Grammatical Gender Affects Bilinguals' Conceptual Gender: Implications for Linguistic Relativity and Decision Making

James N. Forbes*[,1], Diane Poulin-Dubois[2], Magda R. Rivero[1] and Maria D. Sera[3]

[1]*Angelo State University, Department of Psychology & Sociology, ASU Station 10907, San Angelo, TX 76909-0907, USA*

[2]*Centre for Research in Human Development, Concordia University, Department of Psychology (PY-170), 7141 Sherbrooke Street West, Montréal, Québec H4B 1R6, Canada*

[3]*Institute of Child Development, University of Minnesota, 51 East River Road Minneapolis, MN 55455, USA*

Abstract: We used a non-linguistic gender attribution task to determine how French and Spanish grammatical gender affects bilinguals' conceptual gender. French-English and Spanish-English bilingual, as well as English monolingual adults were asked to assign a male or female voice to 32 color drawings depicting people, animals, and common objects. French-English and Spanish-English bilinguals classified items according to French and Spanish grammatical gender respectively. This effect was replicated for French-English bilinguals on those items whose grammatical gender was opposite in French and Spanish. Unexpectedly, Spanish gender similarly affected classifications by Spanish-English and English-Spanish bilinguals, as well as English monolinguals. We discuss how grammatical gender, possible covariates, and the order of L1 and L2 acquisition, affect conceptual gender as well as implications for decision making.

GRAMMATICAL GENDER AFFECTS BILINGUALS' CONCEPTUAL GENDER: IMPLICATIONS FOR LINGUISTIC RELATIVITY AND DECISION MAKING

A sizeable literature of empirical research on linguistic relativity, the effect of language on thought, began accumulating in earnest during the 1950s. Predictably, some studies adduce evidence strongly supporting linguistic relativity [e.g., 1-7] while other studies are more dismissive of linguistic relativity claims [e.g., 8-13]. This is hardly surprising, since linguistic relativity means different things to different researchers who have relied on a wide variety of investigative methods. Mindful of the ambiguity of what different researchers have meant by relativity, Fishman [14, 15] distinguishes four levels of linguistic effects on thinking: from superficial effects of specialized lexical terms to more profound effects of grammatical structure on individuals' behavior. The present study addresses the more substantive effects of language on cognition. Of particular interest is whether differences in grammatical gender between languages affect bilinguals' conceptual gender classifications.

An example of research on how grammatical gender affects categorization is reported by Sera, Berge, and Castillo [6]. They investigated whether monolingual speakers of Spanish and English, two languages with different systems of grammatical gender, also categorize objects differently. Spanish nouns must be classified as masculine (e.g., el fuego) or feminine (e.g., la estrella). English, on the other hand, does not systematically classify nouns into different categories of grammatical gender. Indeed, only some English words even explicitly incorporate a referent's natural gender with word meaning (e.g., boy, girl, ram, yew; but not writer, painter). In Sera *et al.*'s procedure, Spanish and English monolingual adults and children, from 5 to 10 years of age, viewed pictures of everyday items in one of two conditions. In the label plus pictures condition, participants were shown each picture and heard the experimenter label it in their native language. In the picture only condition, participants were shown the same pictured items, but they were not specifically labeled. In both conditions, participants were asked to attribute either a male or a female voice to animate the pictured item. Spanish speakers, unlike English speakers, were more likely to classify items according to Spanish grammatical gender. This tendency was most pronounced when the pictured items were explicitly labeled. Furthermore, the tendency to categorize objects in accord with Spanish grammatical gender was observed by the time children were 8 years old.

Like Sera *et al.* [6], the present study investigated the role of grammatical gender on categorization across different languages: French, Spanish, and English. Comparisons of two or more languages is a propitious use of a naturally occurring opportunity to study meaningful language effects on cognition [2]. But unlike previous research, we compared French-English, English-French, Spanish-English, and English-Spanish bilinguals', performance on the same cognitive task. Using bilinguals who differ in the order (L1 or L2) in which they acquired a grammatical gender system offers a within-participants crosslinguistic perspective. Importantly, a within-participants design comprising bilingual samples is a direct test of language acquisition order effects on conceptual gender.

*Address correspondence to this author at the Angelo State University, Department of Psychology & Sociology, ASU Station 10907, San Angelo, TX 76909-0907, USA; E-mail: james.forbes@angelo.edu

Also in accord with prior research, we focus on a significant language variable: grammatical gender. Conveniently, French and Spanish grammatical gender systems are fundamentally similar, but nonetheless show important dissimilarities. First of all, in both languages nouns have a grammatical gender; they are either masculine or feminine. Secondly, nouns, articles, and adjectives in both languages mark grammatical gender with grammatical morphemes. For instance, a big spoon is rendered, une grande cuillère in French, and una cuchara grande in Spanish. Since neither French nor Spanish are case-based languages, word order does not affect grammatical gender, or semantic meaning. In other words, Spanish and French translations of the spoon are unaffected by sentences such as: "The dish ran away with the spoon.", "The spoon ran away with the dish." Because nouns, articles, and adjectives provide multiple sources of grammatical gender information, both languages are said to be gender loaded.

Notably, French and Spanish grammatical gender differ in one useful and important aspect. Some nouns that are grammatically masculine and feminine in French are the opposite grammatical genders in Spanish. Hence, a cloud is masculine in French un nuage, but feminine in Spanish: una nube. In contrast, English does not systematically classify nouns into different categories of grammatical gender. Indeed, only some English words even explicitly incorporate a referent's gender with word meaning (e.g., boy, girl, but not writer, painter).

The present study assessed adult bilinguals' cognitive performance using different items with the gender classification task developed by Sera *et al*. [6, 16]. We expected French-English bilinguals' categorization to accord with French grammatical gender and Spanish-English bilinguals' categorization to accord with Spanish grammatical gender. Moreover, we expected French-English and Spanish-English bilinguals' classifications to differ most on those items where French and Spanish grammatical gender was opposite.

Sera *et al*. [6] found that English speakers were most likely to assign male voices to artifacts and female voices to natural kinds. Because English lacks a grammatical gender system, we expected English monolinguals' categorization to follow the same pattern: artifacts-male, natural kinds-female. Of further interest was whether the order of grammatical gender system acquisition (L1 or L2), the language in which the task was conducted, and specifically labeling test items affected performance on the gender attribution task.

METHOD

Participants

Thirty-two French-English (F-E) bilinguals and 32 English-French (E-F) bilinguals were selected based on their responses to a bilingual history and fluency questionnaire. This questionnaire identified the order in which bilinguals acquired their languages, their age, language history, percent of total language activity spent in their second language, self-report of fluency for each language, as well as the age at which that fluency was attained. French-English bilingual participants were recruited by telephone from a pool of bilingual volunteers as well as from the general population of a large Canadian university situated in metropolitan Quebec.

Thirty-two Spanish-English (S-E) bilinguals and 32 English-Spanish (E-S) bilinguals were selected based on their responses to English and Spanish versions of the bilingual history and fluency questionnaire. Spanish-English bilingual participants were recruited from among the bilingual student population of an American university in Texas, through an Adult Literacy Council, as well as from local churches.

The 32 F-E bilinguals reported that French was their native language; the 32 E-F bilinguals reported that English was their native language; the 32 S-E bilinguals reported that Spanish was their native language; the 32 E-S bilinguals reported that English was their native language. Furthermore, only those bilinguals whose self-reported fluency in their second language was 3 or higher using a 5-point scale

Table 1. Participant Age, Sex, Language Proficiency Level and Category by Language

	Language				
	French-English	**English-French**	**Spanish-English**	**English-Spanish**	**English**
Demographics					
Age	26.03 (9.02)	26.06 (9.95)	31.84 (11.53)	26.53 (8.5)	23.06 (7.64)
Male	14	14	16	15	16
Female	18	18	16	17	16
Proficiency Level					
French	native language	4.03 (0.87)	none	none	
Spanish	none	none	native language	4.12 (0.86)	
English	4.17 (0.75)	native language	4.24 (0.63)	native language	
Proficiency Category					
Early Learner	18	20	20	17	
Late Learner	14	12	12	15	

anchored at each end by the words "weak" (rating of 1) and "fluent" (rating of 5), and who rated their exposure to their second language at least 50% of the time were invited to participate. Bilinguals who acquired proficiency in their second language before age 7 were classified as early learners; those who acquired proficiency in their second language after age 7 were considered late learners [17,18]. Table **1** shows participants' mean ages, sex, language proficiency level, and language proficiency category.

Data from 32 English monolingual speakers were included in the present study. The English monolinguals were recruited at the University of Minnesota, Minneapolis.

Materials

Forty laminated color drawings of common people, animals, and objects were presented to participants. These items were obtained primarily from Pictures Please! [19], as well as the Spanish and English versions of the Peabody Picture Vocabulary Test, [20], all of which are listed in Table **2**. Items were selected for use as stimuli according to their natural gender, their French grammatical gender, and their Spanish grammatical gender. There were two categories of items: control and test. Control items consisted of 8 drawings, half of which depicted a female character and half of which depicted a male character. Thus, natural gender accorded with grammatical gender (e.g., king, bride) for control items. These items were used to assess whether participants understood the experimental task. Test items consisted of 32 drawings, half of which depicted natural kinds (naturally occurring items) (e.g., apple, butterfly), and half of which depicted artifacts (e.g., book, airplane). Half of the natural kinds items were grammatically feminine in French (e.g., banana, spider), whereas half were grammatically masculine in French (e.g., cloud, tree). Furthermore, four of the artifact items were grammatically masculine in French, but grammatically feminine in Spanish (e.g., bed, kite). Four of the artifact items were grammatically feminine in French, but grammatically masculine in Spanish (e.g., fork, car). Similarly, four of the natural kinds items were grammatically feminine in French, but grammatically masculine in Spanish (e.g., tomato, bat). Four of the natural kinds items were grammatically masculine in French, but grammatically feminine in Spanish (e.g., grape, zebra).

Design

Half the participants were assigned to a No Label condition; half to a Label condition. For the French-English and English-French participants, within each condition, half were administered the task in French and half were administered the task in English. For the Spanish-English and English-Spanish participants, within each condition, half were administered the task in Spanish and half were administered the task in English. All participants viewed the control and test items in 1 of 4 randomly determined orders and were tested individually.

Procedure

Participants were initially asked to complete the bilingual history and fluency questionnaire. Next, participants were told that they would be shown a series of 40 color drawings in the context of a gender attribution task. Experimental stimuli were presented one at time in the label and no label conditions. In the label condition, each item was shown and labeled (e.g., "This is a car;" "Voici une voiture," "Aquí está un carro."). Participants heard the labels in the same language as that used for the instructions and questionnaire they were given. Labels were always preceded by an indefinite article regardless of whether the label was given in French or English. Thus, participants tested in French were not specifically told whether an item was grammatically feminine or masculine. In the No Label condition, items were shown but not specifically labeled (e.g., "Here's another drawing.", "Voici la prochaine image." "Aqui tenemos dibujo numero uno.").

For participants tested in English, the exact instructions were:

Imagine that you are making a movie for children in which some everyday objects come to life. So in this film the objects will be able to talk, sing, and dance. Imagine that you are the director of this movie. I am going to show you drawings of the objects that will be brought to life in the film. Your task is to write down on the next sheet of paper, by each number, whether you think each pictured object should have a man's voice (or a boy's voice) or a woman's voice (or a girl's voice) in the movie. Ok, here is drawing number one. [The experimenter would show the participant the drawing and for half the participants, name the depicted item.] Should this [the experimenter named the depicted item for half the participants] have a woman's voice or a man's voice in the movie?

For participants tested in French and Spanish, the instructions were given in French and Spanish respectively.

After completing the experimental task, participants' knowledge of the test items' grammatical gender was tested. For this test, participants were given a printed list of the 40 test and control items used in the gender attribution task. Item names were printed in French or in Spanish with no article (e.g., feu, fuego, balai, escoba). Participants were asked to identify the grammatical gender for each item by circling "M" (masculine) or "F" (feminine).

RESULTS

We conducted three ANOVAS to compare bilinguals' performance on the classification task. For all three ANOVAS, the mean proportion of times participants assigned a male voice to test items was the primary dependent measure. Mean proportions could range from 0 to 1. A mean proportion of 0 would indicate that no participants assigned a male voice to any of the items within a category. A mean proportion of 1 would indicate that all participants assigned a male voice to all the items within a category.

Classifications According to French Grammatical Gender

The first analysis compared French-English bilinguals', English-French bilinguals', and English monolinguals' classifications according to French grammatical gender. Mean classifications (Table **3**) were analyzed using a 3 (Language: French-English vs. English-French vs. English) x 2 (French Gender: masculine vs. feminine) x 2 (Conceptual Kind: natural vs. artifact) x 2 (Condition: label vs. no label) x 2 (Test Language: French vs. English) mixed ANOVA with repeated

Table 2.   Control and Test Items Used in the Study Listed by language, Grammatical Gender, and Conceptual Kind

	English	**French**	**Spanish**
Control Items			
Masculine	a king a man a giant a boy	un roi un homme un géant un garçon	un rey un hombre un gigante un niño
Feminine	a ballerina a woman a bride a girl	une ballerine une femme une mariée une fille	una balarina una mujer una novia una niña
Test Items			
Artifacts		**Feminine**	
	a table a spoon an arrow a guitar	une table une cuillère une flèche une guitare	una mesa una cuchara una flecha una guitarra
		Masculine	
	a plane a book a telescope a knife	un avion un livre un télescope un couteau	un avión un libro un telescopio un cuchillo
Natural Kinds		**Feminine**	
	a spider an apple a star a feather	une araignée une pomme une étoile une plume	una araña una manzana una estrella una pluma
		Masculine	
	a tree a corn (cob) a sun a fire	un arbre un maïs un soleil un feu	un arbol un maís un sol un fuego
Artifacts		**Masculine**	**Feminine**
	a bed a kite a broom a canoe	un lit un cerf-volant un balai un canöe	una cama una cometa una escoba una canoa
		Feminine	**Masculine**
	a screw a fork a plate a car	une vis une forchette une assiette une voiture	un tornillo un tenedor un plato un carro
Natural Kinds		**Masculine**	**Feminine**
	a grape a cloud a zebra a butterfly	un raisin un nuage un zèbra un papillon	una uva una nube una zebra una mariposa
		Feminine	**Masculine**
	a tomato a peanut a bat a banana	une tomate une cacahouète une chauve-souris une banane	un tomate un cacahuate un murciélago un plátano

measures on French Gender and Conceptual Kind. There were main effects for French Gender, $\underline{F}(1,90) = 21.25$, $\underline{p} < .001$; and Conceptual Kind, $\underline{F}(1,90) = 86.04$, $\underline{p} < .0001$. However, French Gender and Conceptual Kind were both involved in significant two-way interactions with Language. Bilinguals and monolinguals reliably differed in their classi-

Table 3. **Mean Proportion of Times (and Standard Deviations) French-English Bilinguals, English-French Bilinguals, and English Monolinguals Assigned a Male Voice to Items by French Grammatical Gender, Conceptual Kind, and Participants' Language**

	French Gender Items		Conceptual Kinds	
	Masculine	Feminine	Artifacts	Natural Kinds
Language				
French-English	0.67	0.43	0.59	0.50
	(0.15)	(0.13)	(0.14)	(0.12)
English-French	0.54	0.52	0.61	0.44
	(0.12)	(0.12)	(0.10)	(0.14)
English	0.53	0.53	0.64	0.43
	(0.11)	(0.12)	(0.10)	(0.10)

fications of items according to French Gender, $\underline{F}(2,90) = 16.85$, $\underline{p} < .0001$. Bilinguals and monolinguals also reliably differed in their classifications of items according to Conceptual Kind, $\underline{F}(2,90) = 4.47$, $\underline{p} < .01$.

Contrasts of the means involved in the French Gender and Language interaction showed that the French-English bilinguals' classifications of grammatically masculine and feminine items significantly differed, dependent $\underline{t}(31) = 6.25$, $\underline{p} < .0001$. No further reliable differences between grammatically masculine and feminine items were noted. Thus, only French-English bilinguals assigned a male voice for French grammatically masculine items significantly more often than for French grammatically feminine items.

Contrasts of the means involved in the Conceptual Kind and Language interaction showed that all participants assigned a male voice to artifacts significantly more often than to natural kind items. Dependent t-tests were as follows: French-English, $\underline{t}(31) = 2.70$, $\underline{p} = 0.01$; English-French, $\underline{t}(31) = 5.81$, $\underline{p} < .0001$; English monolinguals, $\underline{t}(31) = 8.75$, $\underline{p} < .0001$. However, the magnitude of the difference was greatest for English monolinguals and English-French bilinguals. Indeed, only French-English bilinguals attributions of

a male voice to natural kind items failed to differ from chance (50%): $\underline{t}(31) = 0.09$, p > 0.95

Classifications of Items Whose Grammatical Gender was Opposite in French and Spanish

We expected that French-English and Spanish-English bilinguals' classifications would differ most on items where French and Spanish grammatical gender was opposite. Therefore, a second analysis was conducted for participants' classifications for the 16 items whose grammatical gender is opposite in French and Spanish. These mean classifications (Table **4**) were analyzed using a 4 (Language: French-English vs. English-French vs. Spanish-English vs. English-Spanish) x 2 (French Gender: masculine vs. feminine) x 2 (Conceptual Kind: natural vs. artifact) x 2 (Condition: Label vs. no label) x 2 (Test Language: French vs. Spanish vs. English) mixed ANOVA with repeated measures on French Gender and Conceptual Kind.

As in the previous ANOVA there were significant main effects for Gender, $\underline{F}(1,120) = 19.64$, $\underline{p} < 0.0001$, and Concept, $\underline{p}(1,120) = 25.52$, $\underline{p} < 0.0001$. There were three reliable two-way interactions: French Gender and Language,

Table 4. **Mean Proportion of Times (and Standard Deviations) Bilinguals Assigned a Male Voice to Grammatically Masculine and Feminine Incongruent Items in French and Spanish, by French Grammatical Gender, Conceptual Kind, and Participants' Language**

	French Gender Items		Conceptual Kinds	
	Masculine	Feminine	Artifacts	Natural Kinds
Language				
French-English	0.62	0.49	0.58	0.54
	(0.18)	(0.17)	(0.19)	(0.14)
English-French	0.45	0.57	0.58	0.43
	(0.17)	(0.16)	(0.11)	(0.14)
Spanish-English	0.38	0.61	0.54	0.45
	(0.15)	(0.15)	(0.15)	(0.15)
English-Spanish	0.39	0.58	0.56	0.41
	(0.19)	(0.17)	(0.19)	(0.16)

$\underline{F}(3,120) = 11.54$, $\underline{p} < 0.0001$; Concept and Language, $F(3,120) = 2.36$, $p = 0.05$; as well as Concept and Gender, $\underline{F}(3,120) = 11.54$, $\underline{p} < 0.001$.

Dependent t-tests showed that French-English bilinguals assigned a male voice to masculine items reliably more often than to feminine items, dependent $\underline{t}(31) = 2.84$, $\underline{p} < 0.008$. In contrast, Spanish-English, English-Spanish, English-French bilinguals (dependent ts: $\underline{t}(31) = -4.76$, $\underline{p} < 0.001$; $\underline{t}(31) = -3.68$, $\underline{p} < 0.001$; $\underline{t}(31) = -2.57$, $\underline{p} < 0.01$ respectively) assigned a male voice reliably more often to feminine rather than masculine items. In other words, French-English bilinguals classified items according to French grammatical gender, whereas all other bilinguals classified items according to Spanish grammatical gender.

French-English bilinguals' classifications of conceptual kinds also differed from those of the other participants. Only French-English bilinguals assigned a male voice to artifacts and natural kinds equally often: dependent $\underline{t}(31) = 1.00$, $\underline{p} = 0.325$. Spanish-English, English-Spanish, English-French, bilinguals assigned a male voice reliably more often to artifacts than to natural kind items: $\underline{t}(31) = 2.32$, $\underline{p} < 0.01$; $\underline{t}(31) = 2.88$, $\underline{p} < 0.007$; $\underline{t}(31) = 4.86$, $\underline{p} < 0.001$; $\underline{t}(31) = 6.28$, $\underline{p} < 0.001$; respectively.

Dependent t-test of the means for the two-way interaction involving Conceptual Kind and Gender showed that participants assigned a male voice reliably more often to masculine ($\underline{M} = 0.66$, $\underline{SD} = 0.17$) compared to feminine ($\underline{M} = 0.53$, $\underline{SD} = 0.17$) artifact items: $\underline{t}(127) = 6.10$, $\underline{p} < 0.0001$. Moreover, participants also assigned a male voice significantly more often to masculine ($\underline{M} = 0.50$, $\underline{SD} = 0.19$) compared to feminine ($\underline{M} = 0.43$, $\underline{SD} = 0.19$) natural kinds items: $\underline{t}(159) = 2.97$, $\underline{p} = 0.004$.

Classifications According to Spanish Grammatical Gender

A third analysis compared Spanish-English bilinguals', English-Spanish bilinguals', and English monolinguals' classifications according to Spanish grammatical gender. These mean classifications (Table **5**) were examined using a 3 (Language: Spanish-English vs. English-Spanish vs. English) x 2 (Spanish Gender: masculine vs. feminine) x 2 (Conceptual Kind: natural vs. artifact) x 2 (Condition: Label vs. no label) x 2 (Test Language: French vs. Spanish vs. English) mixed ANOVA with repeated measures on French Gender and Conceptual Kind. There were main effects for Conceptual Kind, $\underline{F}(1,90) = 61.69$, $\underline{p} < .0001$, and Gender, $\underline{F}(1,90) = 122.21$, $\underline{p} < .0001$. However, Conceptual Kind reliably differed across Gender, $\underline{F}(1,90) = 7.63$, $\underline{p} < 0.007$. All speakers assigned a male voice to masculine artifacts ($\underline{M} = 0.71$, $\underline{SD} = 0.15$) and masculine natural kind items ($\underline{M} = 0.60$, $\underline{SD} = 0.21$) reliably more often than would be expected by chance alone (0.50): dependent $\underline{t}(95) = 13.72$, $\underline{p} < 0.0005$; and dependent $\underline{t}(95) = 4.66$, $\underline{p} < 0.0005$ respectively. Furthermore, all speakers reliably assigned a male voice to feminine natural kind items ($\underline{M} = 0.30$, $\underline{SD} = 0.17$) significantly less than would be expected by chance, dependent $\underline{t}(95) = -6.53$, $\underline{p} < 0.0005$. Speakers assigned a male voice to feminine artifacts ($\underline{M} = 0.49$, $\underline{SD} = 0.19$) at a rate which did not differ from chance, dependent $\underline{t}(95) = -0.52$, $\underline{p} > 0.05$.

The ANOVA revealed one further two-way interaction: Conceptual Kind and Language, $\underline{F}(4,150) = 2.48$, $\underline{p} < 0.05$. All speakers attributed a male voice to artifacts reliably more often than to natural kinds: all dependent $\underline{t}s(31) \geq 2.72$, $\underline{p} \leq 0.01$. Further analysis of the means shown in Table **5**, compared to what would be expected by chance (0.50), revealed that Spanish-English bilinguals failed to attribute a male voice to natural kind items ($\underline{M} = 0.48$, $\underline{SD} = 0.12$), less often than chance, $\underline{t}(31) = 0.94$, > 0.05. Thus, the source of the reliable Conceptual Kind and Language interaction was Spanish-English bilinguals who failed to attribute a female voice to natural kind items at a level beyond that expected by chance alone.

Spanish grammatical gender affected categorization by English monolinguals, as well as by Spanish-English and English-Spanish bilinguals. Because English speakers have no knowledge of Spanish grammatical gender, something other than Spanish grammatical gender may have accounted for this pattern of results. As in Sera *et al*. [16], we conducted a post hoc analysis to study the effect of four dichotomous factors other than Spanish gender empirically associated with masculinity and femininity [21, 22]: artifact vs. natural kind; angular or curved; typically used by males or females; and dense or not dense. We asked 23 adults to categorize each of the 32 test items according to the four dichotomous factors associated with masculine and feminine. Adult raters viewed all 32 items, one at a time, in 1 of 2 random orders. For each item, adults rated whether it was a

Table 5.   Mean Proportion of Times (and Standard Deviations) Spanish-English Bilinguals, English-Spanish Bilinguals, and English Monolinguals Assigned a Male Voice to Items by Spanish Grammatical Gender, Conceptual Kind, and Participants' Language

	Spanish Gender Items		Conceptual Kinds	
	Masculine	**Feminine**	**Artifacts**	**Natural Kinds**
Language				
Spanish-English	0.68 (0.14)	0.37 (0.14)	0.57 (0.10)	0.48 (0.12)
English-Spanish	0.65 (0.14)	0.40 (0.16)	0.60 (0.14)	0.44 (0.14)
English	0.64 (0.11)	0.43 (0.10)	0.64 (0.10)	0.43 (0.10)

natural kind or artifact, angular or curved, typically used by males or females, and dense or not dense.

Based on participants' ratings, items were classified as having from 0 to 4 masculine attributes. Two items were rated as having four masculine attributes (viz., artifact, angular, typically used by males, and dense). Three items were rated as having three masculine attributes. Nine items were rated as having two masculine attributes. Ten items were rated as having one masculine attribute, and eight items were rated as having no masculine attributes (viz., natural kind, curved, typically used by females, and not dense). For each group of items, the expected proportion of masculine classifications was obtained by dividing the number of grammatically masculine (in Spanish) items by the total number of items. For instance, the two items that were rated as having four masculine attributes were grammatically masculine in Spanish. So, the expected proportion of classifications for this group of items was 1. Three of the eight items that were rated as having no masculine attributes were grammatically masculine in Spanish. So the expected proportion of classifications for this group of items was 0.375. Table **6** shows the expected and obtained proportion of masculine classifications for groups of items, by number of masculine attributes (0 to 4), and by speakers' language. This analysis shows that the number of masculine attributes and Spanish Grammatical Gender are equally good predictors of all participants' gender classifications.

Voice Attributions to Control Items

All bilingual and monolingual participants overwhelmingly attributed a male voice to natural gender male control items (all $\underline{M}$s > 0.97, all $\underline{SD}$s < 0.07) and a female voice to natural gender female items (all $\underline{M}$s > 0.98, all $\underline{SD}$s < 0.07).

Performance on Grammatical Gender Identification Task and Classifications

Thus far, the analyses do not rule out the possibility that performance on the gender attribution task was determined solely by participants' knowledge of test items' grammatical gender. If knowledge of grammatical gender affected participants' performance on the gender attribution task, then one would expect that within each bilingual group (F-E, E-F, S-E, & E-S) participants who made the fewest grammatical gender errors would differ from participants who made the most grammatical gender errors. The weighted mean number of grammatical gender errors on the grammatical gender identification task was 0.41 (range = 0 – 2, median = 0.0) for F-E bilinguals; 3.81 (range 1 – 7, mdn = 3) for E-F bilinguals; 0.125 (range 0 – 1; mdn = 0) for S-E bilinguals; and 0.28 (range 0 – 2; mdn = 0) for E-S bilinguals.

For each bilingual group, participants were classified as below or above the median number of grammatical gender errors. Then for each bilingual group, participants' performance on the gender attribution task was compared with 4 separate 2 (Grammatical Gender Errors: below mdn vs. above mdn) x 2 (French (or Spanish) Gender: masculine vs. feminine) x 2 (Conceptual Kind: natural vs. artifact) mixed ANOVAs. None of the four ANOVA revealed any reliable effects. Therefore, knowledge of grammatical gender did not affect performance on the grammatical attributions tasks within each of the bilingual groups.

Age of L2 Acquisition and Classifications

It is possible that bilinguals who acquired proficiency in their second language before age 7 may have differed from those who acquired L2 proficiency after age 7 [see 17, 18]. To test this possibility, two separate ANOVAs were performed: one for F-E and E-F bilinguals combined and another for S-E and E-S bilinguals combined. Early L2 acquisition bilinguals acquired L2 proficiency before age 7; late L2 acquisition bilingual acquired L2 proficiency after age 7 (For each bilingual group a 2 (Age of Acquisition: early vs. late) x 2(Gender: masculine vs. feminine) x 2 (Conceptual Kind: natural vs. artifact) x (Condition: Label vs. no label) x 2(Test Language: French (or Spanish) vs. English) mixed ANOVAs. There was no main effect of Age of Acquisition, nor did Age of Acquisition reliably interact with any other factor. Consequently, the age of self-reported proficiency in their second language did not affect bilinguals' performance on the grammatical attribution task.

DISCUSSION

The purpose of this study was to determine how French and Spanish grammatical gender affects bilinguals' conceptual gender. The principal finding was that participants' gender classifications (viz., selection of a male or female voice) varied as a function of grammatical gender. As expected, French-English and Spanish-English bilinguals classified items according to French and Spanish grammatical gender respectively. Also as expected, the French grammatical gen-

Table 6. **Expected and Obtained Proportion of Masculine Classifications for Groups of Items by Number of Masculine Attributes (0 to 4), and Participants' Language**

	Predicted by Spanish Grammatical Gender of Items	French-English	English-French	Spanish-English	English-Spanish	English
Number of Masculine Factors (0–4)						
Four (2 items)	1.00	0.77	0.80	0.82	0.75	0.84
Three (3 items)	0.67	0.64	0.69	0.76	0.64	0.73
Two (9 items)	0.55	0.65	0.61	0.55	0.61	0.62
One (10 items)	0.50	0.52	0.54	0.55	0.58	0.53
0 (8 items)	0.375	0.38	0.31	0.29	0.26	0.28

der effect on classifications was even more pronounced for the 16 grammatically opposed items in French and Spanish: Only French-English bilinguals classified items according to French gender. These findings extend previous cross-linguistic research by using the same gender attribution task with different test items and bilingual rather than monolingual samples [6, 16]. Comparing bilinguals living in the same geographical location, who nonetheless differ in the order of L1 and L2 acquisition, is a more stringent test of language per se on classifications and better controls for non-language effects on classifications. Notably, English-French bilinguals' pattern of classifications mirrored that of English monolinguals for grammatically French items. Indeed, English-French bilinguals' classifications were indistinguishable from those of English monolinguals. Therefore, it seems that French grammatical gender affects categorization, but only for bilinguals who acquire French as their first language.

An alternative interpretation of the gender attribution task would be that it simply tested knowledge of grammatical gender. Hence bilinguals directly applied their grammatical gender knowledge to complete the task. We think that this explanation is unlikely for several reasons. First of all, if bilinguals relied on knowledge of grammatical gender to complete the gender attribution task, we would expect them to have attributed male and female voices to natural kind and artifact items equally often. This is because our design incorporated an equal number of grammatically masculine and feminine natural kind and artifact items. All of the bilinguals attributed a male voice to artifacts at a level reliably different from chance (50%), so they did not make gender attributions equally often. Secondly, if bilinguals relied on direct language cues of grammatical gender to complete the gender attribution task, we would expect to see a test language and label effect. In our study, half the bilingual participants were tested in L1; half the bilingual participants were tested in L2. Moreover, for half the bilingual participants, items were specifically labeled; for half the bilingual participants, items were not specifically labeled. Testing participants in French and Spanish as well as specifically naming test items provided multiple direct sources of grammatical gender that would have primed them to use grammatical gender to complete the gender attribution task. However, neither the language bilinguals were tested in, nor whether items were specifically labeled affected voice gender attributions.

Spanish Gender and the Artifacts *vs*. Natural Kinds Distinction

French gender affected classifications for only the most proficient French language speakers: French-English bilinguals. We expected a similar patter of results for S-E, E-S bilinguals. Unexpectedly, Spanish gender similarly affected classifications by Spanish-English and English-Spanish bilinguals, as well as English monolinguals, all of whom classified items according to Spanish grammatical gender.

As in Sera *et al*. [16], our post hoc analysis showed that four factors: conceptual kind, an item's appearance (angular or linear), whether it is typically used by males or females, and its density (dense or not dense) covary with Spanish grammatical gender. Consequently, we can not distinguish between the effects of Spanish grammatical gender and the

covariates. Speculatively, Spanish grammatical gender may be less arbitrary than many observers have argued [23]. Instead, Spanish gender may be largely based on observable characteristics such as those used in our post-hoc analysis.

Strength of Language Effect on Cognition

A final issue of interest is how strong an effect language has on non-linguistic gender attributions. We argue that the present study shows a meaningful effect of language on cognition—what Fishman [14] called a level four linguistic structure and behavioral concomitants relationship. Examples of superficial language effects on cognition abound such as this poetic assertion that "French has no word for home, and we have no word for strict pleasure" [24, p.5]. French and English speakers may use circumlocutions to express home and strict pleasure, but they clearly have no trouble understanding the underlying concepts. Pinker [12] rightly argues that specialized lexical terms such as schadenfreude, or pleasure at another's misfortune, are not evidence that Germans think differently than speakers of other languages. Indeed, English has adopted innumerable expressions from other languages and from English language innovators to talk about emotions and other concepts that are instantly recognizable, but not yet lexicalized by English speakers [12].

But languages do differ in more fundamental ways such as in how they categorize grammatical gender. Spanish and French gender systems appear only superficially similar. Both languages have two genders or kinds: masculine and feminine. Both gender systems are loaded, because nouns, adjectives, and articles provide multiple sources of morphological gender information. But the gender of items does not necessarily overlap; some words are the opposite genders across the two languages. And, tellingly, only French-English bilinguals, native French speakers who subsequently acquired English as a second language, classified items according to French grammatical gender. French-English bilinguals also demonstrated the artifact-male, natural kinds-female distinction. So it is not the case that French gender prevents French speakers from categorizing the world like speakers of other languages. Rather, French gender affects non-linguistic gender categorization, but the effect does not completely override non-linguistic ways of categorizing experiences.

Implications for Second Language Learners and Future Directions

The language effect on non-linguistic gender attributions may be viewed as a default assumption of the sort that arises in everyday decision making which leads to systematic biases. Tversky and Kahneman [25] originally identified three heuristics- anchoring, availability, representativeness- and the biases associated with each. For instance, the availability heuristic refers to a person's tendency to estimate the likelihood of an event based on how readily examples come to mind, rather than the true probability of an event. When asked "Which is a more likely cause of death in the United States: being killed by falling airplane parts or being killed by a shark?", most people assume death by shark attack is the more likely event. However, the probability of being killed by falling airplane parts is 30 times more likely than being involved in a fatal shark attack in the United States.

But since fatal shark attacks are more widely publicized, people systematically overestimate the actual probability of a fatal shark attack [26]. Similarly, grammatical gender may function as an unconscious heuristic for bilinguals, when asked to make non-linguistic gender attributions. In the absence of motivation or training to do otherwise, bilinguals' default assumptions about gender attributions are reliably biased by the grammatical gender of the first language they learned to speak.

Indeed, Ardila [27] argues that second language learners must be taught the cultural and social components of the second language in order to overcome the "linguistic interference" of their native language. Bilingual children have superior selective attention and better inhibitory control on a card sort task for ignoring perceptual information than monolinguals [28]. It would be interesting to see whether adult bilinguals who were trained to overcome the default assumptions of their native language, would generalize this ability to other non-linguistic decision making tasks. The growing list of advantages to bilingualism may include improved ability to inhibit unwanted automatic responses, increased ability to switch attention, and perhaps improved control over changing from automatic to conscious thinking, all of which would decrease susceptibility to cognitive biases.

ACKNOWLEDGMENTS

REFERENCES

[1]   Brown R, Lenneberg E. A study in language and cognition. J Abnorm Soc Psychol 1954; 49: 454-62.

[2]   Lucy JA. The scope of linguistic relativity: An analysis and review of empirical research. In Gumperez JJ, Levinson SC, Eds. Rethinking linguistic relativity. Cambridge University Press 1996; 37-69.

[3]   Kay P, Kempton W. What is the Sapir-Whorf hypothesis? Am Anthropol 1984; 86: 65-79.

[4]   Olson D. From utterance to text: The bias of language in speech and writing. Harv Educ Rev 1977; 47: 257-81.

[5]   Segalowitz N, Poulin-Dubois D. The structure of categories: Typicality gradients, linguistic familiarity, and crosslinguistic comparisons. Eur Bull Cogn Psychol 1990; 10: 491-512.

[6]   Sera MD, Berge D, Castillo J. Grammatical and conceptual forces in the attribution of gender by English and Spanish Speakers. Cogn Dev 1994; 9: 261-92.

[7]   Stefflre V, Morley L, Castillo-Vales V. Language and cognition in Yucatan: A cross-cultural replication. J Pers Soc Psychol 1966; 4: 112-5.

[8]   Karmiloff-Smith A. A functional approach to child language: A study of determiners and reference. Cambridge University Press 1979.

[9]   Martinez I, Shatz M. Linguistic influences on categorization in preschool children: a crosslinguistic study. J Child Lang 1996; 23: 529-45.

[10]   Perez-Pereira M. The acquisition of gender: What Spanish children tell us. J Child Lang 1991; 18: 571-90.

[11]   Pinker S. The language instinct. New York: Harper Perennial 1995.

[12]   Pinker S. How the mind works. New York: W.W. Norton & Company 1997.

[13]   Takano Y. Methodological problems in cross-cultural studies of linguistic relativity. Cognition 1989; 31(2): 141-62.

[14]   Fishman JA. Systematization of the Whorfian hypothesis. Behav Sci 1960; 5: 323-39.

[15]   Fishman J. The Whorfian hypothesis: Variations of valuation, confirmation and disconfirmation. Int J Sociol Lang 1980; 26: 25-40.

[16]   Sera MD, Elieff C, Forbes J, Clark-Burch M, Rodriguez W, Poulin-Dubois D. When language affects cognition and when it does not: Anal Grammatic Gender Classification 2002; 131(3): 377-97.

[17]   Johnson JS, Newport EL. Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. Cogn Psychol 1989; 21(1): 60-99.

[18]   Johnson JS, Newport EL. Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. Cognition 1991; 39(3): 215-58.

[19]   Abatte M, LaChapelle N. Pictures, please! A language supplement. Tucson, AZ: Communication Skill Builders, Inc. 1984.

[20]   Dunn LM, Dunn LM. Peabody Picture Vocabulary Test—Revised. Circle Pines (MN): Am Guid Serv 1981.

[21]   Leinbach MD, Hort BE, Fagot B. Bears are for boys. Metaphorical associations in young children's gender stereotypes. Cogn Dev 1997; 12(1): 107-30.

[22]   Mullen MK. Children's classifications of nature and artifact pictures into female and male categories. Sex Roles 1990; 23: 577-87.

[23]   Comrie B. Grammatical gender systems: A linguist's assessment. J Psycholinguist Res 1999; 28(5): 457-66.

[24]   Gilbert J. The great fires. New York: Alfred A. Knopf 1994.

[25]   Tversky A, Kahneman D. Judgment under uncertainty: Heuristics and biases. Science 1974; 185(4157): 1124-31.

[26]   Plous S. The psychology of judgment and decision making. Mcgraw-Hill 1993.

[27]   Ardila JAG. (2000Linguistic relativity in foreing language teaching: A didactic proposal. Didactica (Lengua y Literatura) 2000; 12: 43-53.

[28]   Bialystok E, Martin MM. Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. Dev Sci 2004; 7 (3): 325-39.