Journal of Experimental Psychology: General
2002, Vol. 131, No. 3, 377–397

0096-3445/02/$5.00 DOI: 10.1037//0096-3445.131.3.377

When Language Affects Cognition and When It Does Not: An Analysis of Grammatical Gender and Classification

Maria D. Sera and Chryle Elieff
University of Minnesota

James Forbes
Angelo State University

Melissa Clark Burch
University of Minnesota

Wanda Rodríguez
University of Puerto Rico

Diane Poulin Dubois
Concordia University

The focus of this work was on the relation between grammatical gender and categorization. In one set of studies, monolingual English-, Spanish-, French-, and German-speaking children and adults assigned male and female voices to inanimate objects. Results from Spanish and French speakers indicated effects of grammatical gender on classification; results from German speakers did not. A connectionist model simulated the contradicting findings. The connectionist networks were also used to investigate which aspect of grammatical gender was responsible for the different pattern of findings. The predictions from the connectionist simulations were supported by the results from an artificial language-learning task. The results from this work demonstrate how connectionist networks can be used to identify the differences between languages that affect categorization.

Many studies suggest that language affects cognition (e.g., Brown & Lenneberg, 1954; Carroll & Casagrande, 1958; Davidoff, Davies, & Robeson, 1999; Ervin, 1962; Gelman & Gallistel, 1978; Levinson, 1996; Lucy, 1992; Luria, 1961; Martinez & Shatz, 1996; Miller, Smith, Zhu, & Zhang, 1995; Sera, Bales, & del Castillo Pintado, 1997; Sera, Berge, & del Castillo Pintado, 1994; Sera, Reittinger, & del Castillo Pintado, 1991; Sapir, 1912; Whorf, 1956). Just as many other studies argue against such effects (e.g., Bellugi, Bihrle, & Corina, 1991; Fodor, 1983; Furth, 1966; Heider, 1972; Heider & Oliver, 1972; Karmiloff-Smith, 1979; Pérez-Pereira, 1991; Soja, Carey, & Spelke, 1991). For example, Davidoff et al. (1999) reported that the number of basic color terms in a language affects categorization; Heider and Oliver (1972) reported no language effects. Levinson (1996) found that differences across languages in locative expressions affect the representation of spatial relations; Munnich and Landau (1997) found no language effects. Sera et al. (1994) and Martinez and Shatz (1996) argued that grammatical gender systems influence classification; Karmiloff-Smith (1979) and Pérez-Pereira (1991) argued that grammatical and perceptual gender are independent. So a critical issue in the language–thought literature is to understand why language sometimes affects cognition and other times does not.

We address this issue by focusing on the role of grammatical gender in classification. The broader goal of the work is to offer an approach for understanding differences across languages and how those differences might influence other aspects of cognition. Few studies have systematically looked at a particular linguistic system across different languages and how speakers of those languages perform on a related classification task. By examining the effects of grammatical gender on classification across several languages and by modeling these effects, we were able to identify a property of gender systems that leads to the generalization of male and female attributes to inanimate objects. In the remainder of this introduction, we describe the grammatical gender systems under investigation. Then we review past work on the relation between

Maria D. Sera, Chryle Elieff, and Melissa Clark Burch, Institute of Child Development, University of Minnesota; James Forbes, Psychology and Sociology Department, Angelo State University; Wanda Rodríguez, Department of Psychology, University of Puerto Rico; Diane Poulin Dubois, Department of Psychology, Concordia University, Montreal, Quebec, Canada.

This research was supported by a McKnight-Landgrant Professorship and a University of Minnesota Grant-in-Aid to Maria D. Sera. The research was also supported by a grant from the Natural Sciences and Engineering Research Council of Canada to Diane Poulin Dubois. This work was conducted while Chryle Elieff and Melissa Clark Burch held traineeships from the National Institute of Child Health and Human Development. Study 2 was conducted to fulfill the requirements for an honors degree from Indiana University by Chryle Elieff. Portions of this work were presented at the meetings of the Psychonomic Society in Philadelphia in November 1997 and the Society for Research in Child Development in Albuquerque in April 1999.

We thank Milissa Tilton, Christa Niven, Nair Irizarry, Isabelle Marleau, Nicole Atwood, and Annie Ryman for their assistance with data collection. We also thank the children, parents, and staff of Hamline Elementary School, Hamburg Public Schools, Happy Moments Daycare Center, Pensionnat Jèsus-Marie, Ecole Sainte-Marcelline, Ecole Ste-Anne, and the University of Puerto Rico Middle School. Finally, we thank Michael Maratsos, Linda Smith, and Kim Plunkett for their comments on the work.

Correspondence concerning this article should be addressed to Maria D. Sera, Institute of Child Development, University of Minnesota, 51 East River Road, Minneapolis, Minnesota 55455. E-mail: sera@umn.edu

grammatical gender and classification. We conclude the introduction with a review of connectionist models.

The Gender Systems of English, Spanish, German, and French

Natural gender is an attribute of all animals in the world and a salient property for organizing information. Languages differ in how they code natural gender. English codes natural gender primarily through lexical items (e.g., girl–boy, brother–sister, and aunt–uncle) and through some pronouns (e.g., she–he and him–her). English does not assign a gender to all nouns that refer to animates (e.g., doctor) or to nouns that refer to inanimates (e.g., apple). Languages such as Spanish, however, have a grammatical gender system because they mark gender with morphological information that is carried by pronouns, determiners, nouns, and adjectives. One way linguistic gender systems differ from each other is by the number of grammatical categories that carry gender information. For example, whereas Spanish pronouns, determiners, nouns, and adjectives all frequently carry gender markings, only German pronouns and determiners are marked for gender consistently. Languages that mark gender morphologically across several grammatical categories are often described as being "gender loaded." Another difference across languages with grammatical gender systems is the number of grammatical "gender" categories that they have. Some languages have no grammatical gender categories (e.g., English), some have 2 (e.g., Spanish and French), some have 3 (e.g., German and Greek), and some have nearly 20 (e.g., Thai and Sesotho). Of course, in a language with 20 noun classes, the degree to which the noun-based agreement rules actually mark gender is questionable.

The Spanish grammatical gender system works as follows. All Spanish nouns have a grammatical gender. Moreover, the Spanish words for boy, brother, and male cat—*niño, hermano,* and *gato*—all have the same stem as the Spanish words for girl, sister, and female cat: *niña, hermana,* and *gata.* They differ in that the Spanish words that refer to males end in /o/ and the Spanish words that refer to females end in /a/. Therefore, many Spanish words that refer to the same category of people (and animals) often differ only in their morphological ending. For words that refer to animals and people, then, the Spanish morphological endings are a good, albeit imperfect, cue to natural gender. Nouns that do not refer to animals or people share the same morphological endings as nouns that mark the natural gender of males and females. Unlike many of the nouns that refer to animals or people, however, there is only one noun form for words that refer to inanimate objects. For example, the Spanish word for "table," *mesa,* is feminine, and the Spanish word for "telescope," *telescopio,* is masculine. Other components of Spanish noun phrases are also marked for gender. Singular and plural definite and indefinite articles each have masculine and feminine forms. So one would say *una* niña when referring to "a girl" and *un* niño when referring to "a boy." Many Spanish adjectives also carry information about grammatical gender. Thus, the Spanish translations of "tall" when saying "a tall girl" and "a tall boy," *una niña alta* and *un niño alto,* are also marked for gender. The Spanish language, then, is a language that is loaded with grammatical gender markings because the articles, nouns, and adjectives typically carry morphological gender information (see Terrell & Salgués de Cargill, 1979, for a more detailed outline of the Spanish gender system).

German, like Spanish, possesses a grammatical gender system. However, the German system differs from the Spanish system in three ways. First, instead of having just two gender categories, the German has a third, neuter, category. For example, the German translations of "the man," "the woman," and "the girl" are *der Mann* (masculine) *die Frau* (feminine), and *das Mädchen* (neuter), respectively. It should be clear that the relation between grammatical gender and natural gender in German is imperfect, because different articles are used for "the woman" and "the girl." Moreover, the German indefinite article only sometimes carries grammatical gender information. For example, among *ein Mann, eine Frau,* and *ein Mädchen,* only *eine Frau* is distinguished from the others by its indefinite article. Another difference between the Spanish and German gender systems is that German determiners mark case as well as gender. Consider the German translations of "the man" in the sentences "The man scratched the cat" (*Der Mann kratzt die Katze*) and "The cat scratched the man" (*Die Katze kratzt den Mann*). In these sentences, the determiner of "man" changes as the man's thematic role changes. So the German determiners carry information about case and gender simultaneously.

A third difference between the Spanish and German gender systems is that not as many grammatical categories carry gender information in German. German determiners (e.g., definite and indefinite articles) carry information about grammatical gender most frequently and consistently. Nouns are rarely marked for gender. Adjectives are marked for grammatical gender in some cases. For example, the German translation of "tall," when used to describe the tall man (*Der große Mann*), the tall woman (*die große Frau*), and the tall girl (*das große Mädchen*), is the same across all three gender categories (see Fox, 1992, for more information about German gender). We thus asked whether grammatical gender effects would be observed among speakers of German, a language whose gender system varies in several ways from Spanish.

The French and Spanish systems are similar to each other in three ways. First, both French and Spanish possess two gender categories. Second, neither Spanish nor French is a case-based system. The French and Spanish translations of "the girl" are the same in sentences such as "The girl hit the boy" and "The boy hit the girl." Third, articles, nouns, and adjectives in French and Spanish carry morphological gender information, so both languages are gender loaded. The French and Spanish systems differ from each other in that some nouns that are classified in one language according to one gender are classified in the opposite way by the other language (e.g., *bed* is feminine in Spanish but masculine in French). This difference between the Spanish and French gender systems offers a naturally occurring "manipulation" of grammatical gender. So we asked whether Spanish and French speakers agree on the classification of objects that are assigned the same genders by French and Spanish and disagree on the classification of items that are assigned different genders.

Grammatical Gender and Classification

Numerous studies have examined the relationship between grammatical gender and classification. The evidence from some studies indicates effects of grammatical gender on classification; other studies fail to show such effects. In a classic study by Ervin

(1962) on the role of grammatical gender and classification, native speakers of Italian were taught nonsense words that possessed either masculine or feminine Italian affixes. The Italian speakers rated the nonsense words with masculine affixes as more like men and the words with feminine affixes as more like women. Similar findings have been reported in speakers of Arabic (Clarke, Losoff, McCracken, & Rood, 1984; Clarke, Losoff, McCracken, & Still, 1981). However, the results from such studies were criticized because participants were asked to judge words. Thus, the tasks were argued to be directly testing speakers' knowledge of grammatical gender. If so, then their judgments reveal nothing about how grammatical gender affects categorization more generally. This criticism, however, does not address why such results are not uniform. For example, speakers of Finnish, who presumably also know the Finnish grammatical gender system, do not rate words that refer to inanimate objects according to grammatical gender (Clarke et al., 1984; Clarke et al., 1981). Moreover, in a more recent study in which speakers of Spanish were asked to assign either a man's or a woman's voice to pictured objects that were either labeled or not labeled, effects of grammatical gender were observed even when the pictures were not labeled (Sera et al., 1994). These findings suggest that use of the words themselves matters little in these tasks. Indeed, Martinez and Shatz (1996) reported that Spanish-speaking children are more likely than English-speaking children to use grammatical gender as a basis for sorting pictures of inanimate objects in a free classification task. Effects of grammatical gender are not limited to judgments of inanimate objects. Guiora, Beit-Halachmi, Fried, and Yoder (1983) found that children who speak Hebrew, a language with a grammatical gender system, were more likely than children who speak English to correctly label their own gender at 25 months of age. Thus, effects of grammatical gender on classification have been reported across a variety of languages and over a broad range of tasks.

As stated earlier, however, not all studies that have examined the relation between grammatical gender and classification have revealed an interaction between language and categorization. In a task in which participants were asked to rate the masculine–feminine quality of entities denoted by a set of words, Clarke et al. (1981) reported effects of grammatical gender for speakers of Arabic but not for speakers of Finnish. In addition, apparently different findings have emerged from researchers using different methodologies. For example, the results from speakers of Spanish reported by Sera et al. (1994), which show effects of grammatical gender on classification, contrast with those from Pérez-Pereira (1991), which suggest that grammatical and perceptual classification are independent. Pérez-Pereira (1991) examined the relation between grammatical and perceptual gender in Spanish-speaking children following the procedures that Karmiloff-Smith (1979) used to test French-speaking children. In these studies, children were shown pictures of novel animate figures (some pictures depicted males and others depicted females) and were told "This is a picture of a *pifar* (an invented word)." Thus, the referent was assigned both a grammatical gender that was coded by the determiner and noun and a perceptual gender that was illustrated in the picture. Sometimes the grammatical gender matched the perceptual gender, and other times the grammatical and perceptual gender information contradicted. A second picture that illustrated the original figure of a different color followed the presentation of the originally labeled picture. The children were then asked to state the color of the second "pifar." Because the gender of adjectives in Spanish and French agrees with the gender of nouns, the adjective provided by the children indexed the children's classification of the gender of the figure. Pérez-Pereira (1991) and Karmiloff-Smith (1979) found that children generated the adjectives that matched the grammatical information, even when the perceptual information pointed to an opposite gender. Because children ignored the perceptual gender of the figures when generating adjectival endings, these findings suggest that language and perceptual processing of gender are independent. To summarize, the relation between grammatical gender and classification suggested by the literature varies as a function of language and the task that is used to examine the relation.

Connectionism

Connectionism, parallel-distributed processing systems, or neural networks are computational models that can be used to simulate ways that cognitive phenomena might arise. The models are based on the interaction among large numbers of neuronlike units called nodes. The nodes are limited in their representational capabilities—they represent only activations between 0 and 1—so they simulate cognitive phenomena without explicitly representing categories. Within these frameworks, learning involves altering the connections between the nodes, which regulates the flow of activation across the network.

A very important property of these models is that they can be used to simulate conditions that are rare or very difficult to find in nature. They can be used to describe ideal circumstances, or they can be used to parcel out confounding variables to help identify the cause of a particular effect. In addition, by carefully designing sets of simulations, researchers can pit alternative hypotheses against one another. For example, Elman (1991) demonstrated that limitations in memory might actually help children acquire grammar. The insights from Elman's (1991) simulations substantially change our understanding of "critical periods" in language learning. The cross-linguistic study of language–thought relations is another area that may benefit from the use of connectionist simulations.

Goals of the Current Work

The goals of this research were threefold. One was to determine whether the cross-linguistic variability of the gender systems of French, Spanish, English, and German leads to systematic differences in classification by speakers of these languages. We examined the attribution of male and female voices to pictured objects by English, Spanish, and French speakers in Study 1. We included speakers of English as a comparison group and as speakers of a language that possesses few grammatical gender markers. We examined the performance of English and German speakers in the same task in Study 2. In Study 3, we used a connectionist model to simulate our findings as well as findings such as those of Karmiloff-Smith (1979) and Pérez-Pereira (1991). The simulations also identified a linguistic property responsible for the overgeneralization of male- and female-like properties to inanimate objects (and lack thereof), and Study 4 tested the predictions from the simulations.

The second goal of this work was to further investigate the degree to which the assignment of grammatical gender to inanimate objects in languages such as Spanish is arbitrary. Although researchers have often argued that grammatical gender assignments are arbitrary, Sera et al. (1994) found English speakers' attribution of male and female voices to inanimate objects to be reliably consistent with the assignments made in the Spanish language. These results suggested that Spanish grammatical gender assignments were correlated with features of objects associated with males and females by speakers of English. In the current work, we asked whether the Spanish grammatical gender assignments predict the classification of objects by speakers of English, French, and German. If so, it would indicate that these linguistic constructions reflect shared ideas about the properties of objects associated with males and females and that they are not as arbitrary as previously thought. The final goal of this work was to examine the onset of grammatical gender effects, if any such effects exist. Finding that grammatical gender effects emerge at the same age across speakers of languages with different grammatical gender systems would suggest that similar underlying aspects of language are exerting an influence on categorization.

Study 1

If grammatical gender influences categorization, its influence should vary systematically according to differences across languages. French and Spanish are two languages with structurally similar gender systems but lexically different gender assignments. We thus examined male- and female-like classification of items by French and Spanish speakers as a function of the items' agreement on grammatical gender across the two languages. We predicted that speakers of French and Spanish will agree on the classification of items when the items' grammatical genders agree and disagree in their judgments when the languages disagree.

Another goal of Study 1 was to further investigate the apparent correlation between Spanish grammatical gender and English speakers' assignment of male and female voices to inanimate objects reported by Sera et al. (1994). In that study, most of the Spanish grammatically masculine items were artifacts, and the feminine items were natural kinds. The correlation between Spanish gender and classification could simply have reflected the tendency to view artifacts as more male-like than natural kinds (as reported by Mullen, 1990). We counterbalanced the number of grammatically feminine and masculine artifacts and natural kinds in the current study. Finding a correlation between English speakers' voice assignments and Spanish gender in the current study would suggest additional characteristics of objects that are associated with males and females and correlated with Spanish gender.

The final goal of the current study was to extend the previously reported evidence on the influence of grammatical gender on classification among Spanish speakers (e.g., Martinez & Shatz, 1996; Sera et al., 1994) to French speakers, speakers of a language with a similar grammatical gender system. We also asked whether speakers of French were influenced by grammatical gender at the same age as has been reported in Spanish speakers, approximately 7 years (Sera et al., 1994). Similar developmental trajectories would suggest that similar underlying mechanisms (i.e., certain aspects of grammar) influence classification in the two language groups.

Method

Participants. Monolingual speakers of English ($n = 128$), Spanish ($n = 128$), and French ($n = 128$) in four age and grade groups participated: kindergartners, second graders, fourth graders, and adults. The ages and grades of the children in each language group are listed in Table 1. Thirty-two people participated in each age group. Equal numbers of males and females participated in each age and language group. Spanish-speaking adults were recruited from the University of Puerto Rico. French-speaking adults were recruited through the University of Québec at Trois Rivières Québec. English-speaking adults were recruited through the University of Minnesota. Children who participated were recruited through preschools and elementary schools located near the universities just mentioned. All participants were native monolingual speakers of their respective languages.

Materials. A series of 40 colored, laminated drawings, obtained primarily from the English and Spanish versions of the Peabody Picture Vocabulary Test (Dunn & Dunn, 1981, 1986), were presented to all participants. Several drawings required slight alterations. All pictures were selected according to their grammatical gender in the Spanish and French languages. The objects were divided into two sets of items: 8 control items and 32 test items. The control items were used to ensure that participants understood the task. The grammatical gender of the control items was grammatically consistent with natural gender in both Spanish and French. Four of the control items depicted a male, and 4 items depicted a female. Test items were pictures of familiar objects, and these were items were used to assess the impact of the grammatical gender systems on participants' classifications. Eight test items were feminine in both Spanish and French, and 8 were masculine in both languages. In addition, 16 test items with opposite grammatical genders in the two languages were used: 8 were feminine in Spanish and masculine in French, and 8 were masculine in Spanish and feminine in French. In each group of test items, half depicted artificial objects (e.g., plane and book) and half depicted naturally occurring objects (e.g., corn and star). Table 2 lists the items that were used.

Table 1
Ages and Grades of Children Who Participated in Study 1

	English speakers		French speakers		Spanish speakers	
Grade	Range	*M*	Range	*M*	Range	*M*
Kindergarten	5 y 7 mo–6 y 7 mo	6 y 0 mo	5 y 3 mo–6 y 4 mo	5 y 9 mo	5 y 4 mo–6 y 11 mo	6 y 1 mo
Second grade	7 y 3 mo–9 y 1 mo	8 y 2 mo	7 y 7 mo–8 y 11 mo	8 y 5 mo	7 y 4 mo–8 y 11 mo	8 y 2 mo
Fourth grade	9 y 3 mo–10 y 11 mo	10 y 2 mo	10 y 2 mo–11 y 5 mo	10 y 7 mo	9 y 4 mo–10 y 4 mo	10 y 0 mo

Note. y = years; mo = months.

Table 2
Test and Control Items Used in Study 1

English (E)	French (F)	Spanish (S)
	Masculine controls	
a king	un roi	un rey
a man	un homme	un hombre
a giant	un géant	un gigante
a boy	un garçon	un niño
	Feminine controls	
a ballerina	une ballerine	una bailarina
a woman	une femme	una mujer
a bride	une mariée	una novia
a girl	une fille	una niña
	Feminine test items: artificial	
a table	une table	una mesa
a spoon	une cuillère	una cuchara
an arrow	une flèche	una flecha
a guitar	une guitar	una guitarra
	Feminine test items: natural	
a spider	une araignée	una araña
an apple	une pomme	una manzana
a star	une étoile	una estrella
a feather	une plume	una pluma
	Masculine test items: artificial	
a plane	un avion	un avión
a book	un livre	un libro
a telescope	un télescope	un telescopio
a knife	un couteau	un cuchillo
	Masculine test items: natural	
a tree	un arbre	un árbol
a corn	un blè d'Inde	un maíz
a sun	un soleil	un sol
a fire	un feu	un fuego
	Feminine F vs. masculine S test items: artificial	
a fork	une fourchette	un tenedor
a plate	une assiette	un plato
a car	une voiture	un carro
a screw	une vis	un tornillo
	Feminine F vs. masculine S test items: natural	
a tomato	une tomate	un tomate
a peanut	une cacahouète	un maní
a bat	une chauve-souris	un murciélago
a banana	une banane	un plátano
	Masculine F vs. feminine S test items: artificial	
a bed	un lit	una cama
a kite	un cerf-volant	una cometa
a broom	un balai	una escoba
a canoe	un canoë	una canoa
	Masculine F vs. feminine S test items: natural	
a grape	un raisin	una uva
a cloud	un nuage	una nube
a zebra	un zèbre	una cebra
a butterfly	un papillon	una mariposa

Procedure and design. Participants were told that we were going to make a movie in which items came to life, that they were going to be shown a series of pictures, and that they were to assign either a man's voice or a woman's voice to the object that was pictured. Participants were tested in one of two conditions. Half of the participants were in a *pictures plus labels* condition in which they heard the name of the object in addition to seeing the picture. Half of the participants were in a *pictures only* condition in which they were just shown the picture. The English instructions were as follows (French, German, and Spanish translations appear in the Appendix):

> We are thinking about making a new movie in which some everyday objects come to life and sing and dance. I am going to show you a series of pictures of these objects and want you to tell me whether you think each pictured object should have a man's voice or a woman's voice. Okay, here is picture number one [the experimenter would then show one picture to the participant, and for half of the participants the experimenter labeled the picture with its article], should this have a woman's voice or a man's voice in the movie?

Each participant was tested by an experimenter who was fluent in the participant's native language. The researcher recorded the participant's responses. Four randomly determined orders of stimulus presentation were used. The same number of participants viewed each order across conditions. Participants were tested individually. Testing lasted approximately 15 min.

Results

We first examined performance on control items, which had a natural gender. The data from 3 Spanish-speaking 5-year-olds were excluded from further analysis because these children did not classify 50% or more of the control items correctly and thus did not seem to understand the task. The remaining participants classified control items correctly between 96.5% and 100% of the time. Table 3 shows the mean percentages of times participants in each age and language group classified control items correctly. The high percentages indicate that participants understood the task.

The dependent measure used in statistical analyses was the percentage of female voice assignments. Our first analysis examined whether French and Spanish speakers systematically differed in their classifications of items that had different or the same grammatical genders across the two languages. This consisted of a Language (French or Spanish) × Grade (kindergartners, second graders, fourth graders, or adults) × Spanish Gender (masculine or feminine) × French Gender (masculine or feminine) × Conceptual Kind (artificial or natural) analysis of variance (ANOVA). Language and grade level were between-subjects factors, and Spanish gender, French gender, and conceptual kind were within-subject factors. We observed reliable main effects of French gen-

Table 3
Percentages of Times Control Items That Had a Natural Gender Were Classified Correctly by Speakers of Each Language Group as a Function of Age: Study 1

Language group	Kindergartners	Second graders	Fourth graders	Adults
English	99.6	100.0	100.0	99.2
French	99.2	98.0	98.8	97.8
Spanish	97.0	96.5	100.0	100.0

der, $F(1, 244) = 73.48$, $p < .0001$; Spanish gender, $F(1, 244) = 128.42$, $p < .0001$; and conceptual kind, $F(1, 244) = 76.34$, $p < .0001$. We observed reliable two-way interactions between language and French gender, $F(1, 244) = 70.98$, $p < .0001$; between French gender and grade, $F(3, 244) = 6.07$, $p < .0005$; between Spanish gender and language, $F(1, 244) = 41.42$, $p < .0001$; between Spanish gender and grade, $F(3, 244) = 12.64$, $p < .0001$; between conceptual kind and grade, $F(3, 244) = 5.30$, $p < .002$; between Spanish gender and conceptual kind, $F(1, 244) = 9.55$, $p < .002$; and between Spanish and French gender, $F(1, 244) = 17.61$, $p < .0001$. We observed reliable three-way interactions among French gender, language, and grade, $F(3, 244) = 3.31$, $p < .02$; among Spanish gender, language, and grade, $F(3, 244) = 5.78$, $p < .001$; and among French gender, Spanish gender, and conceptual kind, $F(1, 244) = 6.68$, $p < .01$. No other effects or interactions reached significance.

The main question we asked in this study, whether gender classifications varied systematically according to the assignment of grammatical gender across two languages, was addressed by the two-way interaction between French and Spanish gender. This interaction revealed reliable differences in judgments of items that were grammatically masculine in both languages versus items that were feminine in both languages (i.e., when the items agreed in the two languages) but not reliable differences when the grammatical genders of the items disagreed (Tukey's honestly significant difference [HSD] = 3.70, $p < .05$). Thus, the classifications of Spanish and French speakers varied systematically as grammatical gender varied across French and Spanish.

These results are more dramatic in the context of the two 3-way interactions involving French gender, Spanish gender, language, and grade. The means making up these interactions appear in Table 4. The interaction among French gender, language, and grade showed an effect of French gender among French speakers who were in second grade and older (Tukey's HSD = 16.39, $p < .05$). The interaction among Spanish gender, language, and grade showed an effect of Spanish gender among Spanish speakers who were in second grade and older (Tukey's HSD = 13.73, $p < .05$). No effects of French gender were found among Spanish speakers, and no reliable effects of Spanish gender were found among French speakers. These findings indicate that the grammatical gender systems of languages such as French and Spanish exert an effect on classification, starting in second grade.

We interpreted the two-way interaction between Spanish gender and conceptual kind in the context of the three-way interaction among Spanish gender, French gender, and conceptual kind. Results from post hoc analyses (Tukey's HSD = 3.51, $p < .05$) showed that artificial objects are not as susceptible to grammatical gender effects as naturally occurring ones. For natural kinds, items that were feminine in both languages were given female voices more often than items that had different genders across the languages, which were given more voices than items that were masculine in both languages. However, for artifacts, only the items that were feminine in both languages and the items that were masculine in both languages were reliably different. It is not clear why grammatical gender effects are more pronounced within naturally occurring objects than within artificial ones.

To further examine the effects of grammatical gender, we analyzed such effects in each individual stimulus item. Because the effects of grammatical gender were most pronounced among the adults, we compared the number of Spanish- and French-speaking adults who classified each item as masculine or feminine with chi-square tests. Table 5 shows the results of these tests. Of the 16 items with the same genders in Spanish and French, no reliable differences were observed between French and Spanish speakers on 13 of them. There were reliable differences in the classification of *fire, tree,* and *corn.* Both language groups tended to classify *fire* as male, but Spanish speakers did so more often than French speakers (94% vs. 75%). Both language groups also tended to classify *tree* as male, but here that tendency was weaker among Spanish than French speakers (67% vs. 84%). Thus, the only item that was classified differently by the two groups was *corn,* an item that several Spanish speakers spontaneously labeled as *una mazorca* (*cob*), which is feminine in Spanish. We found reliable differences between French and Spanish speakers on 15 of the 16 items for which Spanish and French genders disagree. The only item for which a reliable difference was not observed was *peanut,* which is feminine in French and masculine in Spanish but was classified as male-like by 22 French speakers as well as 20 Spanish speakers. This finding is important, because it is possible to explain the agreement between French and Spanish speakers as due to reliance on perceptual and conceptual features that are correlated with French and Spanish genders. However, it is difficult to explain the *disagreement* between French and Spanish by factors other than grammatical gender. Taken together with the results of the ANOVA, the results from the individual items offer strong support of an effect of grammatical gender on classification.

Our next set of analyses compared the classifications of English and Spanish speakers. Recall that previous work (Sera et al., 1994)

Table 4
Mean Percentages of Female Voice Assignments Made by French- and Spanish-Speaking Children and Adults for Masculine and Feminine Items in Spanish and French: Study 1

	French speakers–French grammatical gender				Spanish speakers–French grammatical gender				French speakers–Spanish grammatical gender				Spanish speakers–Spanish grammatical gender			
	Masculine		Feminine		Masculine		Feminine		Masculine		Feminine		Masculine		Feminine	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Kindergartners	44.5	24.7	52.9	23.6	42.5	25.6	40.1	24.4	47.9	25.3	49.6	23.7	39.2	25.4	43.3	24.5
Second graders	33.8	25.0	57.8	27.7	44.9	27.9	43.4	25.8	42.8	27.6	48.8	30.1	36.1	25.6	52.1	25.7
Fourth graders	32.4	27.8	62.1	29.3	44.4	28.1	48.0	27.7	43.2	32.3	51.4	31.6	30.6	22.5	61.7	24.0
Adults	28.7	25.5	67.0	25.7	48.8	30.9	50.0	29.8	43.8	32.0	51.9	31.5	31.0	23.8	67.8	24.3

Table 5
Numbers of Spanish- and French-Speaking Adults (of 32) Who Assigned a Female Voice to Each Item: Study 1

Item	Spanish gender	French gender	Spanish adults	French adults	χ^2 *p*
		Different Spanish and French genders			
fork	M	F	13	26	<.001
plate	M	F	17	30	<.001
car	M	F	7	14	<.10
screw	M	F	2	11	<.01
tomato	M	F	17	30	<.001
peanut	M	F	12	17	*ns*
bat	M	F	4	16	<.01
banana	M	F	14	23	<.05
bed	F	M	22	6	<.001
kite	F	M	23	11	<.01
broom	F	M	28	15	<.01
canoe	F	M	9	1	<.01
grape	F	M	24	11	<.01
cloud	F	M	24	17	<.10
zebra	F	M	17	6	<.01
butterfly	F	M	30	23	<.05
		Same Spanish and French genders			
table	F	F	28	23	*ns*
spoon	F	F	25	25	*ns*
arrow	F	F	10	15	*ns*
guitar	F	F	16	18	*ns*
spider	F	F	21	21	*ns*
apple	F	F	26	29	*ns*
star	F	F	22	23	*ns*
feather	F	F	22	22	*ns*
airplane	M	M	3	4	*ns*
book	M	M	15	12	*ns*
telescope	M	M	3	3	*ns*
knife	M	M	6	5	*ns*
tree	M	M	12	5	<.05
corn	M	M	19	8	<.01
sun	M	M	13	12	*ns*
fire	M	M	2	8	<.05

Note. M = male; F = female.

suggested a correlation between Spanish gender and English speakers' assignments of male and female voices to inanimate objects. In past work, however, the effects of Spanish gender were correlated with the artifact–natural kind distinction. Thus, we counterbalanced the number of masculine and feminine artifacts and natural kinds in the present study to further explore this finding. First, we analyzed English and Spanish speakers' female voice assignments through a Language (English or Spanish) × Grade (kindergarten, second, fourth, or adult) × Condition (pictures only or pictures plus labels) × Spanish Gender (masculine or feminine) × Conceptual Kind (artificial or natural object) ANOVA. Language, grade, and condition were between-subjects factors. Spanish gender and conceptual kind were within-subject factors. We observed main effects of grade, $F(3, 236) = 3.31$, $p < .05$; Spanish gender, $F(1, 236) = 195.67$, $p < .001$; and conceptual kind, $F(1, 236) = 89.63$, $p < .001$. We observed two-way interactions between language group and Spanish gender, $F(1, 236) = 10.84$, $p < 01$; between grade and Spanish gender, $F(3, 236) = 14.40$, $p < .001$; between condition and grade, $F(3, 236) = 2.70$, $p < .05$; and between grade and conceptual kind, $F(3, 236) = 8.92$, $p < .001$. We observed three-way interactions among language, grade, and Spanish gender, $F(3, 236) = 4.19$, $p < .01$; and among language, grade, and conceptual kind, $F(2, 236) = 4.19$, $p < .01$. No other main effects or interactions reached statistical significance.

We interpreted all of the main effects and lower order interactions in the context of the higher order interactions in which the lower order effects participated. As previously noted, the Spanish speakers who were in second grade or older reliably assigned voices in a manner consistent with Spanish gender (Tukey's HSD = 1.65, $p < .05$). The Grade × Condition interaction indicated that an effect of language gender was observed only when the pictures were labeled for second graders. Providing labels for the pictures did not matter for the other language groups; fourth graders and adults showed an effect of gender even when the pictures were presented alone. Table 6 shows these means. Our post hoc analyses also yielded an "effect" of Spanish gender among speakers of English. Of course, there cannot be true effects of Spanish gender on English speakers. However, we observed reliable differences in the assignment of male and female voices in a manner consistent with Spanish gender among English speakers who were in fourth grade or older. Because we counterbalanced the number of natural kinds and artifacts that we included in each Spanish grammatical gender class, this effect indicates that there are features beyond the artifact–natural kind distinction that are correlated with Spanish gender and are leading English speakers to classify items consistently with Spanish gender.

We conducted post hoc analyses to identify the features correlated with the Spanish grammatical gender that seem to be influencing English speakers' classifications. Initially, we identified four factors often associated with males and females, following Leinbach, Hort, and Fagot (1997) and Mullen (1990): (a) artifact or natural kind, (b) angular or curved, (c) used typically by males or females, and (d) dense or not dense. Each stimulus item used in the present study was then judged on each of these attributes by three of the authors. Next, five groups of items were formed on the

Table 6
Mean Numbers of Female Voice Assignments (of 16 Possible Trials) Made by English and Spanish Speakers for Spanish Masculine and Feminine Items: Study 1

	Pictures plus labels				Pictures only			
	Feminine		Masculine		Feminine		Masculine	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
				Spanish speakers				
Kindergartners	6.2	3.0	4.9	2.6	7.6	2.3	7.6	2.5
Second graders	9.1	2.5	4.7	2.4	7.5	2.7	6.9	2.4
Fourth graders	9.8	2.1	4.8	1.9	9.9	2.6	5.1	1.7
Adults	10.8	3.1	5.2	2.2	10.9	2.7	4.7	2.8
				English speakers				
Kindergartners	7.3	3.9	6.0	3.3	8.2	2.1	6.5	2.1
Second graders	8.2	2.4	6.9	1.8	7.5	2.4	6.3	2.4
Fourth graders	8.8	1.8	6.4	1.4	8.8	1.5	6.6	1.6
Adults	8.9	1.3	5.6	1.9	9.2	2.0	6.1	1.8

basis of the number of factors (0–4) per item judged as having a feminine quality. For example, the spoon was categorized as a curved (female-like), dense (male-like) artifact (male-like) that is typically used by females (female-like).

Seven items were judged as having no feminine factors (bed, canoe, book, screw, telescope, car, and plane); 7 had one feminine factor (kite, arrow, table, tree, bat, fork, and knife); 10 had two feminine factors (zebra, star, spoon, plate, corn, peanut, banana, fire, broom, and guitar); 6 were judged as feminine on three factors (grape, apple, spider, feather, sun, and tomato); and 2 were rated as feminine on all four factors (cloud and butterfly). We then calculated the percentage of feminine classifications for each of these five groups based on the number of feminine factors in each group. Thus, the predicted percentage of feminine classifications for the objects that were rated as feminine on all qualities (cloud and butterfly) was 100%; the predicted classification for items that were rated as feminine on three of the four factors was 75%; the predicted classification for items that were rated as feminine on two of the four factors was 50%; the predicted classification for items that were rated as feminine on one of the four factors was 25%; and the predicted classification for items that were rated as feminine on none of the four factors was 0%.

We then compared the predicted pattern of classification with the classifications made by adult speakers of the three language groups. We included speakers of French to determine whether similar effects would be found among French speakers. (If a true correlation exists between Spanish gender and certain properties, an apparent effect of Spanish gender should be observed for French speakers as well.) Table 7 shows the predicted and observed percentages of feminine classifications in all language groups for groups of objects rated as being feminine on four, three, two, one, or no qualities. The last column in Table 7 shows the grammatical gender of the items in Spanish. The four factors (kind, shape, context, and density) seemed to accurately predict the percentage of feminine classifications made by speakers of the three languages. The Spanish gender of the items also corresponded to the predicted pattern of classifications. This suggests a close correlation between the Spanish grammatical gender system and conceptual and perceptual features that are available to adult speakers of other languages.

We now interpret the remaining reliable effects observed from our ANOVA comparing English and Spanish speakers' judgments. The three-way interaction involving language, grade, and conceptual kind indicated a reliable effect of conceptual kind (Tukey's HSD = 1.24, $p < .05$) among English speakers in fourth grade and older and among Spanish speakers in second grade and older. Table 8 shows the means that made up this three-way interaction. The reliable effect of conceptual kind among Spanish speakers also offers evidence against the possibility that Spanish speakers were explicitly referring to grammatical gender in their judgments: Explicitly referring to grammatical gender should have yielded no effect of conceptual kind among speakers of Spanish, because there were equal numbers of feminine and masculine items within each conceptual kind.

The final reliable effect that emerged from this ANOVA was a two-way interaction between grade and condition indicating that labeling the pictures led to more male-like judgments by kindergartners (Tukey's HSD = 1.39, $p < .05$). Although children and adults have been shown to have a bias in favor of male-like representations (e.g., Khosroshahi, 1989), it is not clear why they were more likely to show this bias when the objects were labeled. Note, however, that even though the three-way interaction involving condition, grade, and language was not reliable, it appears that this condition effect among kindergartners originated from Spanish speakers (see Table 6).

To explore whether aspects of French gender were correlated with features characterized by English speakers as male-like or female-like, our final set of analyses compared the judgments of French and English speakers. We submitted the number of female voice assignments by French and English speakers to a Language (French or English) × Grade (kindergarten, second, fourth, or adult) × Condition (pictures only or pictures plus labels) × French Gender (masculine or feminine) × Conceptual Kind (artificial or natural object) ANOVA. Language, grade, and condition were between-subjects factors; French gender and conceptual kind were within-subject factors. We observed main effects of French gender, $F(1, 239) = 80.12$, $p < .01$, and conceptual kind, $F(1, 239) = 74.77$, $p < .01$, and interactions between grade and French gender, $F(3, 239) = 6.13$, $p < .01$; between condition and French gender, $F(1, 239) = 9.30$, $p < .01$; between grade and conceptual kind, $F(3, 239) = 8.56$, $p < .01$; and between language and French gender, $F(1, 239) = 83.23$, $p < .01$. We observed reliable three-way interactions among language, grade, and French gender, $F(3, 239) = 4.27$, $p < .01$; among language, condition, and French gender, $F(1, 239) = 18.89$, $p < .01$; and among grade, condition, and French gender $F(3, 239) = 3.34$, $p < .05$. We also observed

Table 7
Mean Percentages of Feminine Classifications for Groups of Objects Possessing Zero to Four Feminine Factors by Spanish, English, and French Speakers: Study 1

Number of feminine factors	Predicted by number of feminine factors	Spanish speakers	English speakers	French speakers	Spanish gender
Four (2)	100	76.98	72.50	63.38	100
Three (6)	75	55.98	58.33	57.42	67
Two (10)	50	49.46	49.46	49.43	50
One (7)	25	42.40	42.40	47.43	43
Zero (7)	0	32.03	29.57	33.18	29

Note. The factors were natural kind–artifact, curved–angular, female context–male context, and not dense–dense. The numbers inside parentheses represent the numbers of items that were classified into each set. The last column shows the percentages of these items that are feminine in Spanish.

Table 8
Mean Numbers of Female Voice Assignments Made by English and Spanish Speakers for Artificial and Naturally Occurring Objects (of 16 Possible Trials): Study 1

	English speakers				Spanish speakers			
	Artifacts		Natural kinds		Artifacts		Natural kinds	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Kindergartners	6.9	2.8	7.1	3.0	6.4	3.0	6.8	2.4
Second graders	6.6	1.8	7.8	2.4	6.2	2.0	7.9	2.2
Fourth graders	7.0	2.0	8.3	1.6	6.2	1.6	8.5	1.8
Adults	5.8	1.5	9.2	1.6	7.1	1.5	8.7	1.9

a reliable four-way interaction including language, grade, condition, and French gender, $F(3, 239) = 2.75$, $p < .05$.

We interpret first the Grade × Conceptual Kind interaction, because conceptual kind did not interact with any other factors. Post hoc analyses of this interaction revealed that naturally occurring objects were more likely to be assigned female voices than artificial objects by children who were in second grade and older (Tukey's HSD = 0.90, $p < .05$). Thus, older speakers of French, like speakers of Spanish and English, also classified naturally occurring objects as more female-like than artifacts. We interpret the remaining reliable findings in the context of the four-way interaction involving language, grade, condition, and French gender through Tukey tests (HSD = 3.87, $p < .05$). Table 9 shows the percentage of female voices assigned by English and French speakers in each condition. We did not observe reliable effects of French gender among English speakers or among French-speaking kindergartners. We observed effects of French gender among the French-speaking second and fourth graders when the pictures were presented with their labels but not when the pictures were presented alone. For French-speaking adults, effects of grammatical gender were reliable when the pictures were presented both alone and with their labels. The effect of labeling that we found for French second and fourth graders might be due to the possibility that there exist different French words (with different genders) that can be used to describe the pictures employed. Perhaps the French-speaking children mentally "named" the items differently than we assumed they would. Instead of using *un fue,* the masculine word for "fire," they may have represented it as *une flamme,* which is feminine. Instead of using *une assiette,* the feminine word for "plate," they may have represented it as *un plat,* which is masculine. This might have also occurred among the Spanish-speaking second graders (see Table 6).

Discussion

We designed this study with three goals. Our first goal was to determine whether grammatical gender effects on classification vary as the agreement of gender of items across French and Spanish varies. We found that Spanish and French speakers in second grade and older made grammatically consistent voice assignments more often when the items had concordant genders in French and Spanish than when the languages disagreed on the grammatical gender of the items. Among adults, this pattern of judgments was found for 30 of the 32 individual stimulus items. Finding that classification varies systematically as a function of the grammar of items across two languages provides strong evidence that grammatical gender influences classification. It is difficult to imagine a factor other than grammatical gender that could account for the disagreement between French and Spanish speakers.

Our second goal was to better understand the correlation between Spanish grammatical gender and English speakers' judgments. Because we counterbalanced the number of grammatically masculine and feminine items that were artificial objects and natural kinds in the current study, the results point to factors beyond the natural–artificial distinction that are correlated with Spanish grammatical gender and lead to apparent Spanish gender effects among speakers of English. Distinctions such as the item's shape (angular vs. curved), its density (heavy vs. light), and its common use (whether it is used primarily by males or females) map onto the Spanish grammatical gender assignments and are probably responsible for the systematicity of English speakers' judgments. Thus, the relation between Spanish grammatical gender and the perception of items being male-like or female-like is not arbitrary. From a practical standpoint, these distinctions offer speakers of English a heuristic for guessing the gender of Spanish nouns.

The final goal of this study was to determine whether French speakers are influenced by grammatical gender in a manner similar to Spanish speakers. The effects of French gender on French speakers' judgments indicate that grammatical gender effects are not specific to Spanish but are also evident in speakers of a language with a similar grammatical gender system. Finding that reliable grammatical gender effects first emerge for both speakers of French and Spanish who are in second grade (about 7 years of age) suggests that the same underlying factors are at work in the development of Spanish and French speakers' classifications. These effects of language on categorization may strike one as rather "late." After all, 3- to 4-year-old children have excellent command of grammatical gender in Spanish and French (Karmiloff-Smith, 1979; Pérez-Pereira, 1991), and it is reasonable

Table 9
Mean Numbers of Female Voice Assignments (of 16 Possible Trials) Made by English and French Speakers for French Masculine and Feminine Items: Study 1

	Pictures plus labels				Pictures only			
	Feminine		Masculine		Feminine		Masculine	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
	French speakers							
Kindergartners	8.1	2.7	7.1	2.4	8.9	2.2	7.4	2.1
Second graders	11.1	3.0	4.2	2.8	7.4	2.1	6.6	2.5
Fourth graders	11.6	3.1	4.1	2.5	8.2	2.4	6.3	2.2
Adults	10.9	2.8	3.8	1.4	10.6	2.6	5.4	3.3
	English speakers							
Kindergartners	6.2	3.8	7.1	3.2	7.4	1.8	7.4	2.0
Second graders	7.3	1.9	7.8	1.7	7.3	2.1	6.5	2.8
Fourth graders	7.8	1.4	7.5	1.6	7.6	2.2	7.8	1.8
Adults	7.1	1.8	7.4	2.2	7.8	2.0	7.5	1.4

to expect to find grammatical gender effects as soon as grammatical gender is acquired. However, the late effects of language on cognition that we have observed could be due to two possibilities. One is that our task underestimated the effect of language on cognition. The findings of Martinez and Shatz (1996) suggesting differences between English and Spanish 4-year-olds are consistent with our measure not being sufficiently sensitive to detect subtle language effects. However, it is also possible that a substantial amount of time is necessary for language, after it is acquired, to affect cognition. Our results might suggest as well that the time lag between language acquisition and the infiltration of language into the cognitive system may be in the order of years.

Study 2

Study 2 had three goals. One was to determine whether classifications among German speakers, like the classifications of French and Spanish speakers, are affected by grammatical gender. We selected speakers of German because the German grammatical gender system differs from the French and Spanish systems as follows: (a) The German gender system has three categories (masculine, feminine, and neuter) instead of two; (b) the German gender system is case based, so the morphological markings associated with a noun vary with the noun's case (e.g., nominative, accusative, genitive, or dative); and (c) German noun phrases do not carry as much morphological gender information as French and Spanish noun phrases. The second goal of this study was to determine whether speakers of German, like English speakers, tend to classify inanimates according to Spanish grammatical gender. Finding an effect of Spanish gender among speakers of German would provide further evidence that the relation between Spanish grammatical gender and the classification of items as male- or female-like is not arbitrary. The third goal was to examine when (if ever) the effects of grammatical gender emerge among German-speaking children. Because the German gender system is structurally different from the French and Spanish systems, it is possible that effects of grammatical gender, if they emerge, emerge at a different age for speakers of German.

Method

Participants. The participants were 53 monolingual speakers of German and 50 monolingual speakers of English. Each language group was further divided into three age groups: 5-, 7-, and 9-year-olds. The youngest German-speaking group ranged in age from 4 years 8 months to 6 years 1 month (M = 5 years 4 months). The youngest English-speaking age group ranged in age from 4 years 7 months to 6 years 0 months (M = 5 years 5 months). The next German-speaking age group ranged in age from 6 years 10 months to 7 years 11 months (M = 7 years 6 months). The corresponding English-speaking age group ranged in age from 7 years 0 months to 7 years 11 months (M = 7 years 5 months). The oldest German-speaking children ranged in age from 9 years 1 month to 10 years 4 months (M = 9 years 8 months). The corresponding English-speaking age group ranged in age from 9 years 2 months to 10 years 7 months (M = 9 years 9 months). German speakers were recruited from public day care facilities in Hamburg, Germany, and were tested by an English–German bilingual. The English speakers were recruited from a public day care center in Merrillville, Indiana, and from birth announcements in Bloomington, Indiana. All participants received compensation for taking part. No adult speakers of German were tested, because it is nearly impossible to find a large number of native monolingual German speakers (i.e., German-speaking adults commonly speak one or two additional languages).

Materials. The stimuli consisted of a series of 32 colored, laminated flash cards similar to those used in Study 1. The objects were selected on the basis of their grammatical gender in German and Spanish. The objects were separated into two categories, control items and test items. The 8 control items, which were used as an indication of whether the participants understood the task, consisted of objects that had a clear natural gender (e.g., ballerina and boy). The 24 test items were pictures of familiar objects, and these items were used to determine whether the German grammatical gender system influenced the German speakers' classifications. Specifically, 8 objects that were grammatically masculine and grammatically feminine in both German and Spanish were used. Objects with discordant genders across the two languages were also used, such that we had 4 objects that were masculine in German but feminine in Spanish, 4 objects that were feminine in German but masculine in Spanish, 4 objects that were neuter in German but masculine in Spanish, and 4 objects that were neuter in German but feminine in Spanish. Table 10 describes the pictured objects used.

Procedure. Participants were told that they were going to be shown a series of pictures and that they were to assign either a man's voice or a

Table 10
German and English Items Used in Study 2

German	English
Control items	
die Braut	bride
die Ballerine	ballerina
die Nonne	nun
die Frau	woman
der Junge	boy
der König	king
der Mann	man
der Reise	giant
Masculine	
der Apfel	apple
der Fisch	fish
der Hubschrauber	helicopter
der Hut	hat
der Löffel	spoon
der Schuh	shoe
der Stern	star
der Tisch	table
Feminine	
die Feder	feather
die Gabel	fork
die Gitarre	guitar
die Hose	pants
die Spinne	spider
die Traube	grapes
die Schraube	screw
die Uhr	watch
Neuter	
das Buch	book
das Fahrrad	bicycle
das Fernrohr	telescope
das Feuer	fire
das Haus	house
das Messer	knife
das Ohr	ear
das Schwert	sword

woman's voice to the pictured object. English speakers were given the same general instructions as in Study 1 (see the Appendix for the German instructions). Approximately half of the participants in each age group were shown only the pictures. The other half were shown the pictures and were verbally provided with their accompanying labels and articles. The pictures were presented in two randomly determined orders. Each participant was tested by an experimenter who was fluent in the participant's native language. Participants were tested individually either in a corner of the day care center set aside for the experiment or in a laboratory testing room.

Results

We first report the results from the control items that possess a natural gender. Every participant within every age group gave a female voice to female items and a male voice to male items. Thus, we are confident that the participants understood the task. We then conducted two analyses to examine performance on the test items. The first analysis examined the potential effect of German grammatical gender among German- and English-speaking children. The second analysis examined potential effects of Spanish gender in English and German speakers' classifications.

The classifications made by German- and English-speaking children are shown in Table 11. Our first analysis consisted of a Language (German or English) × Age (5 years, 7 years, or 9 years) × Condition (pictures only or pictures plus labels) × German Gender (masculine, feminine, or neuter) ANOVA. Language, age, and condition were between-subjects factors; German gender was a within-subject factor. We observed a main effect of German gender, $F(2, 186) = 6.46$, $p < .05$, and a two-way interaction between language and German gender, $F(2, 186) = 8.84$, $p < .05$. No other reliable main effects or interactions emerged.

In terms of the Language × German Gender interaction, German speakers were more likely to assign grammatically feminine objects female voices (50.3%) than were English speakers (37.0%). No differences were found between German and English speakers in their classifications of grammatically masculine and grammatically neuter objects. Moreover, we observed no reliable differences among German speakers in their classification of items as a function of German grammatical gender. The grammatically feminine items were not assigned female voices more than 50% of the time. Overall, these findings indicate that the German gender system does not affect classification like the Spanish and French gender systems do.

In Study 1, we found an apparent effect of Spanish gender among English speakers. Thus, our second analysis for Study 2 examined whether speakers of German also assigned voices to objects in a manner consistent with Spanish grammatical gender. This analysis consisted of a Language (German or English) × Age (5 years, 7 years, or 9 years) × Condition (pictures only or pictures plus labels) × Spanish Gender (masculine or feminine) ANOVA. Language, age, and condition were between-subjects factors; Spanish gender was a within-subject factor. Table 12 shows the mean percentages of times children assigned female voices to items that were masculine and feminine according to Spanish gender. The only significant effect observed was a main effect of Spanish gender, $F(1, 99) = 37.45$, $p < .001$. On the whole, both German- and English-speaking children assigned fewer female voices to Spanish masculine objects (35.5%) than to Spanish feminine objects (46.38%). Both groups demonstrated an apparent effect of Spanish gender by assigning more male voices to Spanish masculine items than to Spanish feminine items. The inclusion of a different group of English speakers in this study served as a replication of the finding of Study 1 that English speakers assign male and female voices in a manner consistent with Spanish gender. The inclusion of German speakers extended the apparent Spanish effect to speakers of a language with another grammatical gender system.

Discussion

One goal of this study was to determine whether speakers of German are influenced by the German grammatical system in a categorization task in the same way speakers of French and Spanish are influenced by their respective languages' grammatical systems. Although we observed one reliable difference between the classifications of German and English speakers, the German speakers' judgments did not vary systematically according to German grammatical gender in the same way the judgments of French and Spanish speakers varied. Feminine items were not assigned female voices at above-chance levels. Thus, the specific properties of a language's grammatical gender system (i.e., whether it is case based and gender loaded and has two or three gender categories) seem to determine whether effects of grammatical gender will be observed for male- and female-like judgments. Finding that speakers of German do not make grammatically consistent gender assignments is also evidence against the possibility that our task leads speakers of a language with any grammatical gender system to classify according to gender. If the task itself were generating effects of grammatical gender, we should have observed these effects among German speakers. Another possibility is that factors other than the differences in language

Table 11
Percentages of Times German and English Speakers Assigned Female Voices to German Grammatically Feminine, Masculine, and Neuter Items as a Function of Age: Study 2

	German speakers–German grammatical gender						English speakers–German grammatical gender					
	Feminine		Masculine		Neuter		Feminine		Masculine		Neuter	
Age (years)	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
5	49.7	19.8	45.0	24.2	24.4	19.0	38.6	15.3	35.5	17.8	41.9	17.8
7	46.9	16.1	35.9	17.6	33.6	20.3	33.8	15.5	47.5	17.9	40.6	22.2
9	54.4	22.9	45.6	19.6	35.6	23.0	38.8	12.4	53.8	11.9	36.2	16.5

Table 12
Percentages of Times German and English Speakers Assigned Female Voices to Spanish Grammatically Feminine and Masculine Items as a Function of Age: Study 2

	German speakers–Spanish grammatical gender				English speakers–Spanish grammatical gender			
	Feminine		Masculine		Feminine		Masculine	
Age (years)	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
5	45.1	16.9	34.3	15.8	41.3	12.7	35.6	15.0
7	46.9	16.3	30.7	14.2	47.9	14.2	33.3	13.3
9	47.1	12.8	43.3	15.8	50.0	6.6	35.8	7.5

(i.e., nonlinguistic cultural differences between German speakers and speakers of French and Spanish) are responsible for the different pattern of findings. Study 4 addressed this possibility. It is also worth noting that the oldest German speakers that we tested were in fourth grade, and although we observed gender effects for Spanish and French speakers by second grade, it is possible that German speakers are affected by German gender at an older age than we tested. However, if the effects of German gender occur on a different developmental schedule than Spanish and French effects, this would still suggest that language factors operate on cognition differently in German than they do in Spanish and French.

The second goal of this study was to investigate whether German speakers would systematically assign male and female voices in a manner consistent with Spanish gender, as did English speakers in Study 1. We found apparent effects of Spanish gender on English and German speakers' classifications. These results provide additional evidence indicating that the classification of particular items as masculine or feminine in the Spanish language is not arbitrary. They also suggest that the Spanish grammatical gender classifications may be more "natural" or easier to learn than the grammatical gender assignments made in French and German.

Study 3

In this study, we used a connectionist model to better understand the findings from English, Spanish, and German speakers in Studies 1 and 2. We also used the model to simulate the apparently contradicting evidence regarding the relation between grammatical gender and categorization obtained with different tasks within speakers of a language. For example, the results from Spanish and French speakers in Study 1 indicating that language affects classification contrast with the results from Pérez-Pereira (1991) and Karmiloff-Smith (1979) that are typically used to argue for the independence of language from cognition. Our simulations were designed to mimic both the interaction between grammatical gender and classification that was found in Studies 1 and 2 and the independence between classification and grammatical gender found by Karmiloff-Smith (1979) and Pérez-Pereira (1991). Finally, we used the simulations to pinpoint the critical difference between Spanish and German gender systems leading to the different findings reported in Studies 1 and 2.

First, some caveats regarding our simulations are in order. Our primary reason for using connectionist networks was to address how differences across languages might be studied in a principled manner. We recognize that there exist other ways of formally describing the relation between language and cognition. Similarly, our model is intended neither as a model of language acquisition nor as a model of object perception. Our main purpose for using these simulations was to characterize differences across languages and how these differences may interact with other cognitive processes.

Method

Overview of the model and simulations. All of our simulations involved simple auto-association networks. In these simulations, the network learns to produce as output the input that was presented to it. Thus, there is no "teacher" that provides the network with any more information than was presented as input. Our model consisted of three layers: an input layer containing 42 nodes, a hidden layer of 8 units, and an output layer of 42 nodes. Figure 1 illustrates the structure of the model.

The input to the model consisted of a processed perceptual image and noun phrase. The input layer was not fully interconnected to the intermediate layer of hidden units; instead, four of the hidden units received input from the perceptual input units, and four received input from the noun phrase input units. This pattern of "modular" connectivity was suggested by the results of numerous (unsuccessful) simulations. It was easy to simulate the effects of language on categorization with fully interconnected networks, but not the lack of such effects. The modular architecture we

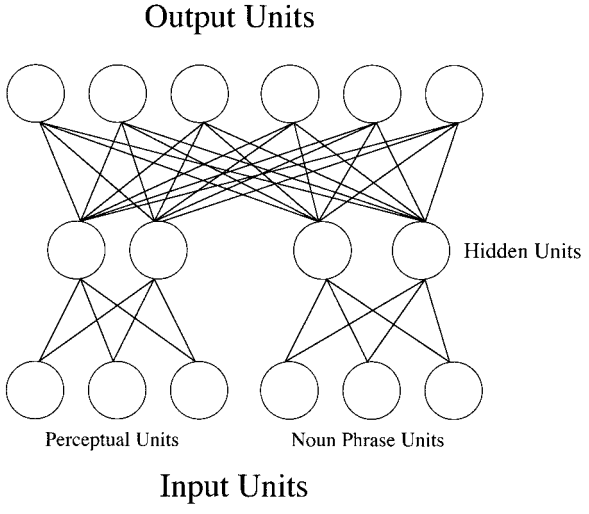

Figure 1. Architecture of the networks that were used in all simulations.

used followed the strategy used by Plunkett, Sinha, Møller, and Strandsby (1992) to simulate word production, comprehension, and vocabulary growth. It is interesting that modular architectures, in which perceptual and language input nodes are connected to different hidden units, appear necessary to simulate a variety of findings in the literature.

All of the simulations involved a training and test phase. During training, the network learned to associate an image with a noun phrase that corresponded to the 18 animate nouns most frequently spoken to children (Wetman & Hauss, 1969). The differences across languages were captured by differences in the coding of the noun phrases. In separate testing phases, networks classified novel test items in a task analogous to the "voices" task used in Studies 1 and 2 and in a task analogous to those used by Karmiloff-Smith (1979) and Pérez-Pereira (1991). The final set of simulations systematically transformed the German gender system into the Spanish gender system and was used to find the critical difference between the Spanish and German gender systems responsible for the results from Spanish versus German speakers in Studies 1 and 2. In one set of these simulations, we "loaded" the German noun phrases with gender as they are in Spanish. In the second set of simulations, we removed case by coding all of the German noun phrases like the nominative case. In the third set of these simulations, we removed the neuter category by coding each noun phrase as masculine or feminine.

Input layer. The input layer contained 42 nodes. Twenty-seven of these nodes coded perceptual and conceptual properties of the image. These nodes were needed to distinguish among the training items. Although our post hoc analyses from Study 1 indicated that only four properties are needed to understand the apparent effects of Spanish gender among English speakers, these properties alone did not distinguish among the training items, because many of the most common nouns spoken to children refer to males and females (e.g., mother, father, and sister) and likely play a major role in establishing the acquisition of grammatical gender classes in the first place. Thus, we had to include more input nodes to distinguish between these items.

We coded the perceptual properties associated with males with values of 1 and the properties associated with females with values of 0. For example, Input Nodes 1 and 2 coded the degree to which the items were angular or curved. Highly curved items were coded with values of 0 on Nodes 1 and 2; angular items were given values of 1 on the same nodes; and items with shapes suggesting both angular and curved components (e.g., a bird) were assigned intermediate values of 0 and 1 on these same input nodes. The next pair of input nodes (3 and 4) coded whether the objects were animate or inanimate. During training, Input Nodes 5 and 6 coded the natural gender of the objects. So, for example, "girl" was coded as female (0, 0) and "boy" as male (1, 1). The output nodes that corresponded to Input Nodes 5 and 6 were probed during testing in an attempt to simulate the classification of inanimate objects in the "voices" task.

The fourth pair of input nodes (7 and 8) coded the age of the items. Input Nodes 9–11 coded spousal, sibling, or parent–offspring kinship relations for animates. These nodes were necessary to distinguish between pairs of inputs such as daughter versus girl and mother versus woman. The next set of input nodes (12 and 13) coded the sizes of the objects. For example, these nodes captured the difference in size between "man" and "boy." Input Nodes 14 and 15 coded the loudness of the items. Thus, for example, "boy" was coded as being louder than "bird." The next pair of input nodes (16 and 17) coded whether the objects were psychologically or physically "warm" versus "cold." For example, "woman" was coded as warmer than "doctor." Input Nodes 18 and 19 represented the strength of the items. Therefore, "man" was coded as stronger than "girl." Input Nodes 20 and 21 represented the object's potential for nurturance. So, for example, "mother" and "woman" were coded as being different in nurturance (Units 20 and 21) but not warmth (Units 16 and 17).

The global shape of the items—long and thin, long and fat, short and thin, or short and fat—was represented by Input Nodes 22 and 23. The next pair of perceptual nodes (24 and 25) represented stereotyped associations with the objects. For example, the inanimate item "arrow" was coded as being associated with men. The final pair of perceptual input nodes (26 and 27) represented the density of the objects. Thus, "bird" was coded as being less dense than "girl." The features were motivated by past research as well as our findings from Studies 1 and 2. We also obtained ratings of the perceptual properties of all items from a group of six adult English speakers and used their ratings to code the perceptual input values in all of the simulations. Table 13 shows how the perceptual input units were coded for two training and test items.

The remaining 15 input nodes coded the linguistic input. Nine of these nodes, the grammatical gender nodes, coded the morphological properties associated with the grammatical gender of the noun phrase that corresponded to each object. The first three noun phrase nodes coded the morphological gender of the indefinite article (1-0-0 for masculine, 0-1-0

Table 13
Examples of How the Perceptual Input Nodes Coded Properties in English and Spanish During the Training and Testing Phases

	Training phase				Testing phase for the "voices" task			
Perceptual property	girl–niña		husband–esposo		king–rey		table–mesa	
Angular–curved	0-1	0-1	1-1	1-1	1-1	1-1	1-1	1-1
Animacy	0-0	0-0	0-0	0-0	0-0	0-0	1-0	1-0
Natural gender	0-0	0-0	1-1	1-1	.5-.5	.5-.5	.5-.5	.5-.5
Age	0-0	0-0	1-0	1-0	1-0	1-0	0-1	0-1
Kinship	1-1-1	1-1-1	1-0-0	1-0-0	1-1-1	1-1-1	1-1-0	1-1-0
Size	0-0	0-0	1-1	1-1	1-1	1-1	1-1	1-1
Loudness	1-0	1-0	1-1	1-1	1-1	1-1	0-0	0-0
Warm–cold	0-0	0-0	0-1	0-1	1-1	1-1	1-1	1-1
Strength	0-0	0-0	1-1	1-1	1-1	1-1	1-1	1-1
Nurturance	0-1	0-1	1-0	1-0	1-1	1-1	1-1	1-1
Global shape	0-1	0-1	1-1	1-1	1-1	1-1	1-1	1-1
Stereotypic associations	0-0	0-0	1-0	1-0	1-0	1-0	0-1	0-1
Density	0-0	0-0	1-1	1-1	1-1	1-1	1-1	1-1

Note. Male-like properties and masculine morphological endings were coded with 1; female-like properties and feminine endings were coded with 0. Intermediate values on dimensions that did not code gender (e.g., size) were coded as 1-0 or 0-1. During the test for the "voices" task, the natural gender input nodes were coded with .5 and .5, and we examined how the networks completed the pattern.

for feminine, 0-0-1 for neuter, and 0-0-0 for other). The next three noun phrase nodes similarly coded the grammatical gender of the noun that referred to the object. The last three noun phrase nodes coded the grammatical gender of the adjective "good" that corresponded to each noun. The output nodes that corresponded to the morphological gender of the adjective were probed during testing in an attempt to simulate the relation between grammatical and perceptual gender in tasks such as the ones used by Karmiloff-Smith (1979) and Pérez-Pereira (1991). The last six language nodes coded each individual noun stem.

Hidden layer. The input layer was connected to eight hidden units. This hidden layer was not fully interconnected to the input layer. The input units coding perceptual and conceptual properties (Inputs 1–27) were connected to four hidden units, and the input units coding linguistic properties (Inputs 28–42) were connected to the other four hidden units.

Output layer. The eight hidden units were fully interconnected to the output layer. The output layer consisted of the same number of units as the input layer.

Training. The model was trained to associate an image (Input Nodes 1–27) with a noun phrase (Input Nodes 28–42) that corresponded to each of 18 animate training items: mother, girl, boy, lady, daughter, son, wife, father, husband, child, doctor, bird, horse, and sister (two versions of doctor, bird, horse, and child [male and female versions] were presented). Each item was presented four times during training.

We simulated the learning of English, Spanish, and German during training as follows. The perceptual units were coded the same way for all three languages. The only difference among the English, Spanish, and German simulations was in the coding of each corresponding noun phrase. Because English does not mark gender morphologically, the pattern of activation of the grammatical gender nodes never corresponded to the natural gender in the perceptual nodes in the English simulation. Thus, the grammatical gender nodes in English were always given the value of 0-0-0. For the Spanish simulation, each pair of grammatical gender nodes (noun, adjective, and article) was coded for grammatical gender (masculine or feminine) according to the Spanish language. We coded Spanish nouns and adjectives that ended in /o/ and /a/ as masculine (1-0-0) and feminine (0-1-0), respectively. Nouns that did not end in /o/ or /a/ were assigned the value of 0-0-0. The Spanish masculine indefinite article, *un*, was coded as masculine, and the indefinite feminine article, *una*, was coded as feminine. There was a close but imperfect correspondence among the grammatical genders of the Spanish nouns, articles, and adjectives. For example, the Spanish nouns "mother" (madre) and "father" (padre) were not coded as either masculine or feminine, but the articles and adjectives included in these noun phrases were. In this simulation, the natural gender of the objects corresponded to some aspect of grammatical gender for all but two training items (e.g., "bird" and "horse" are grammatically masculine in Spanish, but we included both female and male versions during training). Each of the four times each training item was presented in the Spanish simulation, the item was coded exactly the same way.

For the German simulation, each noun, adjective, and definite article was coded in an analogous fashion to the Spanish simulation, with two exceptions. One difference between the Spanish and German simulations was that some components of the German noun phrases were coded as neuter, according to the German language, with a value of 0-0-1. The other coding difference between the German and Spanish simulations was that the German noun phrases were also coded as a function of case (nominative, accusative, dative, and genitive). Different German cases were simulated as follows. Each noun phrase was presented once to the model in each case: nominative, accusative, dative, and genitive. Thus, there were four distinct noun phrases in the German simulation for "the good woman." There is also morphological overlap between German articles. Some indefinite articles within a case are morphologically identical across two genders. For example, in the nominative case, the indefinite masculine article *ein* is the same as the indefinite neuter article *ein.* In these situations, the articles were coded consistently but not like the typical masculine and feminine genders (i.e., as 0-0-0). We coded the grammatical gender of German adjectives in an analogous manner. Table 14 shows how gender morphology for the noun phrase "a good woman" was coded in English, Spanish, and German across different cases.

The simulations were carried out with the software program t-learn. Each item was presented to the network 800 times during training, and because there were 18 training items, each simulation consisted of 14,400 presentations or "sweeps." The learning algorithm used was back-propagation in which the root mean square error was minimized. The results from the simulations that we report involved a learning rate of .25 (see Plunkett & Elman, 1997, for more details about the parameters used by t-learn). Using this learning rate, we presented the networks with the input stimuli in 10 different random orders. Thus, the results from the simulations that we report actually were the mean results from 10 different simulations. We found similar learning curves for all three languages using these parameters. Figure 2 shows the learning curves for all three languages.

Testing. To simulate the results from Studies 1 and 2, we tested the model with the items that appear in Table 15. These were items on which the networks were not trained. All items were presented to the model

Table 14
Coding of the Input Nodes for the Noun Phrase "A Good Woman" in English, Spanish, and German

	Case			
Language	Nominative	Genitive	Dative	Accusative
English	a good woman	a good woman	a good woman	a good woman
Indefinite article	0-0-0	0-0-0	0-0-0	0-0-0
Noun ending	0-0-0	0-0-0	0-0-0	0-0-0
Adjective ending	0-0-0	0-0-0	0-0-0	0-0-0
Spanish	una mujer buena	una mujer buena	una mujer buena	una mujer buena
Indefinite article	0-1-0	0-1-0	0-1-0	0-1-0
Noun ending	0-0-0	0-0-0	0-0-0	0-0-0
Adjective ending	0-1-0	0-1-0	0-1-0	0-1-0
German	eine gute Frau	einer guten Frau	einer guten Frau	eine gute Frau
Indefinite article	0-1-0	0-1-0	0-1-0	0-1-0
Noun ending	0-0-0	0-0-0	0-0-0	0-0-0
Adjective ending	0-1-0	0-0-0	0-0-0	0-1-0

Note. Masculine morphological endings were coded as 1-0-0, feminine endings were coded as 0-1-0, neuter endings in German were coded as 0-0-1, and other endings were coded as 0-0-0.

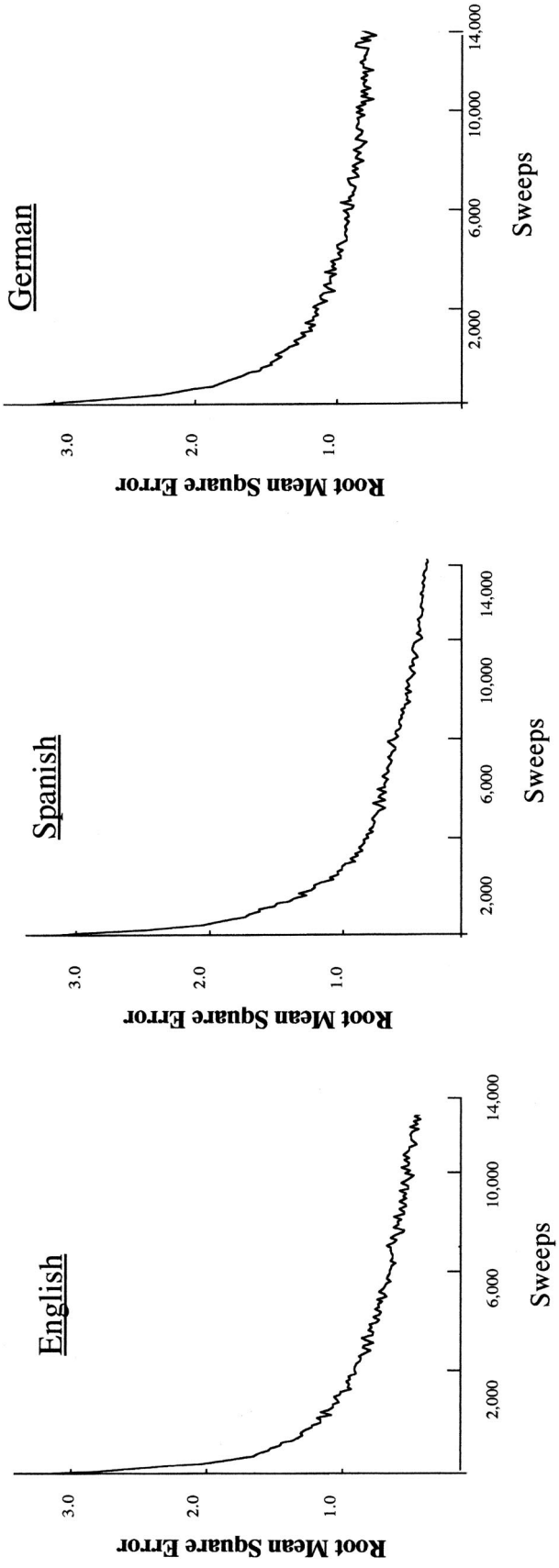

Figure 2. Learning curves after 14,400 sweeps of training the networks with English, Spanish, and German input (learning rate = .25, random seed = 1). Each sweep represents the presentation of one training item to the network.

Table 15
Items Used to Test the Networks' Performance on the "Voices" Task

English	Spanish	German
king	rey (king)	König (king)
man	hombre (man)	Mann (man)
woman	mujer (woman)	Frau (woman)
bride	novia (bride)	Braut (bride)
table	mesa (table)	Gabel (fork)
guitar	guitarra (guitar)	Gitarre (guitar)
arrow	flecha (arrow)	Schraube (screw)
apple	manzana (apple)	Traube (grapes)
grapes	uvas (grapes)	Feder (feather)
feather	pluma (feather)	Spinne (spider)
fork	tenedor (fork)	Tisch (table)
telescope	telescopio (telescope)	Hubschrauber (helicopter)
screw	tornillo (screw)	Löffel (spoon)
bat	murciélago (bat)	Apfel (apple)
fire	fuego (fire)	Fisch (fish)
tomato	tomate (tomato)	Stern (star)

during testing in nominative case, as they were presented to participants in Studies 1 and 2. Four of these test items were control items that possessed a natural gender (e.g., king and bride). The remaining test items were inanimate objects. We probed the model by leaving the input nodes that code natural gender "blank" (with the neutral value of .5) and observing how the model completed the pattern.

To simulate the results from Karmiloff-Smith (1979) and Pérez-Pereira (1991), we used four novel items. The first was coded as perceptually masculine and grammatically masculine. The second was coded as perceptually masculine and grammatically feminine. The third was coded as perceptually feminine and grammatically feminine. The fourth was coded as perceptually feminine and grammatically masculine. In the case of all of these items, the network completed the activation pattern for the grammatical nodes corresponding to the adjective.

Three additional simulation sets systematically varied the differences between the Spanish and German grammatical gender systems. In one set of simulations, we "loaded" the German noun phrases with gender as they are in Spanish. In these simulations, all of the components of the German noun phrase (e.g., article, noun, and adjective) were consistently coded for grammatical gender. In a second set of simulations, we coded all of the German noun phrases like the nominative case. These simulations removed case from the German grammatical gender system. In the third set of simulations, we removed the neuter category by coding all of the elements of the noun phrase during training as masculine or feminine. Thus, for example, in these simulations "woman" was coded with a grammatically feminine article and adjectival ending across all cases.

Results and Discussion

Simulating assignment of gender to objects. The results from the networks that simulated the performance of Spanish, English, and German speakers in Studies 1 and 2 appear in Table 16. As is apparent from Table 16, we were able to closely approximate the pattern of performance by each language group. The correlation coefficients were .931 between the Spanish speakers and the Spanish simulation, .971 between the German speakers and the German simulation, and .962 between the English speakers and the English simulation. In the simulations, as in the language groups, we observed large effects of grammatical gender in Spanish, smaller effects in German, and effects of the artifact–natural kind contrast in English and German. We also observed small effects of Spanish gender in the English simulation and small effects of the natural kind–artifact contrast in the Spanish simulation. By only changing the structure of noun phrases during training, we were able to closely approximate the performance of English, Spanish, and German speakers. We simulated both the effect of grammatical gender on categorization found among Spanish speakers and the lack of effect among speakers of German.

Simulating assignment of gender to adjectives. The results from the simulations that modeled the pattern of performance found in the Pérez-Pereira (1991) and Karmiloff-Smith (1979) studies, in which participants were asked to provide the gender of adjectives, appear in Table 17. The correlation coefficient between the output of this simulation and the responses provided by the 9-year-olds studied by Pérez-Pereira was .984. In both the model and the data, grammatical gender dominated the responses. This pattern of performance among people is typically taken as evidence that language is independent of perception (i.e., modularity). However, both the model and the data suggested that perceptual gender influenced the classifications. For example, when both

Table 16
Percentages of Male-Like Classifications by Adult Speakers of Spanish and English and 9-Year-Old Speakers of German on Grammatically Feminine and Masculine Natural Kinds and Artifacts

	Controls		Natural kinds		Artifacts		
Source	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.	*r*
Spanish speakers	100.0	0.0	76.0	25.1	83.1	43.7	.931
Spanish simulation	94.5	4.5	75.6	0.5	81.2	6.3	
German speakers	100.0	0.0	33.3	31.7	70.0	60.0	.971
German simulation	97.8	2.3	42.4	50.8	73.9	55.4	
English speakers	100.0	0.0	57.3[a]	30.0[a]	79.0[a]	69.7[a]	.962
English simulation	95.7	6.2	78.8[a]	49.8[a]	85.3[a]	72.0[a]	

Note. Percentages of male-like classifications by the simulations trained and tested with input simulating Spanish, German, and English noun phrases are also shown. The correlation coefficients between language groups' classifications and the simulations' classifications are shown at the ends of rows. Masc. = masculine; Fem. = feminine.
[a] Masculine or feminine according to Spanish grammatical gender.

Table 17
Mean Percentages of Times Adjectives Were Classified as Masculine by the Simulation and by the 9-Year-Old Spanish-Speaking Children in the Pérez-Pereira (1991) Study

Perceptual gender	Grammatical gender	Simulation results (%)	Spanish speakers' results (%)
Male	Male	92.9	100
Male	Female	12.9	30
Female	Female	2.5	10
Female	Male	80.0	95

Note. The correlation coefficient between the output of the simulations and the Spanish speakers' judgments was .984.

grammatical and perceptual genders were masculine, the percentages of masculine responses by the children and the model were 100% and 93%, respectively. When the grammatical gender was masculine and the perceptual gender was feminine, the percentages of masculine responses were lower, 95% by children and 80% by the model. These results suggest that the assignment of gender to adjectives describing an object is not completely insulated from perceptual processing of the object's natural gender. They also suggest that at least some of the evidence that has been used to support "modular" accounts of language may arise from statistical patterns in the input.

It appears that the pattern of correlations among natural gender, grammatical gender, and type of gender judgment being made can account for various findings in the literature. In the case of Spanish, there are many cues that are correlated with natural gender; some of these cues are perceptual, whereas others are grammatical. When asked about the gender of an object, speakers of Spanish are likely to rely on all of these cues, some of which are grammatical, leading to effects of grammatical gender in their judgments. Similarly, when asked about the grammatical gender of an adjective, speakers are likely to rely on the pattern of correlations between perceptual and grammatical gender. In the case of adjectives, however, a very good predictor of gender is the gender of the articles and the nouns that compose the noun phrase. Consequently, when asked for the gender of an adjective, people rely on the best predictor of an adjective's gender: the grammatical classification of the corresponding articles and nouns. By our account, the findings of Karmiloff-Smith (1979), Pérez-Pereira (1991), and Sera et al. (1994) and those reported here do not contradict because they all arise from patterns of correlations.

Simulating differences between Spanish and German. The results from the simulations that systematically examined the properties of gender systems that lead to effects of gender in Spanish but not German appear in Table 18. We compared the results from these simulations with the results from the Spanish speakers by using cutoffs of .45 and .55, such that classifications lower than .45 were classified as "feminine" and classifications higher than .55 were classified as "masculine." The simulations that added morphological gender markings to German noun phrases (that "loaded" German noun phrases) and that removed the German case system yielded Spanishlike classifications of grammatically feminine items, but these simulations failed to classify masculine items as masculine. Only the simulation of German without the neuter category yielded the pattern of results observed among Spanish speakers on both masculine and feminine items (see Table 18). These results suggest that the critical difference between Spanish and German is the neuter category. We tested this prediction in Study 4.

Study 4

In this study, we asked whether the difference between Spanish and German speakers we found in Studies 1 and 2 was due to the different number of gender categories in the languages. Three groups of adult, native speakers of English participated. One group learned novel names for 18 animates with two suffixes, similar to the Spanish gender system. A second group learned novel names for the same 18 animates with three suffixes, similar to the German gender system. We then examined how these participants classified novel test items in comparison with a third group that received no language training. The results from the simulations in Study 3 predict that adults who learn two gender categories will be more likely to classify according to Spanish gender than adults who receive no language training. The simulations also predict that such effects will not emerge among adults who learn an artificial language with three gender categories such as German. If, however, the difference between speakers of Spanish (and French) and speakers of German reflects nonlinguistic cultural influences, the type of language training should not matter.

Method

Participants. Fifty-four adult speakers of English with little or no experience in second languages participated. They were assigned randomly to one of three groups. One group was trained in a novel "language" that had two gender categories ($n = 18$), a second group consisted of adults who learned a new language with three gender categories ($n = 18$), and a third, a control group, received no language training but simply completed a classification task ($n = 18$). Data from one person in the control group

Table 18
Percentages of Male-Like Classifications From Simulations and Speakers of German and Spanish in the "Voices" Tasks in Which Different Aspects of the German Gender System Were Transformed to Resemble the Spanish System

	Grammatical gender			
	Masculine		Feminine	
Source	Artifacts	Natural kinds	Artifacts	Natural kinds
German speakers	70.0 (M)	33.3 (F)	60.0 (M)	31.7 (F)
Standard German simulation	73.6 (M)	42.4 (F)	55.4 (M)	50.8 (N)
Loaded German	70.3 (M)	52.4 (N)	32.3 (F)	30.5 (F)
German without case	74.3 (M)	45.0 (F)	22.9 (F)	49.0 (F)
German without neuter	74.8 (M)	62.3 (M)	7.25 (F)	8.4 (F)
Spanish speakers	83.1 (M)	76.0 (M)	43.7 (F)	25.1 (F)

Note. The letters in parentheses represent whether the items were classified as feminine (F), masculine (M), or neither (N) according to the criterion that classifications of 45% or fewer were feminine and classifications of 55% or greater were masculine.

were not used because this individual did not follow the instructions in the classification task. Participants were recruited through an advertisement in the University of Minnesota newspaper and were paid $5 for taking part.

Materials. Fifty-four pictures selected from the Peabody Picture Vocabulary Test were colored and laminated onto 4- × 6-in (10.2- × 15.2-cm) index cards. Eighteen of the cards were used for the groups that received language training. The training items all possessed a natural gender. The stimuli were divided either into two categories or into three categories. For the participants who were trained in the group with two categories, the training stimuli were divided by masculine and feminine; for the participants who were trained in the group with three categories, the stimuli were divided into three groups: masculine, feminine, and both. The control group that received no language training did not receive the training stimuli.

The remaining 36 pictures were test items that were used to examine how language training might affect classification. Four of the 36 test items possessed a natural gender and were used as controls to confirm that participants remembered the suffixes on which they were trained. We tried to select test items that would not be spontaneously classified according to Spanish grammatical gender categories by English speakers. Because past work has shown that natural objects tend to be viewed as female-like and artificial objects as male-like, half of the test items were natural kinds and half were artifacts (16 of each). Half of the natural kind objects were grammatically feminine in Spanish, and half were grammatically masculine in Spanish. Half of the artifacts were grammatically feminine in Spanish and half grammatically masculine in Spanish. We selected the test items so that English speakers in the control group would not spontaneously classify the items according to Spanish gender. If English speakers spontaneously classify the items according to the artifact–natural kind distinction, they should not show an effect of Spanish gender. We attempted to follow the same strategy for German. Twelve of the items were grammatically feminine in German, 10 were masculine, and 10 were neuter. Table 19 describes the pictured test items.

Procedure. Before the experiment, participants completed brief questionnaires describing their experience with second languages. We included only adults who had minimal experience with second languages. All participants were tested individually in a laboratory room.

The training sessions were conducted as follows. The adults who learned the language with two gender categories were shown the 18 training items, and the experimenter labeled each object in English and added either *-wug* or *-dax* as a suffix. For 9 participants, *-wug* denoted masculine gender and *-dax* feminine gender. For the other 9 participants, *-wug* was the feminine suffix and *-dax* the masculine one. For example, *priest* became "priest-wug," and *majorette* became "majorettedax." The experimenter labeled the training cards as often as needed until the participants labeled the cards using the suffixes by themselves 100% correctly. The adults who learned the language with three categories were trained similarly, except that three different suffixes were added to the training items: *-wug, -dax,* and *-riff.* Nine adults in this group learned to add *-wug* to the female items and *-dax* to the male items. The remaining 9 participants added *-dax* to feminine items and *-wug* to masculine items. The suffix *-riff* was always neuter and referred to both males and females. Adults in the control group did not receive any training.

The testing sessions were conducted as follows. After the two groups of adults in the language learning groups learned the suffixes for the training items, they were instructed as follows: "Using *-wug* and *-dax* like before, please label the following items." They were then shown the test items, one at a time. They responded verbally, and their responses were recorded on an answer sheet. The adults who did not receive any language training (the control group) were given the same 36 test items. They were instructed to look at the cards and "divide them into two groups of things that go together." The results from this group should show how speakers of English, which has no grammatical gender system, spontaneously categorize the objects. They should also reveal English-speaking adults' natural tendency to classify the test items according to Spanish or German gender.

Table 19
Descriptions of the Pictured Training and Test Items Used in Study 4

	Test item	
Training item	Artifact	Natural kind
witch	screw	apple
nurse	arrow	bat
majorette	canoe	corn
ballerina	book	sun
queen	bed	butterfly
nun	fork	grapes
girl	telescope	tree
cowgirl	kite	fire
baker	plate	star
pirate	guitar	spider
policeman	airplane	banana
sheriff	spoon	feather
sailor	knife	zebra
boy	car	tomato
giant	broom	cloud
priest	table	peanut
horse	king	woman
bird	man	bride

Results and Discussion

We first report the results from the control group that received no language training. Initially, we designated the two piles of pictures that they produced as Pile 1 and Pile 2 depending on the number of natural kinds and artifacts in each pile. Pile 1 consisted of more natural kinds (88.7% natural kinds and 11.3% artifacts) and Pile 2 (75% artifacts and 25% natural kinds) more artifacts. Then we determined the degree to which the control group's classifications corresponded to Spanish as follows. We calculated the percentage of grammatically masculine items according to Spanish in each pile (i.e., the number of masculine items in Pile 1 vs. Pile 2). Likewise, the percentage of grammatically feminine items in each of the two piles was compared. A *t* test for correlated groups showed no statistically significant difference in the percentage of masculine and feminine items in the two piles, $t(34) = 0.73$. Thus, without training, English speakers did not classify these test items according to Spanish grammatical gender. We followed a similar strategy to determine the degree to which English speakers' spontaneous classifications corresponded to German grammatical gender. However, in these analyses, significant differences emerged between the mean percentage of German masculine items in the different piles, $t(34) = 3.33$, $p < .01$, and the percentage of German feminine items in the two piles, $t(34) = 4.30$, $p < .001$. Thus, English-speaking adults in the control group, who received no language training, spontaneously divided the test items in this study consistently with German gender. Apparently, despite our efforts to avoid such effects, the test items in this study were correlated with German gender.

Next, we examine the classifications from the groups that received language training. All but 2 participants who received the language training applied the suffixes correctly to each of the control items (one of the participants applying the suffixes incorrectly was in the two-category group, and the other was in the

three-category group). One of these participants added suffixes incorrectly to all four control items and was replaced. The other participant added only one suffix incorrectly and was not replaced. The group that learned two categories applied the masculine suffix more often to Spanish masculine than feminine items and the feminine suffix more often to Spanish feminine than masculine items, $t(34) = 3.19$, $p < .01$. This suggests that learning a language with two gender categories that consistently map onto males and females leads people to assign the newly learned suffixes consistently with Spanish grammatical gender. A different pattern of results emerged from the comparisons involving the adults who learned the three suffixes. After receiving training with three "gender" categories, they did not classify according to German gender. The mean percentages of masculine items in different piles assigned different suffixes were reliably different at the $p < .02$ level, $t(34) = 2.68$; however, the effect was in the opposite direction of German gender. Participants were more likely to apply the newly learned feminine suffix to German masculine than to German feminine items. There were no significant differences in application of the newly learned suffixes to the German feminine items, $t(34) = 0.94$.

The results of this study can be briefly summarized as follows. The results from the control group indicate that English speakers spontaneously classified the test items according to German gender. They did not spontaneously classify these items according to Spanish gender. The group that learned a two-category gender-based "language" subsequently classified the test items according to Spanish gender. However, the group that learned a three-category language did not classify the items according to German gender. They tended to apply the newly learned suffixes to half of the test items opposite of German gender. Table 20 shows these results. Admittedly, our two- and three-category artificial gender systems are arguably not much like actual Spanish and German. The Spanish gender system is not as perfectly correlated with natural gender as our two-category training, and the German system is not as poorly correlated with natural gender. Participants in this artificial language learning task also probably learned these categories consciously and explicitly rather than in the unconscious and implicit way in which natural languages are learned. Yet, despite all of these shortcomings, the present results seem to mimic the critical difference that we found between speakers of Spanish and German in Studies 1 and 2. The manner in which something is learned, explicitly or implicitly, may not affect *what* is learned.

How might the results from this study mimic the findings from speakers of Spanish and German? The people in the control group classified the items primarily according to the artifact–natural kind distinction, revealing no effects of Spanish gender but reliable effects of German gender. After learning the two-category gender-based "Spanish-like" artificial language, the participants were primed to base their subsequent classifications on the basis of natural gender and away from their original biases. This led them to classify the test items according to the properties they associate with natural gender that are correlated with Spanish gender, and an effect of Spanish gender emerged. How do these findings simulate our failure to find gender effects among speakers of German? The control group's tendency to spontaneously classify according to the artifact–natural kind distinction is correlated with German gender, leading to effects of German in the free classification task.

Table 20
Mean Percentages of Items Classified According to Spanish and German Gender by English-Speaking Adults Who Were Taught Artificial Languages With Either Two or Three Gender Categories and by Adults Who Received No Language Training: Study 4

	Masculine suffix				Feminine suffix			
	Masculine gender: Spanish		Feminine gender: Spanish		Masculine gender: Spanish		Feminine gender: Spanish	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Two-category language learning	59.4	11.5	40.6	11.5	42.6	13.1	57.4	13.1

	Pile 1 classification				Pile 2 classification			
	Masculine		Feminine		Masculine		Feminine	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
No language training								
Spanish gender	50.9	8.9	49.1	8.9	48.8	4.4	51.2	4.4
German gender	27.2	5.1	50.7	8.1	34.1	3.8	28.1	8.5

	Masculine suffix				Feminine suffix			
	Masculine gender: German		Feminine gender: German		Masculine gender: German		Feminine gender: German	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Three-category language training	27.1	8.5	34.8	8.5	39.0	10.9	38.2	7.7

After taking part in the three-category gender-based "German-like" training, the participants were also primed to base their subsequent classifications away from their original biases. However, it is not clear what bias the three-category training promoted; it is clear only that this training led away from classifying according to German gender.

More important, however, the results of this study suggest that both two- and three-category gender systems may influence classification: Both groups that received language training differed in their classifications from the control group. The Spanish gender system magnifies the tendency to rely on properties associated with males and females. A system such as German magnifies the tendency to rely on other kinds of properties. We might not have found language effects in German in Study 2 because our "voices" task tapped reliance on properties associated with males and females. The results from the current study suggest that effects of German "gender" on classification will not revolve around the properties associated with males and females.

Conclusion

Contradictory findings permeate the language–thought literature. In this work, we focused on the discrepant findings involving grammatical gender. Our main goal was to systematically investigate the relation between grammatical gender and the overgen-

eralization of gender traits to inanimate objects. By examining this relation across four languages with different gender systems, we were able to show how the effect varies as properties of the languages vary. We found that the judgments of Spanish and French speakers varied systematically and predictably with variations in gender assignments across the two languages. We found that the German grammatical gender system does not influence classifications among German speakers in the same way the Spanish and French systems do. We also offered a framework for understanding contradicting results across speakers of different languages on one task—results from the "voices" task among speakers of German and Spanish—as well as contradicting results within language groups across different tasks: Karmiloff-Smith (1979) and Pérez-Pereira's (1991) findings suggesting independence between language and categorization versus our results suggesting an interaction in French and Spanish. We used a connectionist model to simulate the contradictory findings. Although we made a number of assumptions about the nature of learning in these simulations, some of which are likely to be inaccurate, the simulation results demonstrate how a single mechanism can generate apparently contradictory results. Our strategy of closely examining language and task differences enabled us to reconcile many findings in the literature.

Perhaps more important, however, we used this framework as a starting point from which to test hypotheses about the features of grammatical gender systems that lead speakers of Spanish and French, unlike speakers of German, to rely on grammatical gender in their attribution of male and female voices to inanimate objects. Our framework enabled us to study the role that different components of grammatical gender systems (e.g., case, gender loading, and number of categories) play in classification. We found that grammatical factors such as case and gender loading matter little. Instead, our results suggest that a grammatical gender system with only two gender categories, and with a high correlation between grammatical and natural gender, leads to overgeneralization of masculine and feminine traits to inanimate objects. This prediction was confirmed by the results from English-speaking adults in the artificial language learning task used in Study 4. Our approach also enabled us to predict what sorts of language differences and what sorts of tasks are likely to yield interactions between language and thought. We can now use the predictions from our simulations to compare the classification skills of speakers of languages with particular properties. We may also use our model to further investigate the role of grammatical gender in German; even though the grammatical gender system of German failed to lead to overgeneralization of male and female qualities to inanimate objects, it seems to lead its speakers to rely more heavily on other conceptual distinctions. In short, connectionist networks enable us to address questions regarding language–thought relations in a principled fashion as opposed to simply by trial and error. Our work also extends connectionist models. Marcus (1998) has recently argued that connectionist models are encapsulated, predicting performance from only a single task. Our interactive model of language and categorization, which predicts performance in a variety of tasks and could predict performance across a number of categorical and linguistic contrasts, is a counterexample.

We also found that the Spanish gender assignments are not as arbitrary as previously thought. Speakers of English, German, and French assign male and female voices to items corresponding to Spanish gender more often than not. It appears that the Spanish grammatical gender system has a semantic basis that is naturally acquired by speakers of these other languages. Our findings further suggest that properties such as natural–artificial, round–angular, light–heavy, and used by women–men are correlated with Spanish gender assignments. This information should prove useful to nonnative learners of Spanish when trying to remember the gender of Spanish nouns. Instead of having to individually memorize the gender of each noun, remembering these four properties should help Spanish learners make good guesses about a Spanish noun's gender.

Our findings are also relevant for cognitive development. The hypothesis that language affects categorization is a hypothesis about cognitive development. If language is to affect categorization, young children must be flexible in how they classify objects in the world. Our findings indicate that grammatical gender can lead speakers of a language to think about inanimate objects in terms of properties that they associate with males and females. The properties that "pop out" when people think of inanimate objects are the result of a developmental process in which language plays a meaningful role, starting at the age of 7 years.

In summary, we have offered new empirical evidence showing effects of language on categorization, as well as the lack of such effects. We have also offered a framework for understanding a number of contradictory findings in this literature. Our work clarifies the empirical findings and provides a principled point from which to ask future questions.

References

Bellugi, U., Bihrle, A., & Corina, D. (1991). Linguistic and spatial development: Dissociations between cognitive domains. In N. Krasnegor, D. Rumbaugh, R. Schielfelbusch, & M. Studdert-Kennedy (Eds.), *Biological determinants of language development* (pp. 368–398). Hillsdale, NJ: Erlbaum.

Brown, R., & Lenneberg, E. (1954). A study in language and cognition. *Journal of Abnormal and Social Psychology, 49,* 454–462.

Carroll, J. B., & Casagrande, J. B. (1958). The function of language classifications in behavior. In E. E. Maccoby, T. M. Newcomb, & E. L. Hartley (Eds.), *Readings in social psychology* (pp. 18–31). New York: Henry Holt.

Clarke, M., Losoff, A., McCracken, M., & Rood, D. (1984). Linguistic relativity and sex/gender studies: Epistemological and methodological considerations. *Language Learning, 34,* 47–67.

Clarke, M. A., Losoff, A., McCracken, M. D., & Still, J. (1981). Gender perception in Arabic and English. *Language Learning, 31,* 159–169.

Davidoff, J., Davies, I., & Robeson, D. (1999). Color categories in a stone-age tribe. *Nature, 398,* 203–204.

Dunn, L., & Dunn, L. (1981). *Peabody Picture Vocabulary Test—Revised.* Circle Pines, MN: American Guidance Service.

Dunn, L., & Dunn, L. (1986). *Test de vocabulario en Imagenes Peabody* [Peabody Picture Vocabulary Test]. Circle Pines, MN: American Guidance Service.

Elman, J. L. (1991). *Incremental learning, or the importance of starting small* (Tech. Rep. 9101). La Jolla: University of California, San Diego, Center for Research in Language.

Ervin, S. M. (1962). The connotations of gender. *Word, 18,* 249–261.

Fodor, J. (1983). *The modularity of mind.* Cambridge, MA: MIT Press.

Fox, A. (1992). *The structure of German.* Oxford, England: Clarendon Press.

Furth, H. (1966). *Thinking without language.* New York: Free Press.

Gelman, R., & Gallistel, R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Guiora, A., Beit-Halachmi, B., Fried, R., & Yoder, C. (1983). Language environment and gender identity attainment. *Language Learning, 32,* 289–304.

Heider, E. R. (1972). Universals in color naming and memory. *Journal of Experimental Psychology, 93,* 10–20.

Heider, E. R., & Oliver, D. C. (1972). The structure of color space in naming and memory for two languages. *Cognitive Psychology, 3,* 337–354.

Karmiloff-Smith, A. (1979). *A functional approach to child language: A study of determiners and reference.* Cambridge, England: Cambridge University Press.

Khosroshahi, F. (1989). Penguins don't care, but women do: A social identity analysis of a Whorfian problem. *Language and Society, 18,* 505–525.

Leinbach, M. D., Hort, B. E., & Fagot, B. (1997). Bears are for boys: Metaphorical associations in young children's gender stereotypes. *Cognitive Development, 12,* 107–130.

Levinson, S. L. (1996). Frames of reference and Molyneux's question: Crosslinguistic evidence. In P. Bloom, M. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and space* (pp. 109–169). Cambridge, MA: MIT Press.

Lucy, J. (1992). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis.* Cambridge, England: Cambridge University Press.

Luria, A. R. (1961). *The role of speech in the regulation of normal and abnormal behavior.* New York: Liveright.

Marcus, G. F. (1998). Can connectionism save constructivism? *Cognition, 66,* 153–182.

Martinez, I., & Shatz, M. (1996). Linguistic influences on categorization in preschool children: A crosslinguistic study. *Journal of Child Language, 23,* 529–545.

Miller, K., Smith, C., Zhu, J., & Zhang, H. (1995). Preschool origins of cross-national differences in mathematical competence: The role of number-naming systems. *Psychological Science, 6,* 56–60.

Mullen, M. K. (1990). Children's classifications of nature and artifact pictures into female and male categories. *Sex Roles, 23,* 577–587.

Munnich, E., & Landau, B. (1997, November). *Universals of spatial representation in language and memory.* Poster session presented at the 38th Annual Meeting of the Psychonomic Society, Philadelphia, PA.

Pérez-Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language, 18,* 571–590.

Plunkett, K., & Elman, J. L. (1997). *Exercises in rethinking innateness: A handbook for connectionist simulations.* Cambridge, MA: MIT Press.

Plunkett, K., Sinha, C., Møller, M. F., & Strandsby, O. (1992). Symbol grounding of the emergence of symbols? Vocabulary growth in children and a connectionist net. *Connection Science, 4,* 293–312.

Sapir, E. (1912). Language and environment. *American Anthropologist,* n.s., 226–242.

Sera, M. D., Bales, D., & del Castillo Pintado, J. (1997). "Ser" helps speakers of Spanish identify real properties. *Child Development, 68,* 820–831.

Sera, M. D., Berge, C., & 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., Reittinger, E., & del Castillo Pintado, J. (1991). Developing definitions of objects and events in English and Spanish speakers. *Cognitive Development, 6,* 119–142.

Soja, N., Carey, S., & Spelke, E. (1991). Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition, 45,* 101–107.

Terrell, T., & Salgués De Cargill, M. (1979). *Linguística aplicada* [Applied linguistics]. New York: Wiley.

Wetman, J. M., & Hauss, W. (1969). *A spoken word count: Children—Ages 5, 6 and 7.* Chicago: Science Research Associates.

Whorf, B. L. (1956). *Language, thought, and reality.* Cambridge, MA: MIT Press.

Appendix

Instructions Given to Participants in Spanish and French in Study 1 and in German in Study 2

Spanish Instructions

Estamos pensando hacer una pelicula en la que algunos objectos comunes tienen vida y cantan y bailan. Ahora voy a enseñarle una serie de dibujos y quiero que me diga si piensa que la cosa en el dibujo debe tener una voz de un hombre o una voz de una mujer. Muy bien, aquí está el dibujo numero uno [the experimenter would then show one picture to the participant, and for half of the participants the experimenter labeled the picture], esta cosa debe tener voz de un hombre o la voz de una mujer en la pelicula?

French Instructions

Je pense faire un nouveau film dans lequel des objets de la vie de tous les jours prendraient vie. Donc, dans ce film les objets seront capables de parler, de chanter, et de danser. Imagine que tu seras le metteur en scène de ce film. Je vais te montrer une série d'images reprèsentant diverses choses qui prendraient vie dans ce film. J'aimerais que tu me dises, si tu penses que l'objet décrit par l'image devrait être représenté par une voix d'homme (ou de garçon) ou par une voix de femme (ou de fille). Bon! Voici l'image numéro un [the experimenter would then show one picture to the participant, and for half of the participants the experimenter labeled the picture], est-ce que ceci devrait être représenté par une voix de femme ou par une voix d'homme dans le film?

German Instructions

Ich möchte einen Film drehen, in dem ganz normale Sache lebendig werden sollen. Ich werde dir eine Reihe von Bilder diese Dinge zeigen und möchte dich bitten, zu sagen, ob die in den Film eine Männerstimme oder eine Frauenstimme kriegen sollen. Hast du das verstanden? Sollen wir anfangen? Hier ist Nummer Eins [the experimenter would then show one picture to the participant, and for half of the participants the experimenter labeled the picture], Männerstimme oder Frauenstimme?

Received May 31, 2000
Revision received May 3, 2002
Accepted May 6, 2002 ■