

Research Article

Hannah Webber Locklear and Jeff Parker*

Cognitive effects of grammatical gender in L2 acquisition of Spanish: Replicability and reliability of object categorization

<https://doi.org/10.1515/opli-2025-0058>

received September 17, 2024; accepted May 20, 2025

Abstract: This study explores the cognitive effects of L2 Spanish acquisition on L2 speakers' categorization of nouns. We compared monolingual English speakers, L2 Spanish speakers who learned Spanish abroad, and L2 Spanish speakers who learned Spanish in the United States in a task where they assigned adjectives that are stereotypically associated with males or females to nouns of different types. We find effects for both group and noun type; however, contrary to expectations, monolingual English speaker responses were statistically more likely to exhibit a gender-congruence effect than L2 Spanish speaker responses in two groups of nouns, and there was no difference in gender-congruence between the responses of the two L2 Spanish groups. Our results suggest that a speaker's linguistic knowledge of a language with grammatical gender can affect how they categorize nouns in their native language, but that the direction and strength of the effect may differ across groups and tasks. These results contribute to the ongoing discussion on linguistic relativity and the extent to which one's linguistic experience influences their perception of the world, including whether learning a second language always increases gendered effects and to what extent such effects are reliable across experimental tasks.

Keywords: linguistic relativity, Sapir-Whorf, object categorization, language acquisition

Linguistic relativity, also known as the Sapir-Whorf hypothesis,¹ is a theory about the potential connection between how we think and the language(s) we speak (Kay and Kempton 1984). Linguistic relativity covers a range of ideas that differ in both the strength of connection between language and thought, and the type of interaction they exhibit (Lucy 2016, Thierry 2016, for reviews). Linguistic determinism, the idea that linguistic knowledge somehow limits or determines how an individual thinks, is widely rejected. A weaker version of this idea, that language in some way influences how we think, encompasses various versions of linguistic relativity. Since its emergence into mainstream linguistics in the twentieth century, linguistic relativity has been a popular topic of debate among linguistic scholars, both because it is controversial and because of the various ways in which a connection between linguistic structure and cognition can take shape. Linguistic relativity has been investigated in a variety of domains, including the domain of color (e.g., Kay and McDaniel 1978, Davies and Corbett 1997, Regier and Kay 2009, Cibelli et al. 2016), spatial reasoning (e.g., Calderon et al. 2019,

¹ This hypothesis is named after American linguists Edward Sapir and his student, Benjamin Whorf. However, earlier evidence of the hypothesis has been traced back to a German polymath, Wilhelm von Humboldt, from the eighteenth century (Pütz and Verspoor 2000).

* **Corresponding author: Jeff Parker**, Department of Linguistics, Brigham Young University, Provo, United States, e-mail: jeff_parker@byu.edu

Hannah Webber Locklear: Department of Linguistics, Brigham Young University, Provo, United States

Li and Gleitman 2002, Tseng et al. 2016), number marking (e.g., Athanasopoulos 2007, Marian and Kaushanskaya 2005, Zhang and Schmitt 1998), and grammatical gender (see references and discussion below).

Within the domain of gender, there is increasing evidence that grammatical gender affects speakers' categorization of objects. However, there remain various reasons to doubt the strength of the connection between the two and the ways in which they may or may not interact.

A growing number of studies find a connection between the grammatical gender of inanimate objects and the way in which speakers categorize them (e.g., Almutrafi 2015, Belacchi and Cubelli 2012, Beller et al. 2015, Bender et al. 2011, 2016, Lambelet 2016, Pavlidou and Alvanoudi 2018, Phillips and Boroditsky 2003, Vernich et al. 2017). Many such studies that find effects of gender on categorization compare groups of speakers with different native languages. While the evidence is strong, the connection to grammatical gender is not always certain; speakers of different languages may also differ in other ways, e.g., cultural traits. In efforts to remedy this potential challenge to the validity of results, scholars have focused on whether learning a second language affects the way in which speakers categorize objects, making comparisons across groups more clearly related to linguistic knowledge and not other factors (e.g., Kurinski et al. 2016, Kurinski and Sera 2011, Vernich 2017).

Support for the importance of grammatical gender also hinges on the reliability of studies that employ similar experimental methodologies. Various methodological factors are potentially important in experiments investigating linguistic relativity, leading to a literature that provides a mixed picture about the validity of various interpretations of the theory. In an overview of 43 studies on grammatical gender, Samuel et al. (2019) show that differences in experimental design, stimulus selection, language structures, and participant demographics can affect the likelihood of getting positive, mixed or no support for linguistic relativity. Among other things, they show that even fairly small differences in task, e.g., assigning voice/sex to an inanimate object vs judging properties of an inanimate object, affect the likelihood of getting results that support linguistic relativity. The fact that similar tasks, i.e., ones in which participants classify inanimate objects in some way, provide mixed or even conflicting results, brings up questions about the type and strength of connection between grammatical gender and cognitive processes.

For most scholars, mixed results suggest that the connection between language and cognition is somehow dependent on various factors, which leads them to consider questions about what conditions and traits of participants are most likely to produce effects attributable to linguistic relativity. This study is a continuation of the existing literature investigating different conditions and populations of speakers to better understand the ways in which linguistic knowledge does and does not affect cognition. In this article, we present results from an empirical study on gender assignment with monolingual English speakers and two groups of L2 Spanish speakers, one that learned Spanish in the US and one that learned Spanish abroad. Based on related tasks with similar groups of participants, we might expect differences between monolingual English speakers and L2 Spanish speakers. In addition, we might question whether the location of L2 learning might affect speakers' gender categorization. If language is the sole factor in affecting speakers' categorization and not other aspects of language learning, e.g., cultural experience, then we should not expect a difference between groups of speakers who learned their Spanish in different places. Our study provides two unexpected results: first, monolingual English speakers and L2 Spanish speakers differ, but not in the expected direction, and second, the two groups of L2 Spanish speakers do not differ in their gender assignment. Both results suggest we should be cautious in how we interpret results on linguistic relativity. At the same time, we do find an effect attributable to language experience, which can be interpreted as support for linguistic relativity, albeit not in the direction expected.

1 Background

1.1 Grammatical gender and its cognitive effects

In this article, we focus on the potential cognitive effects of language experience in the domain of grammatical gender. Grammatical gender is a type of noun class system in which nouns within a language are classified

based on how other words, e.g., adjectives, agree with them. For example, Spanish nouns are either feminine or masculine depending on the form of associated adjectives (Beatty-Martínez and Dussias 2019). The Spanish word for ‘beautiful’ differs in form depending on the noun that it is describing: *la luna es bonita* ‘the moon is beautiful’ vs *el sol es bonito* ‘the sun is beautiful’. Grammatical gender, thus, provides good conditions to test relativity because it is (1) part of language structure (unlike differences in color terms, for example), (2) unnecessary for communication (some languages do not have grammatical gender), and (3) differs across languages (the noun ‘bed’ is masculine in Italian but feminine in Spanish).

A large set of literature provides evidence that grammatical gender can affect (or ‘interfere’ in the words of Vernich (2017)) speakers’ categorization of inanimate objects (e.g., Almutrafi 2015, Belacchi and Cubelli 2012, Beller et al. 2015, Bender et al. 2011, 2016, Lambelet 2016, Pavlidou and Alvanoudi 2018, Phillips and Boroditsky 2003, Vernich et al. 2017). Effects have been found in speakers of different ages (e.g., children in Belacchi and Cubelli 2012, Flaherty 2001, Saalbach et al. 2012, Sera et al. 2002, Sera et al. 1994, and adults in Belacchi and Cubelli 2012, Beller et al. 2015, Bender et al. 2011, Kurinski et al. 2016, Lambelet 2016) and a variety of tasks, including voice-attribution (e.g., Athanasopoulos and Boutonnet 2016, Kurinski et al. 2016, Vernich 2017), sex assignment (e.g., Belacchi and Cubelli 2012, Pavlidou and Alvanoudi 2018, Sera et al. 1994), properties judgment (e.g., Flaherty 2001, Imai et al. 2014; Saalbach et al. 2012), among others.

Even with this increasing body of evidence for the effect of grammatical gender in categorization, the validity of the link between grammatical gender and effects in cognitive tasks is not unquestioned because differences between groups of speakers may not be solely influenced by grammatical gender. Other differences between the groups in question, like cultural factors, may also play an important role. In order to disentangle the relationship between culture and language, one must single out the specific property of language to be scrutinized and exclude additional and/or alternative cultural factors that could play a role in explaining the cognitive phenomena. Doing this is complicated because most studies involve a cross-linguistic method which compares “two grossly separate groups of speakers” (Beller et al. 2015, 333). The same can be said when comparing speakers of languages with grammatical gender systems to those without – it is difficult to prove that cognitive behavior is not influenced by cultural differences in such a case.

1.2 Validity and cognitive effects of grammatical gender in L2 Spanish acquisition

Recent studies have addressed this issue of validity by studying the cognitive effects of *acquiring* grammatical gender. This allows for a comparison of participants from the same culture: native monolingual speakers of a language without grammatical gender vs native speakers of the same language who also speak a second language with grammatical gender. In fact, over the past two decades, linguists have conducted experiments comparing monolingual English speakers to native English speakers who also speak a grammatically gendered second language (e.g., Forbes et al. 2008, Kousta et al. 2008, Kurinski and Sera 2011). These studies explore the concept of linguistic relativity, but they also avoid the issue of cultural differences among participants.

Kurinski and Sera (2011) focused on the manner in which the acquisition of Spanish in native English-speaking adults affected their categorization of inanimate objects. A total of 50 students of beginning, university-level Spanish courses completed two tasks involving grammatical gender. The first task measured the participants’ acquisition of grammatical gender. In a classroom setting, participants were presented with images of nouns and asked to provide the Spanish words for them (including the definite articles, which show gender agreement). The second task measured the participants’ categorization of inanimate objects. The participants were asked to assign a male or female voice to the same images of nouns used in the first task. The nouns included 8 control items (nouns with biological sex: four male and four female), and the remaining 40 nouns were divided into 4 groups of 10 by the following categories: 1) artificially occurring feminine nouns, 2) artificially occurring masculine nouns, 3) naturally occurring feminine nouns, and 4) naturally occurring masculine nouns (where naturally occurring refers to nouns produced by the natural world, and artificially occurring refers to human-made nouns). Kurinski and Sera (2011) distinguished between naturally and artificially occurring nouns because of a study by Sera et al. (1994), which found that monolingual English speakers

were more likely to classify natural objects as feminine and artificial objects as masculine. The control items ensured that the participants actually matched male voices to males and female voices to females. Finally, the participants completed both tasks four times throughout the academic year. The data showed that change occurred after just 10 weeks of Spanish instruction. As participants' understanding of grammatical gender increased, they were more likely to assign male voices to masculine nouns and female voices to feminine nouns. Specifically, participants most often matched the male voice to artificially occurring masculine objects and the female voice to naturally occurring feminine objects. Kurinski and Sera's (2011) study suggested that acquiring Spanish as a second language impacted participants' categorization of inanimate objects.

A similar study sought to explore the cognitive effects of acquiring grammatical gender by comparing English monolinguals to L2 Spanish speakers whose native language was English. Kaushanskaya and Smith (2016) assembled a list of 22 English nouns, 11 of which had masculine Spanish translations and 11 of which had feminine Spanish translations. The nouns were matched with proper names, taking note to avoid phonological similarities. The proper names were previously proven to be statistically considered either male or female names, and the pairings were then divided into two groups: (1) gender-congruent pairs (the Spanish translation's gender matching the biological sex of the referent) and (2) gender-incongruent pairs (the Spanish translation's gender not matching the biological sex of the referent). English monolinguals, low Spanish-exposure bilinguals, and high Spanish-exposure bilinguals were presented with the pairings in random order and subsequently tasked with recalling the names matched with each noun. Additionally, a small control group of native Spanish speakers was tested to ensure that the procedure would actually bring about the predicted grammatical gender effects, and "the data from this group showed a congruence effect, with gender-congruent proper names retrieved more successfully at testing" (p. 35). Kaushanskaya and Smith (2016) then averaged the correctly-recalled gender-congruent proper names vs the correctly-recalled gender-incongruent proper names, using paired-samples *t*-tests to compare the three groups' performance in accurately recalling the gender-congruent vs gender-incongruent condition. The tests revealed that the English monolinguals and low Spanish-exposure bilinguals were equally accurate at recalling the gender-congruent proper names as the gender-incongruent proper names. However, the high Spanish-exposure bilinguals experienced a significant gender-congruence effect, retrieving the gender-congruent proper names more accurately than the gender-incongruent names. Such results suggested that higher exposure to second language learning can have significant cognitive effects.

1.3 Reliability and cognitive effects of grammatical gender in L2 Spanish acquisition

While studies on language acquisition and grammatical gender provide strong evidence for a link between grammatical gender and cognitive tasks, the evidence is not consistently replicated across different types of languages and across different types of experimental tasks.

Vernich (2017) investigated how speakers of Lithuanian, a language with two grammatical genders, exhibited different patterns of categorization of inanimate objects after learning languages with different gender systems: English (no gender), Russian (three genders), German (three genders and gender-marking articles), and Italian (two genders and gender-marking articles). They found significant differences between the groups of speakers in a sex-assignment task in which speakers assigned a sex (male or female) to inanimate objects. They found that Lithuanian speakers who had learned a second language with a gender system were less likely to assign a sex that matches the Lithuanian gender of inanimate objects. This finding further substantiates a link between gender associations and (second) language knowledge. Interestingly, however, the differences between groups did not correspond to the gender system of the second language. They suggest that "learning a foreign language weakens the strength of the link between each item and its gender" (p. 418). Their results replicate the effect of second language acquisition on grammatical gender and object categorization, but do not replicate the direction of the effect. In studies like Kurinski and Sera (2011), the effects corresponded with gender patterns in the second language, whereas for speakers of a gendered language like Lithuanian, they did not, making the relationship between gender in a second language and object categorization more complex than previously thought.

A second issue of reliability within studies on the effects of grammatical gender is that results are not always replicated, especially across different types of experimental tasks. As mentioned above, Samuel et al. (2019) show that differences in experimental design, stimulus selection, language structures, and participant demographics can affect the likelihood of getting positive, mixed, or no support for linguistic relativity. This leads to questions of how reliable the effects of linguistic gender are in two ways. First, some studies with the same experimental task may find contradictory results (see Boroditsky et al. 2003 vs Mician et al. 2014 for an instance with gender, and Boroditsky 2001 vs Chen 2007 and January and Kako 2007 for an instance in the domain of conceptualization of time). Even if such differences can be attributed to other differences in the studies, e.g., languages involved or speaker attributes, the lack of full replicability indicates that the effects of grammatical gender depend on other factors. Second, studies that employ different experimental tasks differ widely in terms of whether they find support, and in instances where they do find support, how strongly their results support grammatical gender as affecting cognition. Samuel et al. (2019) find that tasks in which participants explicitly assign a voice/sex to inanimate objects are more likely to find results that support relativity (e.g., Almutrafi 2015, Athanasopoulos and Boutonnet 2016, Belacchi and Cubelli 2012, Beller et al. 2015, Bender et al. 2016, Kurinski et al. 2016, Pavlidou and Alvanoudi 2018, Ramos and Roberson 2011, Sera et al. 1994, 2002, Vernich 2017, Vernich et al. 2017) than tasks in which sex is not explicitly mentioned, e.g., when participants are asked to judge the properties of the inanimate objects that are gendered in the language they speak (e.g., Flaherty 2001, Imai et al. 2014, Landor 2014, Mician et al. 2014, Montefinese et al. 2019, Semenuks et al. 2017). The fact that similar tasks, i.e., classifying inanimate objects in some way, provide mixed or even conflicting results, brings up questions about the type and strength of connection between grammatical gender and cognitive processes. Such a lack in replication across studies and task types suggest that the relationship between gender and cognition is complex and interacts with a variety of factors. In this article, we contribute to the ongoing discussion of how this interaction plays out in different tasks and language (learning) situations to better our understanding of what contexts are most likely to reflect an interaction between language structure and cognition.

2 Research questions

In this article, we investigate the following questions:

- (1) To what extent do monolingual English speakers and L2 Spanish speakers differ in how they associate gendered adjectives with nouns of different types?
- (2) To what extent do L2 Spanish speakers who learned their Spanish in a Spanish-speaking country differ from L2 Spanish speakers who learned their Spanish in the United States in terms of how they associate gendered adjectives with nouns of different types?

While similar experiments have been conducted previously, this study is unique in three ways: 1) we employ a novel task in which participants were presented with nouns of different types and asked which adjectives among a set best described the nouns; (2) we compare two distinct groups of second language speakers, participants who learned their Spanish while living in a Spanish-speaking country and participants who learned Spanish primarily in the USA; and (3) all the participants shared similar demographic details, including being born in the United States and members of the same religion, The Church of Jesus Christ of Latter-day Saints (LDS).

3 Methodology

3.1 Participants

Students were recruited from two courses taught at Brigham Young University: SPAN 321 (Third Year Reading, Grammar, and Culture) and WRTG 150 (Writing and Rhetoric) with the assistance of the course professors.

These courses were chosen to provide a monolingual group (WRTG 150) and two groups of L2 Spanish speakers (both from SPAN 321) from the same population of BYU students. The university recommends SPAN 321 as the first Spanish class for students who learned Spanish on their LDS missions. LDS missionaries leave their homes for 18–24 months, during which time they attend an intensive L2 training center for up to 2 months before departing for their country of assignment, which is not chosen by the missionary but rather the leadership of the LDS Church. Once they arrive in the area in which their target language is spoken, missionaries are assigned a fellow missionary companion with whom they work constantly. As a part of the rigorous schedule, missionaries dedicate at least 1 h to official language study each day. The rest of the time is spent meeting, teaching, and serving the people in the area to which the missionaries are assigned. There is relative uniformity in age, mother tongue, and L2 motivation among LDS missionaries (Hansen 2012). The Spanish department at BYU suggests that students who register for SPAN 321 can speak, write, read, and listen at an Advanced Low level, according to the ACTFL Oral Proficiency Interview scale (ACTFL). In order to further measure Spanish language proficiency, participants self-evaluated their language abilities as a part of the survey questions using the Oral Proficiency Interview scale. Self-evaluation of language proficiency is not as objective as external measures of proficiency; however, the validity of self-assessment has been shown to be quite high (see discussion in Oscarson 1989) and self-assessment has been found to significantly correlate with language performance in meta-analyses over dozens of studies including over 68,000 students (Li and Zhang 2021). In addition to evaluating their language proficiency, participants reported whether or not they served an LDS mission and if they served abroad or in the United States. All participants were entered into a drawing from which six were randomly selected to receive \$50 Visa gift cards for their participation.

In total, responses from 149 participants were analyzed: 52 monolingual English speakers, 52 L2 Spanish speakers who served their full LDS mission abroad (36 male, 17 female), and 45 L2 Spanish speakers who did not serve their full mission abroad (15 male, 30 female). Of the 45 L2 Spanish speakers that did not serve their full mission abroad, 27 served their entire mission in the United States, whereas 18 spent some of their service abroad and some in the United States. Due to an error in the survey design, gender and age were not collected for the 52 monolingual English speakers. However, based on the responses from the two subgroups of L2 Spanish speakers, most (if not all) participants likely fell within the age range of 18–24 years old because all but one of the 98 L2 Spanish speakers reported this age range. All 149 participants were members of the Church of Jesus Christ of LDS, making them a relatively homogenous group in terms of religio-cultural background.

3.2 Survey design

Using a similar approach to Kurinski and Sera (2011), a Qualtrics survey was created in which participants were tasked with pairing a list of adjectives typically associated with males or females (Table 1) with a list of ten nouns referring to males and females, ten naturally occurring nouns with feminine Spanish translations, ten naturally occurring nouns with masculine Spanish translations, ten artificially occurring nouns with

Table 1: List of adjectives from Williams and Bennett (1975) used in the survey

Male-associated adjectives		Female-associated adjectives	
Aggressive	Robust	Dependent	Sensitive
Assertive	Self-confident	Dreamy	Sentimental
Autocratic	Severe	Emotional	Soft-hearted
Daring	Steady	Excitable	Sophisticated
Dominant	Stern	Flirtatious	Submissive
Forceful	Strong	High-strung	Talkative
Handsome	Tough	Meek	Unexcitable
Jolly	Unemotional	Mild	

feminine Spanish translations nouns, and ten artificially occurring nouns with masculine Spanish translations (Table 2). Adjectives were selected from a list developed by Williams and Bennett (1975). Williams and Bennett's (1975) study presented participants (50 male and 50 female Euro-American college students) with a list of 300 adjectives commonly used to describe people. The participants were asked to select whether each adjective was more regularly associated with men or women. Williams and Bennett (1975) provide a list of adjectives that were selected by at least $\frac{3}{4}$ of the male and female participants as typically associated with either males or females. In addition, Williams and Bennett classify the adjectives as highly favorable, highly unfavorable, or neutral (based on responses from a different study). For our survey, we use the neutral adjectives that were strongly associated with males and females. Table 1 shows a complete list of adjectives used. We recognize that associations with these adjectives may shift over time; however, we still chose this list because it was empirically based and attempts to avoid the potentially more problematic confound of adjectives that have strong negative or positive connotations.

The survey was conducted entirely in English for all groups. In the survey, adjectives were presented in a random order in pairs of four per question: two male-associated adjectives and two female-associated adjectives. Instead of asking participants to match the nouns with a gendered voice, the survey asked them to select

Table 2: List of nouns based on Kurinski and Sera (2011)

Nouns with biological sex			
Feminine		Masculine	
woman	<i>la mujer</i>	man	<i>el hombre</i>
girl	<i>la chica</i>	boy	<i>el chico</i>
mother	<i>la madre</i>	father	<i>el padre</i>
grandmother	<i>la abuela</i>	grandfather	<i>el abuelo</i>
sister	<i>la hermana</i>	brother	<i>el hermano</i>
Artificially Occurring Nouns			
Feminine		Masculine	
bed	<i>la cama</i>	airplane	<i>el avión</i>
bicycle	<i>la bicicleta</i>	book	<i>el libro</i>
camera	<i>la cámara</i>	building	<i>el edificio</i>
chair	<i>la silla</i>	car	<i>el carro</i>
church	<i>la iglesia</i>	piano	<i>el piano</i>
cup	<i>la taza</i>	plate	<i>el plato</i>
guitar	<i>la guitarra</i>	ship	<i>el barco</i>
house	<i>la casa</i>	shoe	<i>el zapato</i>
school	<i>la escuela</i>	stoplight	<i>el semáforo</i>
shirt	<i>la camisa</i>	train	<i>el tren</i>
Naturally Occurring Nouns			
Feminine		Masculine	
earth	<i>la tierra</i>	air	<i>el aire</i>
flower	<i>la flor</i>	fire	<i>el fuego</i>
fruit	<i>la fruta</i>	forest	<i>el bosque</i>
moon	<i>la luna</i>	lake	<i>el lago</i>
mountain	<i>la montaña</i>	planet	<i>el planeta</i>
sand	<i>la arena</i>	sea	<i>el mar</i>
star	<i>la estrella</i>	sun	<i>el sol</i>
stone	<i>la piedra</i>	tree	<i>el árbol</i>
waterfall	<i>la cascada</i>	universe	<i>el universo</i>
wood	<i>la madera</i>	wind	<i>el viento</i>

We are working to write a children's story in which the following nouns will act as characters. Some of them are inanimate and will come to life in the story.

Please match each noun with an adjective that you feel best represents the noun, if it were to be used as one of the characters in the story.

Figure 1: Survey explanation.

an adjective that they felt most adequately described the nouns listed in Table 2, as mentioned previously (Figure 1). More specifically, participants were told the following.

The list of nouns was inspired by, but not identical to, Kurinski and Sera's (2011) list (see Appendix 1 for original) and was split into two equal lists. One half of the respondents received the first list of nouns, and the other half received the second list of nouns; both lists contained an equal number of nouns from each category and were presented in a random order for each participant. Nouns were separated into two lists in order to reduce the number of overall questions that each participant answered, therefore reducing the survey response time.

The survey was designed and carried out in accordance with the tenets of the Helsinki Declaration and was approved by the authors' institutional review board. Informed consent was obtained from all individuals included in this study. As a part of the survey, participants consented to participate, answered a series of basic demographic questions, and the L2 Spanish speakers reported or self-rated their language abilities (as previously mentioned). The survey lasted approximately 8–12 min, and it contained 25 experimental questions modeled after the sample question in Figure 2.

4 Results

To determine the extent to which factors affected speakers' assignment of adjectives to nouns, we fitted a logistic mixed-effect regression model using the lme4 (Bates et al. 2015) package in R (R Core Team 2020) with congruence (congruent or non-congruent) as the dependent variable and group (monolingual English speakers, L2 Spanish Abroad, L2 Spanish US),² noun type (nouns with biological sex, naturally occurring nouns, artificially occurring nouns) and gender of Spanish translation of the noun (masculine, feminine) and the interaction of group and noun type as independent variables. Random intercepts were included for items. Random intercepts were not included for participants because they resulted in overfitting the model. No random slopes were included in the model.

Grammatical gender did not exhibit a significant effect in the model ($t(46.015) = 0.408$, $SE = 0.067$, $p = 0.685$). Group and noun type exhibited significant main effects in the model. For group, monolingual speakers had greater congruence than L2 Spanish abroad ($t(3685.867) = -2.008$, $SE = 0.0397$, $p = 0.044$) and greater congruence than L2 Spanish US ($t(3684.954) = -2.807$, $SE = 0.0381$, $p = 0.005$). L2 Spanish abroad and L2 Spanish US did not exhibit a significant difference in congruence ($t(3684.244) = -0.688$, $SE = 0.039$, $p = 0.491$). For noun type, nouns with biological sex exhibit greater congruence than naturally occurring nouns ($t(54.043) = -2.027$, $SE = 0.009$, $p = 0.047$) but did not exhibit a significant difference from artificially occurring nouns ($t(54.070) = -0.778$, $SE = 0.096$, $p = 0.440$), and naturally occurring nouns did not exhibit a significant difference from artificially occurring nouns ($t(56.073) = 0.750$, $SE = 0.079$, $p = 0.456$).

² As noted above, 18 participants spent some of their LDS missions abroad and some of their missions in the United States. Because it was not clear if this group should be included with those who served their entire mission abroad, the analysis was run separately, once excluding the 18 participants and once including them with the missionaries who served their entire mission in the United States. In both analyses the significance of factors remained the same. The results presented include the 18 participants with those who served their entire mission in the United States.

Which adjective best describes the following noun:

star

dominant

self-confident

mild

sentimental

Figure 2: Survey sample question.

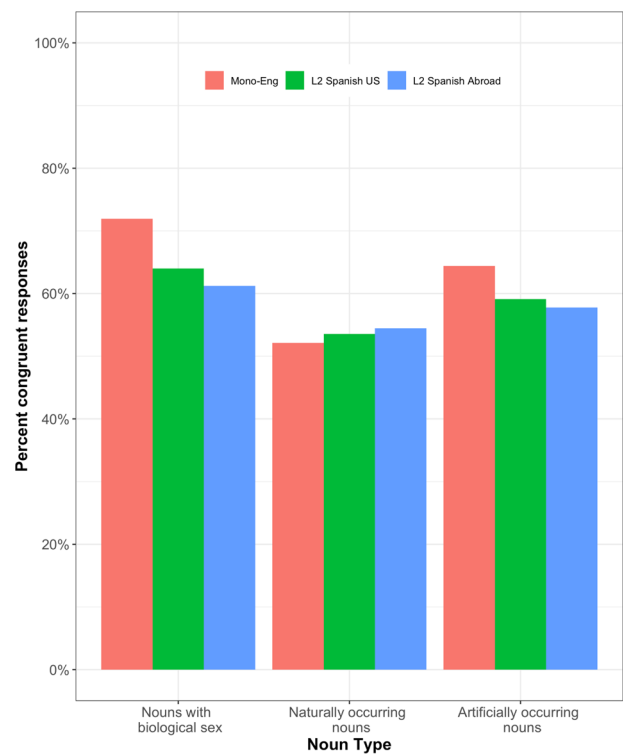


Figure 3: Percentage of congruent responses by noun type and group.

In addition to main effects, there was an interaction between group and noun type. See Figure 3 for a visual representation. To investigate the interaction, we did post-hoc pairwise comparisons within noun type using a pseudo-Bonferroni correction.³ The monolingual English and L2 Spanish Abroad groups exhibit significant differences for nouns with biological sex ($t(3684.9) = 2.807$, $SE = 0.0381$, $p = 0.005$) and fall just short of significance for artificially occurring nouns ($t(3684.9) = 2.807$, $SE = 0.0$, $p = 0.012$). No other groups exhibited

³ The three group by three noun-type interaction results in a large number of pairwise comparisons (36), many of which are uninformative because they cross both group and noun type. A normal Bonferroni correction would thus overpenalize. Instead, we use a pseudo-Bonferroni correction with an alpha value of 0.01 since we only look at only the nine within noun-type pairwise comparisons.

significant differences with noun type (all $p > 0.01$), illustrating that the extent to which groups differ in their congruence was not consistent across noun types.

To further investigate the interaction of noun type and group, we tested the extent to which responses from groups within noun types differed from what would be expected from chance. Given a task in which participants chose between four options, two of which are congruent and two of which are not congruent, not all responses will be more congruent than what we would expect if participants were randomly choosing among the given options. Interestingly, responses to naturally occurring nouns by the monolingual English speakers and L2 Spanish US groups were not significantly different than selecting answers by chance ($\chi^2(1) = 0.930$, $N = 520$, $p = 0.334$; $\chi^2(1) = 2.275$, $N = 450$, $p = 0.131$, respectively), though L2 Spanish Abroad group responses are significantly different from chance ($\chi^2(1) = 4.175$, $N = 529$, $p = 0.041$). In both other noun types (nouns with biological sex and artificially occurring nouns) all three groups chose congruent responses significantly more than non-congruent responses (all $p < 0.001$). These findings illustrate that not only did some groups exhibit greater congruence effects within some noun types, but not all groups within naturally occurring nouns exhibited evidence of congruence greater than chance.

5 Discussion and conclusions

Our results provide an interesting contribution to the nature of grammatical gender and linguistic relativity. On the one hand, we find a significant difference between L2 Spanish speakers and monolingual English speakers in terms of how they associate gendered adjectives with nouns. Such a difference is a classic type of evidence for linguistic relativity – two groups that differ in only their linguistic experience behave differently when performing a specific task in which grammatical gender is relevant. Furthermore, the location in which L2 Spanish speakers learned the language did not have any effect. If the difference between groups were due to something other than linguistic experience, e.g., cultural experience, we might expect a difference between the two groups of L2 Spanish speakers. The fact that such a difference was not present provides additional confidence in ascribing the difference between Spanish speakers and monolingual English speakers to linguistic experience instead of some other (potentially uninvestigated) confounding variable. In terms of validity, these results provide potential support for possible interpretations of linguistic relativity.

At the same time, the results also provide strong reasons to question our confidence in support for linguistic relativity. The fact that monolingual speakers had a gender congruence effect for nouns with biological sex is unsurprising. This effect validates that there is a connection between nouns with biological sex and the association of the adjectives used in the task with males and females. However, it is surprising that L2 Spanish speakers did not exhibit the same effects across groups of nouns as the monolinguals. The difference between monolingual English speakers and L2 Spanish speakers for nouns with biological sex was not in the expected direction – monolingual English speakers exhibited greater gender congruence than L2 Spanish speakers. In previous studies that investigated effects of gender assignment for L2 Spanish speakers, the effect was the opposite – L2 Spanish speakers exhibited greater congruence in gender assignment tasks when compared among difference proficiency levels, including monolingual English speakers (Kaushanskaya and Smith 2016), or following learners as they progressed in Spanish proficiency (Kurinski and Sera 2011). If there were a strong cognitive association between the grammatical gender of Spanish translations and the same nouns presented in English in these other tasks, we should expect that association to persist in tasks involving gendered adjectives, like the one presented in this article. One primary difference between our task and Kaushanskaya and Smith (2016) and Kurinski and Sera (2011) is that gender and natural sex were not explicitly mentioned in our study. Participants were not asked to assign a male or female voice to objects, making our task represent a less-direct association than previous tasks. This corresponds with the findings of Samuel et al. (2019) that tasks involving properties of objects rather than voice choice or sex assignment are much less likely to find (full or mixed) support for relativity. Our results, thus, add to the existing studies that suggest effects of gender are not consistent across experimental tasks. What is surprising in our results is not only that the results from other tasks are not replicated in our study, but that the largest differences, e.g., monolinguals vs L2 Spanish speakers in nouns with biological sex, are significant in the opposite direction as expected. This suggests that L2 Spanish acquisition affects speakers' perception of nouns with biological sex in that they do not

perceive them as inherently masculine or feminine as much as monolingual speakers do, or that their association of adjectives with males and females differs from monolingual speakers. In other words, L2 Spanish acquisition does seem to affect the interaction of gender and categorization in regard to adjective assignment. It is, of course, possible that the methods we employed had confounds we could not or did not directly control or investigate. The association of the adjectives could have changed over time, the gender of the participants (collected for some groups and not others) could have an effect, the country in which speakers learned their Spanish may have mattered, the self-assessment of learner's proficiency could be flawed, or there may have been some other unnoticed variable. However, none of those possible methodological challenges readily explain why there were differences between groups of speakers who differ specifically in terms of their linguistic experience. None of these variables should produce an effect of gender congruence for monolingual English speakers but not produce the same effect for L2 Spanish speakers, unless learning Spanish affected their categorization of some objects. One possible explanation, which we see as more motivated than potential methodological challenges, is that the arbitrary nature of gender in Spanish affects the association between nouns with biological sex and adjectives, even when the task is presented in English. This could explain why L2 speakers exhibit less gender-congruence than monolingual speakers, though it cannot fully explain why those differences do not persist across all noun types. This explanation is consistent with Vernich (2017), who investigated Lithuanian speakers who learned English, Russian, German, or Italian. Vernich found that knowledge of a second language with a gender system weakened the connection between nouns and their gender in Lithuanian, suggesting that "learning a foreign language does appear to interfere – even if only slightly – with the standard pattern exhibited by monolingual" (p. 432). This is particularly interesting because Vernich's participants were native speakers of a language with two genders, whereas our participants' native language does not have gender. Furthermore, Vernich used a task in which participants assigned sex (male or female) to inanimate objects, a type of task that is more likely (based on Samuel et al. 2019) to exhibit gender effects than our task, in which natural sex and gender were not explicitly mentioned. The fact that our results pattern similarly with a different set of language types and experimental tasks is interesting because it suggests that learning a language may weaken gendered associations in particular combinations of languages and task types.

Our results illustrate that not only can task type affect the strength of evidence for relativity, but it can also reverse the direction of the support altogether. In a counterintuitive way, this finding contradicts the expected, straightforward interpretation of linguistic relativity based on expectations of gendered effects in studies with different tasks, while at the same time providing some evidence for the general idea that one's linguistic experience can affect categorization in tasks involving inanimate nouns and nouns with biological sex. We should be cautious not to overstate the impact and direction of linguistic experience on cognitive tasks involving gender association. Our results, thus, suggest that the interaction of linguistic knowledge and cognition is nuanced in that the task type and linguistic experience affect both the significance and direction of effects that can be interpreted as evidence of linguistic relativity.

Acknowledgments: The authors are grateful for comments from anonymous reviewers.

Funding information: The authors gratefully acknowledge funding from the College of Humanities, Brigham Young University, for providing funding for participant incentives.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission. HWL designed the experiment in consultation with JP. HWL recruited participants and did the initial analysis and manuscript draft. JP did the statistical analysis, data visualization, and expanded the discussion.

Conflict of interest: The authors state no conflict of interest.

Data availability statement: The data sets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

- Almutrafi, Fatimah. 2015. *Language and Cognition: Effects of Grammatical Gender on the Categorisation of Objects*. PhD diss., Newcastle University.
- Athanasopoulos, Panos and Bastien Boutonnet. 2016. "Learning Grammatical Gender in a Second Language Changes Categorization of Inanimate Objects: Replications and New Evidence from English Learners of L2 French." In *Crosslinguistic Influence in Second Language Acquisition*, edited by Rafaela A. Alonso, 173–92. Blue Ridge Summit, PA: Multilingual Matters. doi: 10.21832/9781783094837-011.
- Athanasopoulos, Panos. 2007. "Interaction Between Grammatical Categories and Cognition in Bilinguals: The Role of Proficiency, Cultural Immersion, and Language of Instruction." *Language and Cognitive Processes* 22 (5): 689–99. doi: 10.1080/01690960601049347.
- Bates, Douglas, Martin Mächler, Ben Bolker, and Steven Walker. 2015. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software* 67 (1): 1–48. doi: 10.48550/arXiv.1406.5823.
- Beatty-Martínez, Alicia L. and Paola E. Dussias. 2019. "Revisiting Masculine and Feminine Grammatical Gender in Spanish: Linguistic, Psycholinguistic, and Neurolinguistic Evidence." *Frontiers in Psychology* 10: 751. doi: 10.3389/fpsyg.2019.00751.
- Belacchi, Carmen and Roberto Cubelli. 2012. "Implicit Knowledge of Grammatical Gender in Preschool Children." *Journal of Psycholinguistic Research* 41 (4): 295–310. doi: 10.1007/s10936-011-9194-y.
- Beller, Sieghard, Kjersti F. Brattebø, Kåre O. Lavik, Randi D. Reigstad, and Andrea Bender. 2015. "Culture or Language: What Drives Effects of Grammatical Gender?" *Cognitive Linguistics* 26 (2): 331–59. doi: 10.1515/cog-2014-0021.
- Bender, Andrea, Sieghard Beller, and Karl Christoph Klauer. 2011. "Grammatical Gender in German: A Case for Linguistic Relativity?" *Quarterly Journal of Experimental Psychology* 64 (9): 1821–35. doi: 10.1080/17470218.2011.582128.
- Bender, Andrea, Sieghard Beller, and Karl Christoph Klauer. 2016. "Crossing Grammar and Biology for Gender Categorisations: Investigating the Gender Congruency Effect in Generic Nouns for Animates." *Journal of Cognitive Psychology* 28 (5): 530–558. doi: 10.1080/20445911.2016.1148042.
- Boroditsky, Lera, Lauren Schmidt, and Webb Phillips. 2003. "Sex, Syntax, and Semantics." In *Language in Mind: Advances in the Study of Language and Thought*, edited by Dedre Gentner and Susan Goldin-Meadow, 61–79. Cambridge, MA: MIT Press.
- Boroditsky, Lera. 2001. "Does Language Shape Thought? Mandarin and English Speakers' Conceptions of Time." *Cognitive Psychology* 43 (1): 1–22. doi: 10.1006/cogp.2001.0748.
- Calderon, Emilio, Silvia De Pascale, and Eleni Adamou. 2019. "How to Speak 'Geocentric' in an 'Egocentric' Language: A Multimodal Study among Ngigua-Spanish Bilinguals and Spanish Monolinguals in a Rural Community of Mexico." *Language Sciences* 74: 24–46. doi: 10.1016/j.langsci.2019.04.001.
- Chen, Jia-Ying. 2007. "Do Chinese and English Speakers Think about Time Differently? Failure of Replicating Boroditsky (2001)." *Cognition* 104 (2): 427–36. doi: 10.1016/j.cognition.2006.09.012.
- Cibelli, Emily, Yang Xu, Joseph L. Austerweil, Thomas L. Griffiths, and Terry Regier. 2016. "The Sapir-Whorf Hypothesis and Probabilistic Inference: Evidence from the Domain of Color." *Plos One* 11 (7): e0158725. doi: 10.1371/journal.pone.0158725.
- Davies, Irené R. and Greville G. Corbett. 1997. "A Cross-Cultural Study of Colour Grouping: Evidence for Weak Linguistic Relativity." *British Journal of Psychology* 88 (3): 493–517. doi: 10.1111/j.2044-8295.1997.tb02653.x.
- Flaherty, Mary. 2001. "How a Language Gender System Creeps into Perception." *Journal of Cross-Cultural Psychology* 32 (1): 18–31. doi: 10.1177/0022022101032001005.
- Forbes, Joan N., Danielle Poulin-Dubois, Maricela R. Rivero, and Maria D. Sera. 2008. "Grammatical Gender Affects Bilinguals' Conceptual Gender: Implications for Linguistic Relativity and Decision Making." *The Open Applied Linguistics Journal* 1 (1): 68–76. doi: 10.2174/1874913500801010068.
- Hansen, Louise. 2012. *Second Language Acquisition Abroad: The LDS Missionary Experience*, vol. 45. Amsterdam: John Benjamins Publishing.
- Imai, Masae, Lisa Schalk, Holger Saalbach, and Hitoshi Okada. 2014. "All Giraffes Have Female-Specific Properties: Influence of Grammatical Gender on Deductive Reasoning About Sex-Specific Properties in German Speakers." *Cognitive Science* 38 (3): 514–36. doi: 10.1111/cogs.12074.
- January, Daniel and Elena Kako. 2007. "Re-evaluating Evidence for Linguistic Relativity: Reply to Boroditsky (2001)." *Cognition* 104 (2): 417–26. doi: 10.1016/j.cognition.2006.07.008.
- Kaushanskaya, Margarita and Sara Smith. 2016. "Do Grammatical-Gender Distinctions Learned in the Second Language Influence Native-Language Lexical Processing?" *International Journal of Bilingualism* 20 (1): 30–9. doi: 10.1177/1367006915576830.
- Kay, Paul and Chad K. McDaniel. 1978. "The Linguistic Significance of the Meanings of Basic Color Terms." *Language* 54 (3): 610–46. doi: 10.2307/412789.
- Kay, Paul and William Kempton. 1984. "What Is the Sapir-Whorf Hypothesis?" *American Anthropologist* 86 (1): 65–79. doi: 10.1525/aa.1984.86.1.02a00050.
- Kousta, Stavroula T., David P. Vinson, and Gabriella Vigliocco. 2008. "Investigating Linguistic Relativity Through Bilingualism: The Case of Grammatical Gender." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34 (4): 843. doi: 10.1037/0278-7393.34.4.843.
- Kurinski, Elise and Maria D. Sera. 2011. "Does Learning Spanish Grammatical Gender Change English-Speaking Adults' Categorization of Inanimate Objects?" *Bilingualism: Language and Cognition* 14 (2): 203–20. doi: 10.1017/S1366728910000179.
- Kurinski, Elise, Emily Jambor, and Maria D. Sera. 2016. "Spanish Grammatical Gender: Its Effects on Categorization in Native Hungarian Speakers." *International Journal of Bilingualism* 20 (1): 76–93. doi: 10.1177/1367006915576833.

- Lambelet, Anne. 2016. "Second Grammatical Gender System and Grammatical Gender-Linked Connotations in Adult Emergent Bilinguals with French as a Second Language." *International Journal of Bilingualism* 20 (1): 62–75. doi: 10.1177/1367006915576832.
- Landor, Rebecca. 2014. "Grammatical Categories and Cognition Across Five Languages: The Case of Grammatical Gender and Its Potential Effects on the Conceptualisation of Objects." PhD diss., Griffith University.
- Li, Min and Xiaoming Zhang. 2021. "A Meta-Analysis of Self-Assessment and Language Performance in Language Testing and Assessment." *Language Testing* 38 (2): 189–218.
- Li, Ping and Lila Gleitman. 2002. "Turning the Tables: Language and Spatial Reasoning." *Cognition* 83 (3): 265–94. doi: 10.1016/S0010-0277(02)00009-4.
- Lucy, John A. 2016. "Recent Advances in the Study of Linguistic Relativity in Historical Context: A Critical Assessment." *Language Learning* 66 (3): 487–515. doi: 10.1111/lang.12195.
- Marian, Viorica and Margarita Kaushanskaya. 2005. "Autobiographical Memory and Language in Bicultural Bilinguals." In *ISB4: Proceedings of the 4th International Symposium on Bilingualism*, edited by J. Cohen, K. T. McAlister, K. Rolstad, and Jeff MacSwan, 1478–86.
- Mickan, Annika, Michael Schiefke, and Anatol Stefanowitsch. 2014. "Key is a Llave is a Schlüssel: A Failure to Replicate an Experiment from Boroditsky et al. 2003." *Yearbook of the German Cognitive Linguistics Association* 2 (1): 39–50. doi: 10.1515/gcla-2014-0004.
- Montefinese, Marianna, Elisabetta Ambrosini, and Esa Roivainen. 2019. "No Grammatical Gender Effect on Affective Ratings: Evidence from Italian and German Languages." *Cognition and Emotion* 33 (4): 848–54. doi: 10.1080/02699931.2018.1483322.
- Oscarson, Mats. 1989. "Self-Assessment of Language Proficiency: Rationale and Applications." *Language Testing* 6 (1): 1–13.
- Pavlidou, Theano-Sofia and Angeliki Alvanoudi. 2018. "Conceptualizing the World as 'Female' or 'Male': Further Remarks on Grammatical Gender and Speakers' Cognition." In *Selected Papers on Theoretical and Applied Linguistics from ISTAL23*, edited by N. Topintzi, N. Lavidas, and M. Moumtzi. Thessaloniki: School of English, Aristotle University of Thessaloniki.
- Phillips, William and Lera Boroditsky. 2003. "Can Quirks of Grammar Affect the Way You Think? Grammatical Gender and Object Concepts." *Paper presented at the 25th Annual Meeting of the Cognitive Science Society, Boston, MA, August*.
- Pütz, Martin and Marjolijn H. Verspoor. 2000. *Explorations in Linguistic Relativity*. Amsterdam: John Benjamins Publishing Company. <http://ebookcentral.proquest.com/lib/byu/detail.action?docID=729091>.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. <https://cran.r-project.org/>.
- Ramos, Susana and Daniel Roberson. 2011. "What Constrains Grammatical Gender Effects on Semantic Judgements? Evidence from Portuguese." *Journal of Cognitive Psychology* 23: 102–11. doi: 10.1080/20445911.2011.466795.
- Regier, Terry and Paul Kay. 2009. "Language, Thought, and Color: Whorf Was Half Right." *Trends in Cognitive Sciences* 13 (10): 439–46. doi: 10.1016/j.tics.2009.07.001
- Saalbach, Holger, Masae Imai, and Lisa Schalk. 2012. "Grammatical Gender and Inferences about Biological Properties in German-Speaking Children." *Cognitive Science* 36: 1251–67. doi: 10.1111/j.1551-6709.2012.01251.x.
- Samuel, Stephanie, Gemma Cole, and Mark J. Eacott. 2019. "Grammatical Gender and Linguistic Relativity: A Systematic Review." *Psychonomic Bulletin & Review* 26 (6): 1767–86. doi: 10.3758/s13423-019-01652-3.
- Semenuks, Anna, William Phillips, Ioana Dalca, Chia-Chi Kim, and Lera Boroditsky. 2017. "Effects of Grammatical Gender on Object Description." In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, edited by G. Gunzelmann, A. Howes, T. Tenbrink, and E. Davelaar, 1060–5. Austin, TX: Cognitive Science Society.
- Sera, Maria D., Catherine A. Berge, and José del Castillo Pintado. 1994. "Grammatical and Conceptual Forces in the Attribution of Gender by English and Spanish Speakers." *Cognitive Development* 9 (3): 261–92. doi: 10.1016/0885-2014(94)90007-8.
- Sera, Maria D., Claire Elieff, Joan Forbes, Mary C. Burch, William Rodríguez, and Danielle P. Dubois. 2002. "When Language Affects Cognition and When It Does Not: An Analysis of Grammatical Gender and Classification." *Journal of Experimental Psychology: General* 131: 377–97. doi: 10.1037/0096-3445.131.3.377.
- Thierry, Guillaume. 2016. "Neurolinguistic Relativity: How Language Flexes Human Perception and Cognition." *Language Learning* 66: 690–713. doi: 10.1111/lang.12186.
- Tseng, Chia, Anna B. Carstensen, Terry Regier, and Yang Xu. 2016. "A Computational Investigation of the Sapir-Whorf Hypothesis: The Case of Spatial Relations." In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, edited by A. Papafragou, D. Grodner, and D. Mirman, 2231–6.
- Vernich, Lina, Rūta Argus, and Lina Kamandulytė-Merfeldienė. 2017. "Extending Research on the Influence of Grammatical Gender on Object Classification: A Cross-Linguistic Study Comparing Estonian, Italian and Lithuanian Native Speakers." *Eesti Rakenduslingvistika Ühingu Aastaraamat* 13: 223–40. doi: 10.5128/ERYa13.14.
- Vernich, Luca. 2017. "Does Learning a Foreign Language Affect Object Categorization in Native Speakers of a Language with Grammatical Gender? The Case of Lithuanian Speakers Learning Three Languages with Different Types of Gender Systems (Italian, Russian and German)." *International Journal of Bilingualism* 23 (2): 417–36. doi: 10.1177/1367006917728593.
- Williams, Joan E. and Susan M. Bennett. 1975. "The Definition of Sex Stereotypes via the Adjective Check List." *Sex Roles* 1 (4): 327–37. doi: 10.1007/BF00287224.
- Zhang, Shaoming and Bernd Schmitt. 1998. "Language-Dependent Classification: The Mental Representation of Classifiers in Cognition, Memory, and Ad Evaluations." *Journal of Experimental Psychology: Applied* 4 (4): 375. doi: 10.1037/1076-898X.4.4.375.

Appendix 1

Artificial		Natural	
Feminine	Masculine	Feminine	Masculine
<i>La cocina</i> “kitchen”	<i>El inodoro/baño</i> “toilet”	<i>La papa</i> “potato”	<i>El tomate</i> “tomato”
<i>La mesa</i> “table”	<i>El carro</i> “car”	<i>La lechuga</i> “lettuce”	<i>El océano</i> “ocean”
<i>La casa</i> “house”	<i>El autobús</i> “bus”	<i>La luna</i> “moon”	<i>El limón</i> “lemon”
<i>La cama</i> “bed”	<i>El piano</i> “piano”	<i>La montaña</i> “mountain”	<i>El arroz</i> “rice”
<i>La bicicleta</i> “bicycle”	<i>El libro</i> “book”	<i>La manzana</i> “apple”	<i>El parque</i> “park”
<i>La oficina</i> “office”	<i>El plato</i> “plate”	<i>La nariz</i> “nose”	<i>El sol</i> “sun”
<i>La iglesia</i> “church”	<i>El video</i> “video”	<i>La fresa</i> “strawberry”	<i>El maíz</i> “corn”
<i>La guitarra</i> “guitar”	<i>El avión</i> “airplane”	<i>La leche</i> “milk”	<i>El huevo</i> “egg”
<i>La carta</i> “letter”	<i>El teléfono</i> “telephone”	<i>La playa</i> “beach”	<i>El viento</i> “wind”
<i>La ropa</i> “clothes”	<i>El tren</i> “train”	<i>La oreja</i> “ear”	<i>El ojo</i> “eye”

Note: from Kurinski, Elise, and Maria D. Sera. 2011. “Does Learning Spanish Grammatical Gender Change English-Speaking Adults’ Categorization of Inanimate Objects?” *Bilingualism: Language and Cognition* 14 (2): 203–20. <https://doi.org/10.1017/S1366728910000179>.