origin. As noted by Boroditsky and Schmidt (2000) with respect to instructions given in different languages, "one cannot be sure that the words used for "same" mean the same thing in both languages. If in one language the word for "same" is closer in meaning to "identical", while in the other language its closer to "relationally similar" speakers of different languages may behave differently…" (p.3). The participants in this study were all native Spanish speakers, but it is possible that the way they interpreted the instructions varied as a function of the colloquial use of Spanish in their country of origin. It is beyond the scope of this project to systematically examine the reason why these participants interpreted the instructions differently. However it is interesting to note that their was a substantial overlap in the country of origin between the excluded and included participants, suggesting that country of origin can not fully account for the drastic difference in the interpretation of the task[4].

4 Excluded participants were from Colombia (3); Spain (2), Chile (1); Puerto Rico (2), whereas the included participants were from Colombia (3); Spain (8); Chile (3); Puerto Rico (2); Mexico (7); South America (1); Argentina (1); Dominican Republic (1).

3.1.2 Design

A 2 gender match (matched vs. unmatched) by 2 gender type (biological and grammatical gender vs. grammatical gender) within-participants design was used.

3.1.3 Stimuli

The stimuli were 200 Spanish nouns presented to participants in 80 pairs. The words were all in a singular form, and were noncognates with English (i.e., not similar in orthography and phonology to their English translation equivalents). Half the words were grammatically feminine in Spanish and half were grammatically masculine. Although all the target words (those appearing second in the pair) had inanimate referents (i.e., referents without inherent biological gender), half of the primes presented to participants had inanimate referents and half had animate referents (describing men and women). As mentioned earlier, grammatical gender follows biological gender in Spanish, and therefore the nouns referring to humans and animals are characterized by a biological gender in addition to the grammatical gender. Because animate nouns have both grammatical and biological gender, we call animate-inanimate pairs the *biological and grammatical gender* condition, and the inanimate-inanimate pairs the *grammatical gender* condition. Each participant was presented with 40 animate-inanimate prime-target pairs (biological and grammatical condition), and 40 inanimate-inanimate pairs (grammatical gender condition). In both conditions, half of the pairs matched in gender and half did not. All the participants saw the same 80 targets in a randomized order. For each target word, roughly half[5] of the participants saw it preceded by a prime that matched in gender, and the other participants saw it preceded by a prime that did not match in gender. In the biological and grammatical gender condition the 40 targets were preceded by the feminine or masculine form of an animate noun (e.g., chica or chico which mean girl and boy, respectively) and in the grammatical gender condition the 40 targets were preceded by either a feminine inanimate noun or a masculine inanimate noun (e.g., toalla or barco which mean towel and boat, respectively) matched for Spanish frequency (taken from Carmona, et al., 1998) and length. See Table 6 for examples of stimuli in the different conditions. Eight different pairings of primes and targets were randomly created, and the order of presentation of the prime-target pairs within each list was randomized. Effort was made to ensure that none of the pairs conveyed an obvious associative relation between the prime and the target. Table 7 presents the properties of the stimuli of Experiment 2.

[5] Unfortunately, we were unable to match the number of participants in each version of the experiment exactly.

Table 6. Example stimuli from Experiments 2, 3A, and 3B

	Gender Type					
	Grammatical Gender Primes		Biological & Grammatical Gender Primes		Syntactic Gender Primes	
Experiment (s)	2 & 3A		2, 3A, & 3B		3B	
Target	Feminine inanimate	Masculine inanimate	Feminine animate	Masculine animate	Feminine adjective	Masculine adjective
Feminine inanimate	ensalada - camisa	barco - camisa	esposa - gira	esposo - gira	determinada - falda	determinado - falda
Masculine inanimate	toalla - hombro	recurso - hombro	chica - cuchillo	chico - cuchillo	lista - asiento	listo - asiento

Note: Ten pairs of each type were presented for a total of 80 pairs in each experiment.

Table 7. Properties of the stimuli used in Experiments 2, 3A, and 3B

Condition	B&G Animate Primes		B&G Inanimate Targets		Grammatical Inanimate Primes		Grammatical Inanimate Targets		Syntactic Adjective primes	
Experiment (s)	2, 3A, & 3B		2, 3A, & 3B		2 & 3A		2, 3A, & 3B		3B	
Gender (n)	**M** (40)	**F** (40)	**M** (20)	**F** (20)	**M** (20)	**F** (20)	**M** (20)	**F** (20)	**M** (40)	**F** (40)
Mean Spanish Frequency	182 (345)	133 (349)	169 (145)	164 (136)	177 (82.5)	136 (209)	170 (51.6)	167 (69.3)	180 (101)	135 (106)
Mean Spanish Length (number of letters)	7.4 (1.9)	7.7 (2.1)	5.8 (2.4)	5.8 (1.8)	7.4 (1.7)	7.1 (1.6)	5.7 (1.0)	5.8 (1.6)	6.9 (1.6)	6.9 (1.7)

Note: Standard deviations are shown in parentheses. Spanish frequencies were taken from Carmona et al. (1998). There are no significant differences in frequency between conditions, gender, and position (i.e., prime vs. target) (all *Fs* < 1). There is a significant position effect on length, F (1,270) = 39.96, MSE = 3.20, p < .001, such that primes (M = 7.23) are longer on average than targets (M = 5.74). Importantly, the interaction of position with condition or gender is not significant (p > .60).

3.1.4 Procedure

Participants completed the experiment on a computer through a web-based interface at their convenience. Instructions were presented in Spanish on the computer screen, and encouraged completion of the experiment in one sitting. Similar to Experiment 1, participants were presented with a list of word pairs (in Spanish), and were asked to rate how similar in meaning each pair was on a scale of 1 to 7, with 1 indicating "completely different", and 7 indicating "exactly the same". They were to complete this rating task at their own pace. Each participant was presented with 80 pairs to be rated, preceded by two examples. Upon completion of the rating task, they were presented with a language history questionnaire on the computer, similar to the one used in Experiment 1 that had been translated into Spanish.

3.2 RESULTS

Data by participants were analyzed using a repeated measures ANOVA with gender match and gender type as within participants variables, and are reported as F_1. Data by items were analyzed using a repeated measures ANOVA with gender match as a within item variable and gender type as between items variable, and are reported as F_2[6]. The mean semantic similarity ratings for items from each condition are presented in Table 5.

The main effect of gender type was marginally significant by participants, F_1 (1, 26) = 3.89, MSE = .09, p = .059, suggesting that grammatical gender condition pairs were rated

[6] F_2 was computed based on the target items, which were primed by related or unrelated items.

marginally more similar (M = 2.21) than biological and grammatical condition pairs (M = 2.10). However this effect was not significant by items, F_2 (1, 78) = 1.77, MSE = .36, $p > .10$. No other effect was significant in the analysis by participants or by items ($p > .10$).

Therefore, our main hypothesis, that pairs that match in gender will be rated as more similar in meaning than pairs that do not share a gender was not supported by our data ($p > .10$). Furthermore, our secondary prediction that the interaction between match and gender type, indicating that the gender match effect will be stronger in the biological and grammatical condition compared to the grammatical gender condition was not supported either (*F*s < 1).

We did not obtain a significant effect of grammatical gender match using the same paradigm (and ten more items) as that in Experiment 1, which gave rise to significant effects of biological gender match. Although a lack of power should be considered (the observed power for the match effect was .37), it is interesting to note that the pattern of means actually suggests that pairs that match in gender are rated as *less* similar in meaning (M = 2.11) than pairs that do not match in gender (M = 2.20).

3.3 DISCUSSION

Grammatical gender match effects were not found using the semantic similarity rating task, suggesting that native Spanish speakers do not perceive pairs that match in grammatical gender as more similar in meaning than pairs that do not match in this syntactic property. However, it is still possible that by means of shared features, activation in a more automatic task will give rise to a grammatical gender match effect. In Experiment 3 we therefore conducted a primed naming task, in which native Spanish speakers produced targets that were preceded by a prime that did or did not match in gender. The linguistic relativity hypothesis would predict that pairs that match in grammatical gender would be processed as semantically related pairs, which would yield a facilitation effect in this semantic priming paradigm. However, if gender is processed as a syntactic feature, activated only when syntactic gender agreement requires it, no priming effect should be expected.

4.0 EXPERIMENT 3A: SEMANTIC PRIMING OF GRAMMATICAL GENDER IN SPANISH

To examine the extent to which grammatical gender match guides performance during on-line word processing, we used a semantic priming paradigm. In this experiment, native Spanish speakers performed a semantic priming task, in which they were asked to name out loud the second of two words presented one at a time (i.e., the target). A naming task was selected for this study for two main reasons. First, the semantic priming literature suggests that the naming task is less open to strategic control when short presentation parameters are used (Lucas, 2000; Neely, 1991). Second, by requiring participants to produce the target, we could explore whether gender is selected or activated when a bare noun is to be produced. Schriefers (1993) proposed that an abstract gender feature of the noun needs to be selected for production, and this gives rise to competition when the target noun is incongruent in gender with a distracter, resulting in a gender congruency effect (but see Costa, Kovacic, Fedorenko, & Caramazza, 2003, for a different account of the gender congruency effect). This account would predict that producing a bare noun, as in our paradigm, involves selecting, or activating the abstract gender node. In accordance with this formulation, Plemmenou, Bard, and Branigan (2002) examined gender priming with native Greek speakers. They reasoned that if there is a single gender node shared by all words of that gender, subsequent production of a word sharing the same gender as its prime should be facilitated. They focused their investigation on noun-adjective pairs, and found that masculine nouns indeed produced significant priming effects when producing the subsequent adjective (color of a picture). They concluded that gender priming is a result of facilitated reaccess to a gender node, rather than spreading activation to all gender matching elements in the lexicon. However, in their study participants had to reaccess the gender node to produce the adjective, because gender is not an inherent property of the adjective (but rather structurally conditioned). In the current experiment the targets are nouns, for which gender is an inherent property (i.e., a given noun's gender is determined irrespective of its syntactic environment), and therefore noun production does not require reaccessing the gender node. If gender priming is obtained under these conditions, it is likely a result of spreading activation or automatic activation of the gender node, rather than a result of controlled selection of a gender node for syntactic production purposes.

We hypothesized that if grammatical gender guides the internal organization of semantic representations, and if pairs that match in gender are more similar in meaning, such pairs should prime each other and facilitate naming performance more than pairs that do not match in gender. The stimuli used were identical to those used in Experiment 2. Importantly, half of the pairs matched in grammatical gender and half did not.

4.1 METHOD

4.1.1 Participants

Twenty-four native Spanish speakers (7 men; mean age 33.2 years) participated in this experiment. They were recruited from the Pittsburgh community, and were paid $5 for their participation. Although the participants had been living in the US for five years on average, Spanish was their native and dominant language as indicated by their self ratings on a language history questionnaire. The data from an additional two participants were lost due to technical failures, and the data from another participant was removed from the analysis because his overall reaction time and his self-rated proficiency scores implied that English, rather than Spanish, was his dominant language. The statistical analyses were conducted on data from a final set of 24 participants. The language history questionnaire data for this final set of participants are shown in Table 1.

4.1.2 Design

A 2 gender match (match vs. unmatched) by 2 gender type (biological and grammatical gender vs. grammatical gender) within-participants design was used.

4.1.3 Stimuli

The stimuli were identical to those used in Experiment 2. The order of presentation of the prime-target pairs within each version was randomized for each participant by the computer program (Eprime software, Psychology Software Tools, Pittsburgh, PA).

4.1.4 Procedure

Participants were tested individually in a quiet room. Instructions were given in English. Participants were presented with a Spanish prime followed by a Spanish target word at the center of a computer screen. They were instructed to name (read aloud) in Spanish the target word as quickly and accurately as possible. Their verbal responses were tape recorded for later coding of accuracy. Prior to the presentation of each prime, a fixation cross was presented until the participant initiated the beginning of the trial by pressing a key on the computer keyboard. The prime was presented for 150 ms followed by a blank screen for 50 ms, resulting in a 200 ms SOA. These short presentation parameters were chosen in an attempt to minimize strategic control (following recommendations form Lucas, 2000; McNamara, 2005; Neely, 1991). The target was then presented until the participant initiated a verbal response, for a maximum of four seconds. A schematic representation of the trial procedure can be seen in Figure 2. Reaction time (RT) was recorded by the computer program in milliseconds (ms) from the onset of target presentation to the onset of articulation. Participants were given 10 practice trials (composed of non-critical words) before beginning the experimental trials.

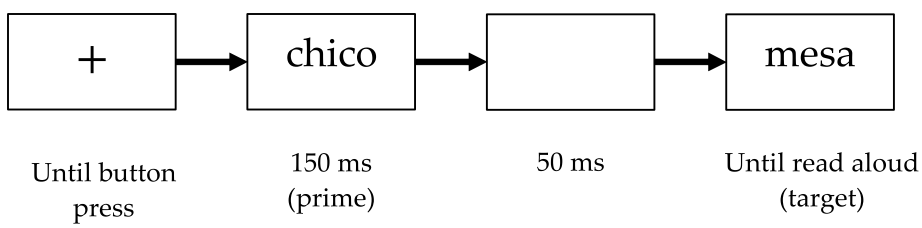

Figure 2. Trial procedure used in the primed naming task

At the end of the naming task, participants completed a printed version of the language history questionnaire used in Experiment 2, including additional questions about immersion experience. This questionnaire was written in English, but participants were provided with a Spanish version if they had difficulty reading the English version.

4.2 RESULTS

4.2.1 Data trimming

Based on the distribution of scores, reaction times below 350 ms or above 850 ms, those that were 2.5 SDs above or below a given participant's mean response for correct trials, and from trials on which the voice key failed (either incorrectly registered a response or failed to register a response) were removed from the analyses and were treated as missing values in the analysis by participants. These procedures resulted in the exclusion of less than 7% of the data. Reaction time analyses were conducted using data from correct trials only.

Data by participants were analyzed using a repeated measures ANOVA with gender match and gender type as within participants variables, and are reported as F_1. Data by items were analyzed using a repeated measures ANOVA with gender match as a within item variable and gender type as between items variable, and are reported as F_2. The mean reaction time and percent correct for items from each condition are presented in Table 5.

4.2.2 Reaction time

The main hypothesis, that targets would be processed more quickly in gender matched pairs compared to unmatched pairs, was not supported by the reaction time data: The effect of gender match was not significant in the analysis by participants, $F_1 < 1$, or in the analysis by items, $F_2 < 1$. The effect of gender type was not significant, F_1 (1, 23) = 2.89, MSE = 227.99, $p > .10$; F_2 (1,

78) = 1.52, MSE = 930.32, $p > .20$, nor was the interaction, $F_1 < 1$; $F_2 < 1$. Therefore, having a gender match between prime and target did not affect speed of processing of the target.

4.2.3 Accuracy data

In the accuracy analysis we again found no significant effect of gender match in the analysis by participants, $F_1 < 1$, or in the analysis by items, $F_2 < 1$. As with the reaction time data, the effect of gender type, $F_1 < 1$, $F_2 < 1$, and the interaction, F_1 (1, 23) = 2.86, MSE = .001, $p > .10$; F_2 (1, 78) = 2.61, MSE = .002, $p > .10$, also were not significant.

4.3 DISCUSSION

We found no effect of grammatical gender match on the speed or accuracy of naming the target. Although it is possible that such an effect would be observed with more participants or items (because the observed power for the match effect was only .06), it is interesting to note that the pattern of means suggests that gender match pairs were processed more slowly (M = 531.45) and less accurately (M = 97.50) than unmatched pairs (M = 530.40, M = 97.70), in contrast to what would be predicted by the linguistic relativity hypothesis.

5.0 EXPERIMENT 3B: SEMANTIC AND SYNTACTIC PRIMING OF GRAMMATICAL GENDER IN SPANISH

We did not find an effect of match in grammatical gender in the semantic priming task used in Experiment 3A. To examine if this null effect could be due to the specific task parameters used (e.g., SOA, naming, etc.), we created a second version of this experiment in which the composition of stimuli was different. In this version, half of the pairs were comprised of adjectives as primes and the same inanimate nouns as targets. This composition can give rise to syntactic priming, in that adjectives that don't match the target noun in gender may create a syntactic mismatch to which the participants may be sensitive. Such syntactic priming effects have been shown in previous research. For example, Bates, Devescovi, Hernandez, and Pizzamiglio (1996) auditorily presented adjective-noun pairs to native Italian speakers and found an advantage in reaction time for matched pairs in three tasks: (1) word repetition; (2) gender monitoring; and (3) grammatical judgment. We therefore included a grammatical gender and syntactic gender condition in the present experiment; half the pairs in each condition matched in gender and half did not. We examined whether a semantic grammatical gender match effect (between two nouns in the biological and grammatical gender condition) would be observed in the context of the expected syntactic priming effect, in the adjective-inanimate noun condition.

5.1 METHOD

5.1.1 Participants

Fifteen native Spanish speakers (7 men; mean age 31.7 years) participated in this experiment. Ten additional participants were excluded from the analysis (three had learned another language along with or before Spanish; data or language history background information for two additional participants were lost because of technical errors; five additional participants were excluded for not following the task instructions, as indicated by very low accuracy or very high reaction times). The language history questionnaire data for the final set of fifteen participants are shown in Table 1.

5.1.2 Design

A 2 match type (biological and grammatical gender vs. syntactic gender) by 2 gender match (matched vs. unmatched) within-participants design was used.

5.1.3 Stimuli

The stimuli were 160 Spanish nouns and 40 Spanish adjectives. Each participant saw 80 pairs. Half of the pairs comprised the *biological and grammatical gender* condition (animate noun prime – inanimate noun target) identical to those used in Experiment 3A. The other half of the pairs comprised the s*yntactic gender* condition (adjective prime- inanimate target). As in Experiment 3A, half of the pairs matched in gender, half of the stimuli were feminine, and stimuli were matched across conditions for Spanish frequency (taken from Carmona, et al., 1998) and length. See Table 6 for examples of stimuli in the different conditions, and Table 7 for the properties of stimuli in the different conditions.

Eight different pairings of primes and targets were randomly created, and the order of presentation of the prime-target pairs within each list was randomized for each participant by the computer program (E-Prime software, Psychology Software Tools, Pittsburgh, PA). As in Experiment 3A, an effort was made to ensure that none of the pairs conveyed an obvious associative relation between the prime and the target.

5.1.4 Procedure

The procedure was identical to that of Experiment 3A.

5.2 RESULTS

5.2.1 Data trimming

The same trimming and exclusion criteria as in Experiment 3A were employed, resulting in the exclusion of 8.4% of the data. Data by participants were analyzed using a repeated measures ANOVA with match type and gender match as within participants variables, and are reported as F_1. Data by items were analyzed using a repeated measures ANOVA with gender match as a within item variable and match type as between items variable, and are reported as F_2. The mean reaction time and percent correct for each condition are presented in Table 5.

5.2.2 Reaction time

The main hypothesis, that targets would be processed more quickly in gender matched pairs compared to unmatched pairs, was not supported by the reaction time data: The effect of gender match was not significant in the analysis by participants, F_1 (1, 14) = 3.06, MSE = 160.79, $p >$ .10, or in the analysis by items, $F_2 < 1$. The effects of match type, F_1 (1, 14) = 2.03, MSE = 172.58, $p > .17$; F_2 (1, 78) = 1.36, MSE = 994.43, $p > .24$, and the interaction, $F_1 < 1$; $F_2 < 1$, were not significant either.

Therefore, having a gender match between prime and target did not affect speed of processing of the target, in either the biological and grammatical condition or in the syntactic gender condition.

5.2.3 Accuracy data

In the accuracy analysis we again found no significant effect of gender match in the analysis by participants, $F_1 < 1$, or in the analysis by items, $F_2 < 1$. As with the reaction time data, the effect of match type, $F_1 < 1$; $F_2 < 1$, and the interaction, F_1 (1, 14) = 1.14, MSE = .001, $p > .30$; F_2 (1, 78) = 1.91, MSE = .003, $p > .17$, did not reach significance[7].

[7] Experiment 3B was also administered to a group of native English speakers who were learners of Spanish (1-4 semesters) to examine if gender match is more salient to learners. Because learners are trying to learn the gender of nouns, it is possible that they construct unique strategies that would allow them to better remember the gender assignment, such as thinking of semantic aspects of the gendered nouns that can correlate with the gender assignment (see Sera et al., 2002, p. 385, for similar ideas). With 26 participants the effects of gender match, match type, and the interaction between them were not significant in the RT, F_1 (1, 25) = 1.16, MSE = 567.07, $p > .29$; $F_1 < 1$; $F_1 < 1$, or the accuracy data, all *Fs < 1*.

5.3 DISCUSSION

In contrast to our hypothesis, and to previous research (e.g., Bates et al., 1996), we did not obtain a gender match priming effect in the syntactic gender condition, in which adjectives either did or did not match the targets that followed them in gender. We should keep in mind that with only 15 participants, the observed power for the match effect in the adjective condition was only 0.26.

In addition, there are three possible reasons why we did not find this type of priming effect. First, although adjectives can precede the noun in Spanish, it is more common that adjectives follow the noun. This may have made our native Spanish speakers less sensitive to the syntactic violation than Bates et al.'s native Italian speakers, because prenominal adjectives are more common in Italian.

Second, we used a naming task in the semantic priming paradigm. As mentioned before, we chose naming and not lexical decision because naming is considered less open to strategic control (Lucas, 2000; Neely, 1991). We also wanted to allow the recruitment of the production system of our participants to encourage gender selection. However, because Spanish has a transparent orthography (i.e., regular correspondence between graphemes and phonemes), participants may have been able to read the targets by computing phonology from orthography, without activating the semantic representations of the words (for similar considerations with Spanish see Dominguez, de Vega, & Cuetos, 1997; see Tabossi & Laghi, 1992, for a comparison of English and Italian). If indeed the meaning was not activated, we would have less reason to expect lexico-semantic effects. A lexical decision task with carefully-chosen nonwords foils is more likely to require participants to access meaning representation when they perform the task.

Finally, the binary nature of gender may operate differently than other semantic features in a priming paradigm. It is therefore necessary to establish that gender match can lead to facilitation in a priming task under different conditions. To establish this we adopted a primed lexical decision task, in which participants are more likely to activate the lexical item to perform the task, given that the nonwords are pronounceable pseudowords. In addition to employing the lexical decision task, Experiment 4 examines biological gender match in English, and its effect in a priming paradigm, in which pure semantic match is also manipulated.

6.0 EXPERIMENT 4: PRIMED LEXICAL DECISION OF BIOLOGICAL GENDER IN ENGLISH

The results of Experiment 1 indicated that biological gender influences the perceived meaning similarity of pairs of English words. In the present experiment we extend these results and examine whether this semantic dimension influences online word processing, specifically in a primed lexical decision task. Given that biological gender contributes to the perceived similarity of word pairs, we expect animate target words to be processed more quickly and/or accurately in a priming task in pairs that share a biological gender than in pairs that do not share a biological gender. Thus, we expect “mother” to be identified as a word more quickly when it follows a prime like “princess” relative to when it follows a prime like “prince”. As in Experiment 1, we included a control condition with meaning related and unrelated pairs of words that do not convey biological gender. This will allow us to examine the effects of biological gender match in the context of other semantic dimensions.

6.1 METHOD

6.1.1 Participants

Ninety-six right handed native English speakers (38 males; mean age 18.6 years) participated in this experiment for Introductory Psychology class credit. Nine additional participants were replaced because their language history questionnaire indicated that they had learned another language along with English, or that they were left handed.

6.1.2 Design

A 2 relatedness (related vs. unrelated) by 2 relation type (biological gender vs. pure semantic) within-participants design was used.

6.1.3 Stimuli

The stimuli consisted of 318 words and 280 orthographically legal and pronounceable nonwords. The words used in this experiment included the 186 critical words from Experiment 1, and an additional 122 words (18 taken from McRae & Boisvert, 1998; and 114 taken from Plaut & Booth, 2000) that served as primes or targets in the filler nonword trials. The distribution of words in each condition is shown in Figure 3.

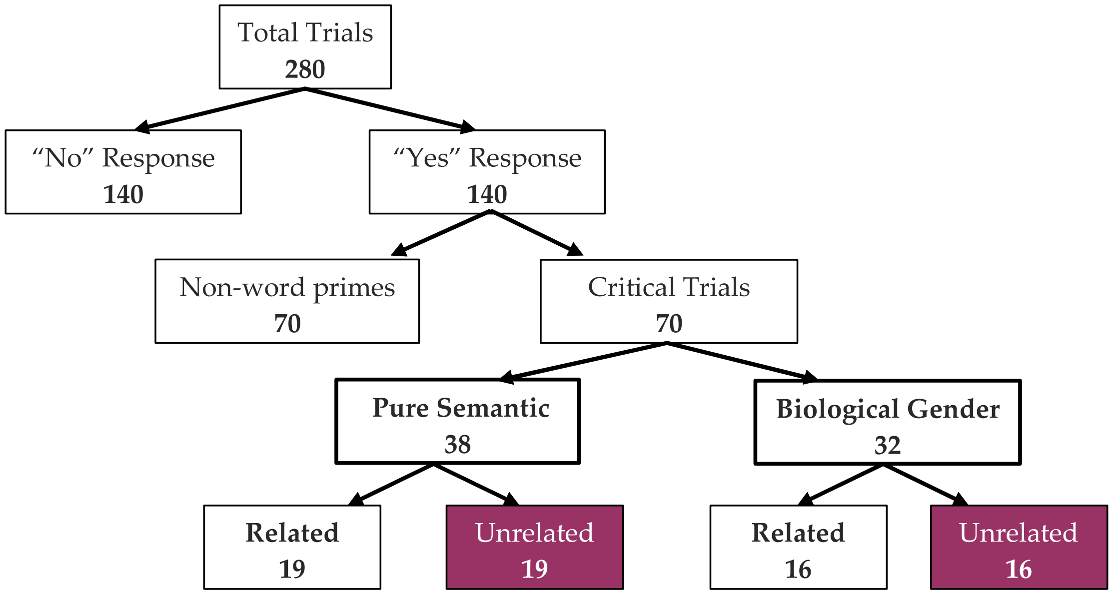

Figure 3. A schematic distribution of trials in Experiment 4

A total of 280 nonword stimuli were used, all pronounceable and orthographically legal. Of these, 84 nonwords were taken from the set used by Plaut and Booth (2000), and an additional 196 nonwords were constructed using the Elexicon (Balota et al., 2002), and the ARC nonword databases (Rastle, Harrington, & Coltheart, 2002). Nonword primes were matched in length to the critical primes, and nonword targets were matched in length to the critical targets.

Each participant completed 280 trials (70 word-word pairs, 70 nonword-word, 70 word-nonword, and 70 nonword-nonword). On each trial a letter string was presented as a prime followed by a second letter string to which participants made a word status judgment (lexical decision). Participants were instructed to make the lexical decision only to the second letter string, and therefore in this stimuli configuration they were to make a “yes” decision on half the trials (140) and a “no” decision on the remaining half (140 trials), yielding a 0.5 nonword ratio (following recommendations from McNamara, 2005).

Half of the 140 yes trials were nonword-word pairs, and the other 70 pairs created the critical relatedness manipulation. Of these 70 trials, 35 were related pairs and 35 were unrelated pairs, creating a 0.25 relatedness proportion (35 out of 140). Of the 35 related pairs, 16 pairs were related in their matching biological gender (e.g., mother-queen) and 19 were related in a pure semantic relation (e.g., whale – dolphin). Similarly, of the 35 unrelated pairs, 16 were unrelated in that the gender of the referents was not the same (e.g., princess- boy), and 19 were semantically unrelated pairs taken from McRae and Boisvert (1998) and Plaut and Booth (2000).

Generally, related primes were matched for length and frequency to the unrelated primes. Furthermore, eight different versions were created, counterbalanced across participants. For each version, related pairs were matched to unrelated pairs in terms of length and frequency of the primes and targets, and the mean RT to make a lexical decision to the target (taken from Elexicon, Balota et al., 2002).

Each of the eight versions had different prime-target pairings of the gender condition, and had a different set of 16 targets in the pure semantic condition that were preceded by a related prime. In each version, the 35 related pairs (16 gender condition, and 19 pure semantic condition) were not significantly different from the 35 unrelated pairs in terms of length, frequency, and mean RT to make a lexical decision to the target (taken from Elexicon, Balota et al., 2002).

6.1.4 Procedure

Participants were tested individually in a quiet room. Instructions were given in English. Participants were told that they would be presented with pairs of letter strings, and that they were to read both letter strings but respond only to the second string. As soon as the second string appeared they were to make a lexical decision as quickly and accurately as possible, by pressing the Yes button (the “f” key on the keyboard) with their left index finger if it was a real word in English, or pressing the No button (the “j” key on the keyboard) with their right index finger if it was not a real word in English.

On each trial a fixation cross appeared in the center of screen until the participant initiated the beginning of the trial by pressing the space bar. When the space bar was pressed, the prime letter sting was presented for 150 ms followed by a blank screen for 50 ms, resulting in a 200 ms SOA. These short presentation parameters, identical to those used in Experiment 3, were chosen to minimize strategic control (Lucas, 2000; Neely, 1991). Then the target letter string appeared until the participant responded. Accuracy and reaction time were recorded by the computer program (Eprime software, Psychology Software Tools, Pittsburgh, PA). Reaction times were calculated in milliseconds from the onset of target presentation to the onset of the participant’s response. A schematic representation of a trial procedure is presented in Figure 4. A total of 280 critical trials, preceded by 10 practice trials were completed. Upon completion of the lexical decision task, participants filled out a short language history questionnaire similar to that used in Experiment 1.

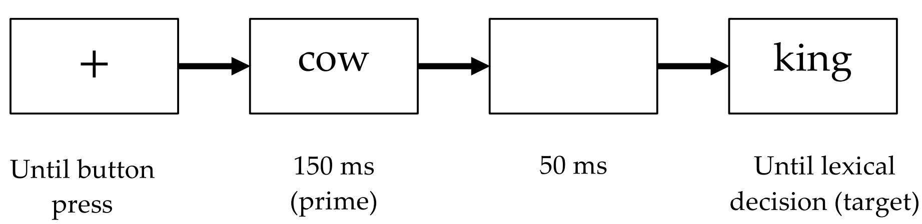

Figure 4. Trial procedure used in the primed lexical decision task

6.2 RESULTS

6.2.1 Data trimming

The statistical analyses were conducted on data from a final set of 96 participants. Based on the distribution of scores, reaction times below 350 ms or above 2000 ms, and those that were 2.5 SDs above or below a given participant's overall mean response were removed from the analyses and were treated as missing values in the analysis by participants. These procedures resulted in the exclusion of 7.1% of the data. Reaction time analyses were conducted using data from correct trials only.

Data by participants were analyzed using a repeated measures ANOVA with relatedness and relation type as within participants variables, and are reported as F_1. Data by items were analyzed using a repeated measures ANOVA with relatedness as a within item variable and relation type as between items variable, and are reported as F_2[8]. The mean reaction time and percent correct for items from each condition are presented in Table 5.

6.2.2 Reaction time

The main effect of relation type was significant in the analysis by participants, F_1 (1, 95) = 16.53, MSE = 2547.40, $p < .001$, suggesting that pure semantic condition targets were processed more quickly (M = 703.61) than gender condition targets (M = 724.55). However, this effect

[8] As in Experiment 1, F_2 was computed based on the target items, which were primed by related or unrelated items. A total of 105 targets were included in these analyses, 38 in the pure semantic condition and 67 in the gender condition (as explained in Footnote 1). Filler non-word trials were not included in the analyses.

should be interpreted with caution, because the effect was not significant in the analysis by items, F_2 (1, 103) = 1.86, MSE = 10735.87, $p > .10$, and because, as in Experiment 1, stimuli in the two conditions were not matched for important linguistic characteristics[9].

Of relevance to our main hypothesis, the effect of relatedness did not reach significance in the analysis by participants, F_1 (1, 95) = 2.15, MSE = 2920.91, $p > .10$, or by items, F_2 (1, 103) = 2.79, MSE = 2425.69, $p > .09$. It is worth noting, that although this effect is not significant, the pattern of means does follow our predictions with targets in related pairs being processed more quickly (M = 710.03) than targets in unrelated pairs (M = 718.13). The interaction between relation type and relatedness also was not significant by participants $F_1 < 1$, or by items, $F_2 < 1$.

6.2.3 Accuracy data

In the accuracy analysis, the effect of relation type was marginally significant in the analysis by participants, F_1 (1, 95) = 3.71, MSE = .002, p =.057, with targets in the pure semantic condition responded to marginally more accurately (M = .975) than targets in the gender condition (M = .966). This effect was not significant in the analysis by items, F_2 (1, 103) = 1.07, MSE = .004, $p > .30$[10]. There were no other significant effects in the accuracy data, Fs>1.

---

[9] When we covaried length and log frequency, the effect was still not significant in the analysis by items, $F_2 < 1$.

[10] When we covaried length and log frequency, the effect of relation type was still not significant in the analysis by items, $F_2 < 1$. However, in this repeated measures ANCOVA the effect of relatedness became significant, F_2 (1, 101) = 5.60, MSE = .001, $p < .05$. Relatedness also significantly interacted with log frequency F_2 (1, 101) = 5.47, MSE = .001, $p < .05$. Because the effect of relatedness was not significant in the F_1, and only became significant in F_2 when length and log frequency are taken into account, and furthermore was only reliable in the accuracy and not in the reaction time analysis, we should interpret it with caution.

6.3 DISCUSSION

Contrary to our predictions, we did not obtain a significant relatedness effect using the primed lexical decision task. Even in the pure semantic condition, which was constructed using stimuli that have been shown to produce significant priming effects in the past (McRae & Boisvert, 1998; Plaut & Booth, 2000), the effect was not significant. However, it is important to keep in mind that effect sizes of pure semantic priming are relatively small (Cohen's d = .39, Lucas, 2000), and that the pattern of means does follow our prediction, in that related pairs in the pure semantic condition were processed more quickly and accurately (M = 697.79, M = .98) than unrelated pairs (M = 709.43, M = .97). The difference in the gender condition between related (M = 722.27, M = .97) and unrelated pairs (M = 726.82, M = .97) is considerably smaller.

7.0 GENERAL DISCUSSION

In four experiments we examined the semantic role of biological and grammatical gender for speakers of English and Spanish. With respect to biological gender, as denoted by the English language, we found that native English speakers consider pairs of words that share biological gender (e.g., queen-cow) to be more similar in meaning than pairs that do not share a gender (e.g., king-waitress). However, match in biological gender was not sufficient to produce a priming effect in a lexical decision task (Experiment 4). With respect to grammatical gender, as denoted by Spanish, we found that in contrast to the linguistic relativity hypothesis, pairs that match in grammatical gender did not elicit higher semantic similarity ratings from native Spanish speakers compared to unmatched pairs (Experiment 2), and furthermore these pairs were not processed more quickly or accurately in a primed naming task (Experiments 3A and 3B).

Foundalis (2002) points to a fundamental difference between languages that employ a natural gender system (such as English) and languages that employ a formal or grammatical gender system, like Spanish. In a natural gender language, the gender can be predicted based on the semantics of the nouns (in English reflected only in personal, possessive, and reflexive pronouns). Speakers of such languages can therefore easily notice the correspondence between sex and gender, because the majority of masculine or feminine nouns are nouns with respective sex related features. We would therefore expect gender effects in these languages to be semantic in nature. Foundalis claims that speakers of grammatical gender languages, such as Spanish, can not avoid noticing the un-relatedness of sex and gender, because only the minority of masculine and feminine nouns truly posses the corresponding sex related features (i.e., denote people). In his analysis of the correlation between grammatical gender assignments across 14 Indo-European languages and intuitions of monolingual English speakers, he points to the remarkable consistency in the ratings made by English speakers (supporting their tendency to assume gender is semantically based) but to the low magnitude of correlation between grammatical gender languages (that are not part of the same subfamilies). This point of view can be considered as an extension of the Sex and Gender Hypothesis discussed by Vigliocco et al. (2005). According to Foundalis, an association between sex and gender is not only absent in a three gendered language like German, but it is also absent in any formal gender language, because not enough correspondence between sex and gender is present to allow generalization to other nouns.

Indeed, when we consider the semantic similarity rating results we see that native English speakers perceive pairs that match in gender as more similar in meaning. It is therefore possible that biological gender exerts an influence on the organization of semantic representations, but it is not likely to be a strong dimension by which lexical representations are grouped. Furthermore, biological gender in English is inherently part of only a subset of animate referents (e.g., “mother” and “bull” but not “speaker” or “elephant”). Because we made use only of animate nouns that obligatorily denote a biological gender, we suspect these effects to be even weaker for other animate nouns that only probabilistically or stereotypically denote a biological gender.

Grammatical gender did not appear to influence perceived similarity rating or primed naming performance in Spanish, in contrast to the predictions made under the linguistic relativity hypothesis. Specifically, these results do not support the Similarity and Gender Hypothesis discussed by Vigliocco et al., (2005) which would predict that words that share a gender would be perceived as more similar in meaning, and would behave as semantically related by virtue of their shared linguistic contexts. Because in Spanish grammatical gender is coupled with biological gender for animate nouns, and because we did not include an animate-animate pair condition, we can not speak to the validity of the Sex and Gender Hypothesis, which would predict semantic effects of grammatical gender only or mostly for animate referents.

If one assumes a functional separation between syntactic and conceptual aspects of language, our results are to be expected, because the grammatical distinction of gender should be processed separately from the meaning of objects. In accordance with this approach, Bowers et al. (1999) showed that participants were faster or equally fast at making a semantic decision (artifact vs. natural kind) to pairs of pictures compared to pairs of words, but were much faster responding to the words, compared to pictures, when making a grammatical gender decision (choosing the appropriate determiner) or a count/mass decision which is more syntactic in nature. The authors take their results to suggest that syntactic features (such as grammatical gender) are strictly linked to the lexical representation of a word rather than to its corresponding concept.

As mentioned earlier, the naming task we employed when examining the on-line effects of grammatical gender did not necessarily require participants to activate the lexical representation to perform the task because of the transparent nature of Spanish orthography. Presumably, a primed lexical decision task would allow such semantic effects to arise. However, when using this paradigm with English speakers, we did not find significant biological gender match effects, which leaves less reason to expect semantic effects of grammatical gender to arise in such a paradigm. Furthermore, if the naming task itself was what hindered our ability to detect an effect we would expect to observe a grammatical gender match effect in other paradigms. In our study, we did not observe such an effect in the perceived similarity rating task. In addition, Versace and Allain (2001) did not observe a grammatical gender match effect when they asked their French speaking participants to make a semantic categorization.

However, it would be premature to conclude that the prediction of the linguistic relativity hypothesis, that grammatical gender will manifest semantic influences, is categorically wrong. First a few important characteristics of our experiments need to be considered, and second, as will be discussed below, some recent research suggests semantic effects of grammatical gender are evident under certain circumstances.

When examining the null effect observed in some of the current experiments we need to consider the issue of power. Experiments 2, 3, and 4, in which no significant match effects were observed, suffered from low observed power ranging from .06 (Experiments 3A and 3B - semantic priming of grammatical gender in Spanish) to .37 (Experiment 2 – similarity rating of grammatical gender in Spanish) and .31 (Experiment 4 – primed lexical decision with biological gender in English). This is partly due to the small effect sizes of semantic priming, especially using a naming task. However, in the experiments that examined grammatical gender match in Spanish (i.e., Experiments 2, 3A), the pattern of means suggested an advantage for the unmatched pairs over the matched pairs.

Second, the native Spanish speakers who took part in our experiments were very diverse in their backgrounds. In Experiment 3A and 3B we were not able to control for the age at which our participants started to learn a second language. It is possible that grammatical gender becomes internalized as part of the semantic representation of objects only when the child is not exposed to other languages with a grammatical gender system that assigns gender differently. Such an exposure is likely to emphasize for the child the arbitrariness of grammatical gender assignment compared to biological gender. Indeed, when we examined the semantic effects of grammatical gender for native English speakers who learned Spanish later in life, we found no effect of matched gender. These results would be expected if grammatical gender is not internalized for words learned later on, when the arbitrariness of the grammatical gender is hard to ignore.

In Experiment 2 we were better able to control the age at which participants started learning a second language, however because this experiment was completed via a web-based interface, there was great variability in the participants' country of origin. In his discussion of linguistic relativity, Lucy (1997) considers this an advantage, because it allows a separation of potential effects of language from effects of culture. The diversity of our participants allows us to assume that any observed effect on semantic representation is a result of the shared language spoken by our participants and not a result of a shared culture. However it is possible that this variability in background, which may have introduced colloquial differences in language use, hindered our ability to detect a significant effect of grammatical gender match.

Finally, some recent evidence suggests that perhaps under specific conditions grammatical gender exerts semantic influences. Lotto, Paolieri, Cubelli, and Job (2007) presented Italian, Spanish, and English speakers with pairs of pictures and asked them to indicate whether the two pictures in each pair belonged to the same semantic category (i.e., a semantic categorization task). Importantly, half of the pairs had pictures that referred to objects with the same grammatical gender in Italian (congruent gender) but different grammatical genders in Spanish (incongruent gender), and in the other half the pictures' referents belonged to different genders in Italian but the same gender in Spanish. The results suggest that both the Italian and the Spanish speakers, but not the English speakers, performed the semantic categorization task more quickly when the pair of pictures matched in grammatical gender in their native language than when the genders of the two pictures were incongruent. These results suggest that even in a conceptual task that uses pictures and not words, participants' performance was influenced by the match in grammatical gender in their native language.

It appears, then, that if grammatical gender plays a semantic role, it may be limited to representation of some nouns but not others (e.g., animate but nor artifacts, Vigliocco et al., 2005), to some languages but not others (e.g., Italian and Spanish, but not German, Sera et al., 2002; Vigliocco et al., 2005), and perhaps to participants with certain linguistic backgrounds (e.g., monolinguals not exposed to other languages early in life, children above second grade, Sera et al. 1994; 2002). Furthermore, the effect may not be strong enough to manifest itself in every task.

As for biological gender, we have clearly shown that for speakers of a natural gender system like English, nouns that share a biological gender are perceived as more similar in meaning, but this similarity is not sufficient to facilitate on-line word processing in a primed lexical decision task. The semantic role played by biological gender for speakers of Spanish or other grammatical gendered languages remains to be discovered.

BIBLIOGRAPHY

Andonova, E., D'Amico, S., Devescovi, A., & Bates, E. (2004). Gender and lexical access in Bulgarian. *Perception & Psychophysics, 66,* 496-507.

Balota, D. A., Cortese, M. J., Hutchison, K. A., Neely, J. H., Nelson, D., Simpson, G. B., Treiman, R. (2002). The English Lexicon Project: A web-based repository of descriptive and behavioral measures for 40,481 English words and nonwords. http://elexicon.wustl.edu/, Washington University.

Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Perception & Psychophysics, 58,* 992-1004.

Boroditsky, L., & Schmidt, L. (2000). Sex, Syntax, and Semantics. *Proceedings of the 22nd Annual Meeting of the Cognitive Science Society*, Philadelphia, PA.

Boroditsky, L., Schmidt, L., & Phillips, W. (2003). Sex, Syntax, and Semantics. In Gentner & Goldin-Meadow (Eds.,) *Language in Mind: Advances in the study of Language and Cognition*. MIT Press: Cambridge, MA.

Bowers, J. S., Vigliocco, G., Stadthagen-Gonzalez, H., & Vinson, D. (1999). Distinguishing language from thought: Experimental evidence that syntax is lexically rather than conceptually represented. *Psychological Science,10,* 310-315.

Carmona, J., Cervell, S., Màrquez, L., Martí, M. A., Padró, L., Placer, R., Rodríguez, H., Taulé, M., & Turmo, J. (1998, May). An environment for morphosyntactic processing of unrestricted Spanish text. *Proceedings of the 1st International Conference on Language Resources and Evaluation, LREC,* pp. 915—922, Granada, Spain.

Costa, A., Kovacic, D., Fedorenko, E., & Caramazza, A. (2003). The gender congruency effect and the selection of freestanding and bound morphemes: Evidence from Croatian. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 1270-1282.

Dominguez, A., de Vega, M., & Cuetos, F. (1997). Lexical inhibition from syllabic units in Spanish visual word recognition. *Language and Cognitive Processes, 12,* 401-422.

Eberhard, K. M., Scheutz, M., & Heilman, M. (2005). An empirical and computational test of linguistic relativity. In B. B. Bara, L. Barsalou, & M. Bucclarelll (Eds.), *Proceedings of the 27th Annual Cognitive Science Society Conference* (pp. 618-623). Stresa, Italy. Mahwah, N.J.: Lawrence Erlbaum Associates, Inc

Flaherty, M. (2001). How a language gender system creeps into perception. *Journal of Cross-Cultural Psychology, 32,* 18-31.

Foundalis, H. E. (2002, August). *Evolution of gender in Indo-European languages*. Paper presented at the 24th Annual Conference of the Cognitive Science Society, Fairfax, VA.

Gentner, D., & Goldin-Meadow, S. (2003). *Language in mind: Advances in the study of language and thought*. Cambridge, MA, US: MIT Press.

Konishi, T. (1993). The semantics of grammatical gender: A cross-cultural study. *Journal of Psycholinguistic Research, 22,* 519-534.

Lotto, L., Paolieri, D., Cubelli, R., & Job, R. (2006, November). *The effect of grammatical gender on the semantic categorization of pictures: evidence from Italian, Spanish, and English.* Poster presented at the 47th Annual Meeting of the Psychonomic Society, Houston, Tx.

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review, 7,* 618-630.

Lucy, J. A. (1997). Linguistic Relativity. *Annual Review Anthropology, 26,* 291-312. Palo Alto, CA: Annual Reviews Inc.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments & Computers, 28,* 203-208.

Marx, E. (1999). Gender processing in speech production: Evidence form German speech errors. *Journal of Psycholinguistic Research, 28,* 601-621.

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. Essays in cognitive psychology, New York, NY, US: Psychology Press.

McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 558-572.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G.W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264-336). Hillsdale, NJ, Englandiates, Inc.

Phillips, W., & Boroditsky, L. (2003). Can Quirks of grammar affect the way you think? Grammatical gender and object concepts. *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*. Boston, MA.

Plaut, D. C., & Booth, J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychological Review, 107*, 786-823.

Plemmenou, E., Bard, E. G., & Branigan, H. P. (2002). Grammatical gender in the production of single words: Some evidence from Greek. *Brain and Language, 81*, 236-241.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. Quarterly Journal of Experimental Psychology, 55A, 1339-1362.

Scheutz, M. J., & Eberhard, K. M. (2004). Effects of morphosyntactic gender features in bilingual language processing. *Cognitive Science, 28,* 559-588.

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

Schweppe, J., & Rummer, R. (2007). Shared representations in language processing and verbal short-term memory: The case of grammatical gender. *Journal of Memory and Language, 56*, 336-356.

Sera, M. D., Berge, C. A .H., & del Castillo-Pintado, J. (1994). Grammatical and conceptual forces in the attribution of gender by English and Spanish speakers. *Cognitive Development, 9,* 261-292.

Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodríguez, W., & Dubios, D. P. (2002). When language affects cognition and when it does not: An analysis of grammatical gender and classification. *Journal of Experimental Psychology: General, 131*, 377-397.

Tabossi, P., & Laghi, L. (1992). Semantic priming in the pronunciation of words in two writing systems: Italian and English. *Memory & Cognition*, *20,* 303-313.

Tokowicz, N., Kroll, J. F., De Groot, A. M. B., & Van Hell, J. G. (2002). Number-of-translation norms for Dutch-English translation pairs: A new tool for examining language production. *Behavior Research Methods, Instruments, and Computers, 34*, 435-451.

Tokowicz, N., Michael, E. B., & Kroll, J. F. (2004). The roles of study-abroad experience and working-memory capacity in the types of errors made during translation. *Bilingualism*: *Language and Cognition, 7,* 255-272.

Versace, R., & Allain, G. (2001). Negative priming in a gender decision task and in a semantic categorization task. *Acta Psychologica, 108,* 73-90.

Vigliocco, G., Vinson, D. P., Indefrey, P., Levelt, W. J. M., & Hellwig, F. (2004). Role of grammatical gender and semantics in German word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 483-497.

Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical gender effects on cognition: implications for language learning and language use. *Journal of Experimental Psychology: General, 134,* 501-520.

Whorf, B. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (J.B. Carroll, Ed.). Cambridge, MA: MIT Press.