

Noun canonicity in heritage speakers and monolingual speakers of Spanish

Canonicidad nominal en hablantes de herencia y hablantes monolingües de español

Abstract

The present investigation examines the role of Spanish noun gender-correlated endings when accessing gender agreement in two different linguistic populations: Spanish-English bilingual heritage speakers and monolingual speakers of Spanish. This study analyzed data from 34 monolingual speakers of Spanish from the Dominican Republic and 44 heritage speakers of Spanish born in the United States who completed a picture naming task in Experiment 1 (determiner-noun agreement) and a picture description task in Experiment 2 (noun-adjective agreement). Results found that canonicity, particularly overt gender marking cues, seemed to have a facilitatory effect for monolingual speakers as seen by significantly faster naming times with transparent nouns on both experiments RTs analysis. However, within the heritage speaker group, no canonicity effects were found on either experiments RTs analyses indicating a difference between the monolingual and bilingual group. There was, however, an effect of noun canonicity in the accuracy rates of both experiments in the monolingual and bilingual group, specifically, a facilitatory effect of transparency, consonant with other studies with heritage speakers and/or monolingual speakers of Spanish (Hur, Lopez and Sanchez 2020; Montrul, Davidson, De La Fuente and Foote 2014; Montrul, De La Fuente, Davidson and Foote 2013; Alarcón 2011; Montrul, Foote and Perpiñán 2008). The main effects of canonicity and frequency found in Experiment 1 monolingual's RTs analysis and the interaction between frequency and canonicity found in Experiment 2 strongly suggest that canonicity along with noun frequency (noun lexico-syntactic information) have a facilitatory effect in the gender agreement process. Therefore, these results corroborate the predictions of the Two-Route Hypothesis proposed by Gollan and Frost (2001) and psycholinguistic studies that have found a lexico-syntactic route and word-form route to gender access (Caffarra, Jansen and Barber 2014). Findings in the present study contribute to the understanding of different processing mechanisms in two adult populations and to our knowledge of the overall role of predictive nominal endings in gender agreement processes in both monolingual and heritage speakers of Spanish.

Autoría

DÁMARIS MAYANS

Colby College, United States

dmayans@colby.edu

<https://orcid.org/0000-0002-6785-9392>

Para citar este artículo:

Mayans, D. (2023). Noun canonicity in heritage speakers and monolingual speakers of Spanish, *ELUA*, 39, 63-84. <https://doi.org/10.14198/ELUA.21688>

Recibido: 13/01/2022

Aceptado: 20/10/2022

© 2023 Dámaris Mayans



Licencia: Este trabajo está sujeto a una licencia de Reconocimiento 4.0 Internacional de Creative Commons (CC BY 4.0). <https://creativecommons.org/licenses/by/4.0/>

Keywords:

canonicity; heritage speakers of Spanish; lexico-syntactic features; gender agreement

Resumen

La presente investigación examina el papel de los fonemas nominales predictivos de género en operaciones de concordancia de género en dos poblaciones lingüísticas diferentes: hablantes bilingües de herencia español-inglés y hablantes monolingües de español. Este estudio analizó datos de 34 hablantes monolingües de la República Dominicana y 44 hablantes de herencia de español nacidos en los Estados Unidos que completaron una tarea de denominación de imágenes en el Experimento 1 (concordancia entre determinante y sustantivo) y una tarea de descripción de imágenes en el Experimento 2 (concordancia entre sustantivo y adjetivo). Los análisis estadísticos de tiempo de reacción indicaron que la canonicidad tuvo un efecto facilitador para los hablantes monolingües ya que eran más rápidos con sustantivos transparentes que con sustantivos opacos. Dentro del grupo de hablantes de herencia, no se encontraron efectos de canonicidad en los análisis de los tiempos de reacción de ambos experimentos. No obstante, hubo un efecto de canonicidad en el análisis de las tasas de errores de ambos experimentos, específicamente, un efecto facilitador de la transparencia, en consonancia con otros estudios con hablantes de herencia del español y/o monolingües de español (Hur, Lopez y Sanchez 2020; Montrul, Davidson, De La Fuente y Foote 2014; Montrul, De La Fuente, Davidson y Foote 2013; Alarcón 2011; Montrul, Foote y Perpiñán 2008). El efecto principal de canonicidad y frecuencia en el Experimento 1 y la interacción entre frecuencia y canonicidad encontrada en el Experimento 2 sugieren que la canonicidad, particularmente las terminaciones nominales predictivas, y la frecuencia nominal (información léxico-sintáctica del sustantivo) tiene un efecto facilitador en el proceso de concordancia de género. Por lo tanto, estos resultados corroboran las predicciones de la Hipótesis de las Dos Rutas propuesta por Gollan y Frost (2001) y estudios psicolingüísticos que han encontrado una ruta léxico-sintáctica y una ruta que hace uso de los marcadores nominales (ortográficos o fonémicos) para el acceso al género (Caffarra, Jansen y Barber 2014). Los hallazgos en el presente estudio contribuyen a la comprensión de los diferentes mecanismos de procesamiento en dos poblaciones de hablantes y a nuestro conocimiento del papel general de las terminaciones nominales predictivas en los procesos de concordancia de género en hablantes de español monolingües y de herencia.

Palabras clave:

canonicidad; hablantes de herencia del español; características léxico-sintácticas; concordancia de género.

1. INTRODUCTION

Grammatical gender, unlike semantic gender, is an inherent lexico-syntactic property of nouns (Carroll 1989; Biran and Friedman 2012) that is determined in an arbitrary way; that is, the referent itself does not inherently determine its gender. Most nouns in Spanish have no relation between their grammatical gender and any concept of feminine or masculine properties in the real world (e.g., *la mesa*-_{N.FEM} “the table”, Harris 1991). In some languages, such as German, the gender of the noun is difficult to predict only from its overt gender marking cues. Spanish uses a binary system where all nouns are assigned to one of the two genders (Lloret and Viaplana

1998; Alarcos 1999). Unlike German, Spanish gender-correlated endings are mainly canonical or transparent; that is, most masculine nouns end in *-o* and most feminine nouns end in *-a* (Green 1988), representing 68.15% of all nouns (O'Rourke and Van Petten 2011). The rest of the nouns have ambiguous gender marking cues and are called non-canonical or opaque nouns. Irregular nouns (also considered non canonical nouns) are a small set of nouns that use canonical/transparent noun endings, *-a* and/or *-o*, but agree with the opposite gender. The present investigation does not include nouns with irregular gender marking. Below are some examples of transparent/canonical and opaque/non-canonical nouns.

Table 1. Transparent/opaque masculine and feminine nouns in Spanish

Transparent		Opaque	
Feminine -a ending	Masculine o — ending	Feminine -e or consonant ending	Masculine -e or consonant ending
cama 'bed'	vaso 'glass'	mente 'mind'	reloj 'watch'
lámpara 'lamp'	plato 'plate'	nariz 'nose'	sol 'sun'
planta 'plant'	zapato 'shoe'	cicatriz 'scar'	puente 'bridge'

Psycholinguistic research using different methodologies has shown varied results about the way speakers utilize gender word-form cues in language production and comprehension. Some studies, most of them administered to monolingual speakers, have shown that gender correlated endings impact gender processing (Bates *et al.* 1996; Taft and Meunier 1998). This view has been called the Reliable Cue Hypothesis (Gollan and Frost 2001). Other research has suggested that gender correlated endings are not necessary to retrieve gender (Badecker *et al.* 1995; Miozzo and Caramazza 1997; Vigliocco *et al.* 1997). Another hypothesis, the Two-Route Hypothesis (Gollan and Frost 2001), acknowledges the existence of two routes to access gender, one that retrieves the lexico-syntactic information in nouns and another route that uses word-form cues, both present in the lexical access process. Lastly, there has been intensive research on the different mechanism the brain utilizes to access regular and irregular forms, such as the case of irregular verbal morphology in English (Marslen-Wilson 2007; Jackendoff 2002; Ullman 2001; Pinker 1991, 1999). Under the Dual Mechanism Model, the processing occurs in two linguistic processes: one rule-based for regular morphology and one that relies on memory for irregular morphology.

There is scarce psycholinguistic research that measures the role of predictive nominal endings in gender processing in heritage speakers of Spanish. Research carried out by

Montrul, Davidson, De La Fuente and Foote (2014) with heritage speakers of Spanish found a facilitatory effect of noun transparency and some offline studies carried out with second language learners and heritage speakers of Spanish also found that participants were more accurate with transparent nouns (Hur *et al.* 2020; Montrul *et al.* 2013; Alarcón 2011; Montrul *et al.* 2008)

The motivation of the present study was to provide evidence of the role of noun canonicity (opaque and transparent gender marking phonemes) in monolingual and heritage speakers of Spanish and to observe whether gender phonological cues are utilized differently by different linguistic populations. Heritage Speakers (HSs) are bilinguals who grew up in families which speak a different language than the dominant language of the mainstream society they live in (Valdés 2005). When they start schooling, heritage speakers in the United States become increasingly exposed to the dominant language of the mainstream society, therefore, their exposure to their languages changes, resulting in shifting linguistic dominance to the second, exposed language (Valdés 2005; Montrul 2012). For the present investigation, young adult bilinguals born in the United States with different degrees of language dominance in Spanish were recruited.

Findings on how gender features are retrieved and whether no hyphen: gender cues override other lexico-syntactic information in monolingual speakers are varied and psycholinguistic research on how heritage speakers utilized nominal predictive endings is very scarce. Furthermore, the different methodologies utilized (further explained below)—with some being more metalinguistic in nature, such as grammaticality judgment tasks or gender decision tasks—might have compelled the speakers to use strategies other than those that would typically be recruited in gender processing. Therefore, the aim of the present study is to provide further evidence on the impact of noun gender-correlated endings in monolingual and bilingual speakers by employing a more implicit task that does not require speakers to use their metalinguistic knowledge of gender.

2. NOMINAL AGREEMENT STUDIES

2.2. Monolingual speakers

An ongoing issue in gendered languages like Spanish is whether noun canonicity (opaque/transparent nouns) affects gender retrieval. Spanish is overall considered a transparent language with respect to gender phonological endings and the type of agreeing elements that encode gender (e.g., articles, adjectives). As stated above, one of the views considers form-based cues (i.e., overt gender markings, such as *-a* or *-o* in Spanish, on articles, determiners, suffixes) the most reliable type of cue for gender retrieval. Following the terminology used in Gollan and Frost (2001), this was called the Reliable Cue Hypothesis, and it is supported by studies that have found strong effects of gender-marking cues in determiners and nouns (Bates *et al.* 1996; Taft and Meunier 1998). These studies mostly used gender decision tasks in which participants indicated the gender of target words. However, using an online methodology, particularly event-related potentials (ERP) Caffarra, Siyanova-Chanturia, Pesciarelli, Vespignani and Cacciari (2015) studied the role of predictive nominal endings in monolingual speakers of Italian using word-by-word sentence comprehension task with determiner-noun pairs that agree and disagree with the referent noun. Their study found evidence “that the processing system can rapidly detect formal cues to gender during online sentence comprehension, as predicted by the reliable cue hypothesis” (Caffarra *et al.* 2015: 9). A second view, the Lexico-Syntactic Access View, is based on studies that claim that gender can be retrieved without the help of overt gender marking cues at the form level. These studies use the Tip of the Tongue paradigm (TOT) and this phenomenon can be explained as the temporary stage in which speakers cannot retrieve a word from memory or the word is not strongly activated and speakers cannot recall it (Meyer and Bock 1992). TOT studies examined the relationship between grammatical gender and lexical access. These studies have been carried out in different languages (e.g., Italian, French, Spanish, and German) with healthy participants (Caramazza and Miozzo 1997;

Miozzo and Caramazza 1997; Vigliocco *et al.* 1997) and with participants with language impairments such as anomia (Badecker *et al.* 1995; Gonon *et al.* 1989; Kulke and Blanken 2001). Results from both populations have provided evidence that speakers can retrieve gender even when they do not access gender lexico-syntactic information, that is, without the help of noun gender cues, i.e. *-o* or *-a* in Spanish (or when they have a deficit, in the case of participants with anomia). The authors of these studies (Badecker *et al.* 1995; Miozzo and Caramazza 1997; Vigliocco *et al.* 1997) suggest that gender retrieval can be done without relying on gender-correlated endings, meaning that there is no difference in gender retrieval between canonical/transparent and non-canonical/opaque nouns.

A third competing view, called the Two-Route Hypothesis (Gollan and Frost 2001), is represented by studies that acknowledge the presence of two routes to retrieve gender, one purely syntactic and another that relies on overt gender marking cues. Gollan and Frost (2001) studied gender retrieval in Hebrew (a language without articles), using a gender decision task and a timed grammaticality judgment task (GJT). In Experiment 1, which consisted of naming the gender of the noun that appeared on the screen researchers found strong effects (showing faster naming time) of regularity (masculine and feminine marked nouns), replicating prior work using the same task (Bates *et al.* 1996; Taft and Meunier 1998). Regular nouns provided the fastest RTs and the fewest errors, indicating that gender-marked cues in nouns impact access to gender. However, in Experiment 2, researchers tested retrieval of grammatical gender in noun-adjective sequences that appeared on the screen and participants had to judge the grammaticality of the sequence via keyboard. Results on RTs and accuracy indicated that participants had the same difficulty making grammaticality judgments on regular nouns as they did on irregular nouns. In both the RTs and error analysis of grammatical pairs, the researchers did not encounter a significant main effect of regularity. In other words, there was no difference between regular and irregular nouns. However, there was

a significant effect of regularity in the pairs that were judged ungrammatical, in both the RTs and error analysis. In other words, when the researchers analyzed the ungrammatical pairs, they found an effect on regularity. The difference found in Experiment 2 (lack of effect of regularity on grammatical responses) was interpreted as suggesting that “the detection of correct gender agreement was far less affected by gender-marking” (Gollan and Frost 2001: 642). To explain the contradictory results on ungrammatical responses, the researchers suggested a two-route hypothesis, in which the first route “derives gender from its correlation with gender marking at the level of form” (Gollan and Frost 2001: 644). Yet, they suggested that a second route (a lexical one) is necessary because access to gender was done as easily and accurately for gender marked and unmarked forms of grammatical pairs—in other words, without any effect of gender predictive nominal endings. Using different psycholinguistic methodology, Caffarra, Jassen and Barber (2014) provided more evidence of the existence of two routes to access gender. Caffarra *et al.* (2014) conducted an ERP study using the visual half field (VHF) paradigm while monolingual speakers of Spanish performed a grammatical judgment task (a comprehension task) in which they assessed the grammaticality of the determiner-noun pairs that were displayed on the computer screen. They interpreted their ERP results as evidence that the lexical route recovers gender abstractly while word-form cues are used at different stages of gender processing.

Another model that offers interesting predictions on how speakers utilize regular/irregular marking cues (in our experiment gender-correlated endings) is the Dual Mechanism Model (Marslen-Wilson 2007; Jackendoff 2002; Ullman 2001; Pinker 1991, 1999). As previously mentioned, this model has been extensively tested with regular and irregular verbal morphology in English. This model proposes a mechanism that computes regular morphology and another mechanism that relies on memory for irregular morphology. Based on this assumption, lexical frequency does not play a role with high-frequency regular nouns because regular nouns are

computed and use a rule-based application system. However, for high-frequency irregular nouns, since they are stored in memory and do not need to be regularly computed, these nouns have an advantage in naming times over high-frequency regular nouns. Opposite to the pattern predicted for high-frequency nouns, low-frequency irregular nouns have a disadvantage over low-frequency regular nouns since irregular nouns have to be looked up in the mental lexicon. There is not much research that compares the processing patterns of regular and irregular verbal morphology with the patterns of irregular nominal endings in languages with nominal agreement. Resende and Mota (2017) studied these patterns testing the Dual Mechanism Model with transparent/opaque and regular/irregular noun endings in monolingual speakers of Brazilian Portuguese. Investigators manipulated word frequency and gender marking cues (opaque/transparent and regular/irregular) in a timed gender selection task with determiner-noun and noun-adjectival agreement. This study did not find evidence in favor of the Dual Mechanism Model to process predictive nominal endings since they found faster naming times in high-frequency transparent nouns than high-frequency opaque nouns, against the predictions of the model.

Apart from the previous hypothesis and models presented, it is worth mentioning that grammatical gender processing has also been examined extensively using the picture-word interference paradigm (Rosinski, Golinkoff and Kukish 1975). This paradigm consists in a naming task in which objects/pictures/words appear on the screen along with a distractor object/picture/word. These distractors can be semantically or phonologically related or not or they may be manipulated to exhibit similar or different agreement features (e.g., gender features: picture of a feminine noun with a picture of a masculine noun as an example of syntactic unrelated distractor). Studies that focus on grammatical gender with the picture-word interference paradigm manipulate the grammatical gender of the distractor word to determine whether there is a gender congruency effect; that is, whether the grammatical gender of the distractor mis/matches the target noun



and affects lexical access of the target item. Findings have provided evidence that overt gender grammatical cues are used during the agreement process (For evidence in Spanish, see Urrutia, Domínguez and Álvarez 2009 and for a meta analysis on the subject, see Sá-Leite, Luna, Tomaz, Fraga and Comesaña 2022).

2.2. Bilingual Heritage Speakers

Regarding research on the role of noun canonicity with bilingual heritage speakers of Spanish, Montrul *et al.* (2008) administered a written gender recognition task (determiner-noun agreement) and oral picture description task (noun-adjective agreement) that included both transparent and opaque nouns. Accuracy results found that HSs produced a large number of errors with Spanish nouns that were opaque (or non-canonical), particularly in noun-adjective agreement. Similar to Montrul *et al.* (2008), Alarcón (2011) examined the role of transparent gender marking cues in the production of agreement by HSs of Spanish using a written gender recognition task and an oral picture description task. HSs were more accurate in producing agreement with transparent/canonical nouns than with opaque/non-canonical nouns. Montrul *et al.* (2013) also carried out a study with HSs of Spanish and noun canonicity. In this study, participants were required to produce nouns in their diminutive form in Spanish (e.g., *un pez* ‘a fish’, diminutive: *un pececito* ‘a little fish’; *una cruz* ‘a cross’, diminutive: *una crucecita* ‘a little cross’). They also found an effect of noun transparency/canonicity showing that both groups were less accurate when the nouns had opaque gender marking cues. Recent research tested heritage speakers of Spanish with different language proficiency (Hur *et al.* 2020). In their study, they administered two oral production tasks examining determiner-noun and noun-adjective agreement with canonical/transparent and non-canonical/opaque gender marking cues. Similar to the present investigation, Hur’s *et al.* (2020) study divided the heritage speaker’s population into two groups according to their dominance level or, as they mentioned, proficiency based on their productive vocabulary knowledge

as measured by the Multilingual Naming Test Gollan (MINT) (Gollan, Weissberger, Runnqvist, Montoya and Cera 2012). Results found that opaque nouns lead to higher error rates in line with previous studies (Montrul 2013, 2008; Alarcón 2011). However, they did not report any results on the role of noun gender-correlated endings relative to heritage speakers’ proficiency level. Using online methodology, Montrul, Davidson, De La Fuente and Foote (2014) conducted three spoken word recognition tasks where heritage speakers of Spanish had to select the grammaticality of the sentence. Results indicated that HSs had faster RTs and were more accurate with nouns that were transparently marked for gender. Furthermore, research using the picture-word interference paradigm has also provided evidence of a facilitatory effect of noun canonicity in bilingual speakers. However, studies have mainly focused on late second language learners not heritage speakers of Spanish. (see Sá-Leite, Fraga and Comesaña 2019 for a review on the topic).

To summarize, the research presented above, with some studies using metalinguistic tasks (gender decision, grammatical judgments, and gender completion tasks), has provided different evidence on how gender features are retrieved by monolingual speakers and whether gender-cues override other lexico-syntactic information are varied. With regard to research on the role of gender cues in heritage speakers of Spanish, research carried out on the effects of gender correlated endings in HSs of Spanish is very scarce (Hur *et al.* 2020; Montrul *et al.* 2013; Montrul *et al.* 2014; Alarcon 2011; Montrul 2008). Moreover, some of the research discussed above used offline methodology and/or administered experiments that were also metalinguistic in nature resembling classroom-taught metalinguistic tasks that might be unfamiliar to heritage speakers of Spanish (Hur *et al.* 2020; Alarcon 2011; Montrul 2008).

This study contemplates the main competing hypotheses on the subject of predictive nominal endings in gender retrieval: The Reliable Cue Hypothesis (Bates *et al.* 1996; Taft and Meunier 1998; Caffarra *et al.* 2015) that considers predictive nominal endings as a reliable cue

to access gender. The Two-Route Hypothesis (Gollan and Frost 2001) acknowledges the presence of two routes to retrieve gender, one purely lexico-syntactic and another route that relies on word-form cues (phonological/orthographic information in nouns). Lastly, this study also considers the predictions of the Dual Mechanism Model (Marslen-Wilson 2007; Jackendoff 2002; Pinker 1991, 1999; Ullman 2001) that pose that the processing of regular morphology is rule-based while irregular morphology is stored in memory. This study did not consider a purely lexico-syntactic view (Badecker *et al.* 1995; Miozzo and Caramazza 1997; Vigliocco *et al.* 1997) to access gender since the bulk of research carried out to date with different methodologies (also presented above) supports that gender marking cues are utilized in the retrieval of gender features.

In conclusion, the present study compares RTs and accuracy rates of two populations of HSs of Spanish and a group of monolingual speakers of Spanish to observe the effect of predictive nominal endings (transparent/canonical nouns) in gender agreement in different linguistic populations. By employing an oral production task that does not require speakers to use their metalinguistic knowledge of gender, the present investigation aims to provide results on the impact of word-form cues and contribute to the ongoing theories about how predictive nominal endings are utilized in gender processing.

3. CURRENT STUDY AND RESEARCH QUESTIONS

As stated above, the aim of this study is to investigate the role of noun canonicity in heritage speakers and monolingual speakers of Spanish so as to determine whether predictive nominal endings are the most reliable cue to process gender (the Reliable Cue Hypothesis; Bates *et al.* 1996; Taft and Meunier 1998) or whether speakers utilized both lexico-syntactic and word-form information (Two-Routes hypothesis: Gollan and Frost 2001). This study will also test the predictions of the Dual Mechanism Model (Marslen-Wilson 2007; Jackendoff 2002; Pinker 1991, 1999; Ullman 2001) that also poses a dual route mechanism

to process gender. Furthermore, the purpose of this investigation was to determine whether different linguistic populations, monolingual speakers, and two groups of HSs of Spanish with different dominance, utilize gender word-form cues differently when accessing gender. To that end, the research questions are presented below:

RQ (1): What is the role of noun canonicity (transparent and opaque nouns) in gender agreement in monolingual speakers and HSs of Spanish?

RQ (2): Do different linguistic populations differ in the way gender-correlated endings are utilized?

RQ (3) Are the noun canonicity effects similar to other dual-route models?

To answer these research questions, the noun lexical frequency (high- and low-frequency) of our experimental stimuli (canonical/transparent and non-canonical/opaque nouns) was manipulated. By manipulating the lexical frequency different outcomes were expected. Below the predictions:

RQ (1) Predictions: According to the Reliable Cue hypothesis (Bates *et al.* 1996; Taft and Meunier 1998; Caffarra *et al.* 2015), transparent nouns will be retrieved faster and more accurately than opaque nouns showing that grammatical gender is most easily retrieved with the help of transparent/canonical endings. An alternative outcome, which is partially compatible with the previous view, is that the retrieval of gender features will depend on both gender-correlated endings and also lexical information (lexical frequency of the noun), assuming a dual-route model, or as Gollan and Frost (2001) named it, the Two-Route Hypothesis. This hypothesis assumes that both word-form cues, and lexical information play a role in retrieving gender. Therefore, it will be expected that the gender of high-frequency nouns will be easily accessed, but also, gender marking phonemes



will play a role in the retrieval of gender, therefore high-frequency transparent nouns will be retrieved faster than high-frequency opaque nouns (as observed in Alarcón 2011, Resende and Mota 2017) and low-frequency transparent nouns will be retrieved faster than low-frequency opaque nouns since predictive nominal endings will help in gender retrieval.

RQ (2) Predictions: This study also explored how gender marking cues are utilized by different populations. The HSs group was divided into two groups based on language dominance that was measured using the Multilingual Naming Test Gollan (MINT) (Gollan *et al.* 2012) (further details about the MINT test in section 4.2.). Since the present investigation administered two oral productions tasks measuring participants RTs and accuracy, predictions are based on some offline studies that used similar oral production experiments that were conducted with HSs of Spanish (Hur *et al.* 2020; Alarcón 2011; Montrul *et al.* 2013) and Montrul's *et al.* (2014) spoken word recognition research that measured RTs and accuracy rates in HSs of Spanish. Therefore, in line with previous studies, overt gender-correlated endings, transparent/canonical nouns, will be predictive of higher accuracy rates in monolingual and bilingual speakers (both groups). Regarding naming times, the effects of predictive nominal phonemes will be seen in faster reaction times with canonical/transparent nouns in both HSs groups and also the monolingual speakers' group (Montrul *et al.* 2014). Additionally, since there are no clear predictions of whether monolingual and bilingual speakers differently utilize predictive nominal endings, due to heritage speakers' divided frequency of use between the two languages (Gollan 2008) and the fact that English does not have overt phonological/orthographic nominal endings, one of the possible outcomes is the lack of canonicity effects in the HSs' accuracy or/and RTs analysis. Similarly, among the bilingual group, based on the same rationale explained before, if dominance plays a role, HSs with less dominance will exhibit the

lowest activation of word-form information, therefore, a possible lack of canonicity effect in this group.

RQ (3) Predictions: Based on the predictions of the Dual Mechanism Model (Marslen-Wilson 2007; Jackendoff 2002; Pinker 1991, 1999; Ullman 2001) participants will be faster with high-frequency opaque nouns than high-frequency transparent nouns because high-frequency opaque nouns are stored in memory and not computed, therefore, opaque high-frequency nouns will have an advantage over high-frequency transparent nouns. With low-frequency, the opposite pattern will be observed, participants will be faster with low-frequency transparent nouns than with low-frequency opaque nouns since low-frequency opaque nouns would need to be looked up in memory, while low-frequency transparent nouns will be simply computed (rule-based).

4. METHODS

4.1. LEAP-Q and Language Background Questionnaire

Participants completed the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian, Blumenfeld and Kaushanskaya 2007) in Spanish. This questionnaire elicits self-reported information about bilingual linguistic background such as the age of acquisition and history of both past and present language exposure of the bilingual in different contexts. The questionnaire also provides information about the bilingual's self-reported language proficiency and dominance. Apart from the LEAP-Q, participants also fill a short language background questionnaire (LBQ) that was designed by the experimenter to elicit other information that the LEAP-Q did not prompt such as Spanish and English usage, schooling in Spanish and English, Spanish proficiency in writing, time spent in a Spanish speaking country, and parents and grandparents' place of birth. In the LBQ, monolingual and bilingual participants assessed their frequency of usage of Spanish and English on a scale of 1 to 8 (where 1 was "everyday" and 8 was "never").

4.2. Standard Proficiency Measures: DELE, MELICET and MINT

To further assess proficiency in Spanish, apart from administering the self-reported tests mentioned above, monolinguals and bilinguals completed the DELE (*Diploma de Español como Lengua extranjera*), a standardized Spanish grammar exam usually administered to L2 speakers of Spanish. English proficiency was measured in the bilingual population by administering an adapted version of the standardized Michigan English Language Institute College English Test (MELICET). In both tests, participants must select from a drop-down menu the proper grammatical option.

Furthermore, to determine Spanish dominance, the Multilingual Naming Task (MINT) (Gollan *et al.* 2012) was administered to bilingual speakers. This test is similar to the Boston Naming Task (BNT; Goodglass, Kaplan and Weintraub 1983), but it was designed specifically for bilingual speakers. As mentioned by the developers of the MINT, the BNT underestimates bilinguals' ability in Spanish. The Multilingual Naming Task has been proven to be a better evaluation of bilingual performance than the BNT; it avoids cognates, maximizes proficiency in language-specific knowledge, and diminishes influence from the non-target language (Gollan *et al.* 2012). In the present investigation, the term more/ less dominant in Spanish is used to describe both of our heritage speakers' groups instead of more or less proficient (lacking proficiency) because both groups were fluent in Spanish, thus proficient. More or less dominance in the minority language better explains the status of Spanish at the moment of the investigation, describing heritage speakers' Spanish dominance as a dynamic continuum where dominance can change depending on the exposure to the specified language (Valdés 2014).

4.3. Vocabulary Test Experiment 2

This vocabulary test consists of providing the names with the corresponding article of the experimental pictures that appear in Experiment 2. The vocabulary test was

carried out because Experiment 2 is a Picture Description Task, and participants must use only the adjective to describe what appears on the screen without uttering the name of the item they saw on the screen; they describe what appeared on the screen using the adjectives *claro/a* or *borroso/a* (clear or blurry). Therefore, participants needed to identify the name and article of the experimental pictures to verify that they were familiar with the name of the picture, and hence, its grammatical gender. The results of this vocabulary test were considered during data coding to discard noun trials where the speakers did not know the target referent or provide the incorrect article.

4.4. Participants

A total of 34 monolingual Spanish-speaking undergraduate students from Pontificia Universidad Católica Madre y Maestra (PUCMM) in Santiago de los Caballeros, Dominican Republic (24 men and 10 women, mean age= 20.8 years, standard deviation=2.54) voluntarily took part in the study. Monolingual participants were born in the Dominican Republic and lived there for their entire life. This group self-reported to be dominant in Spanish, even when they self-reported to have basic proficiency or no proficiency in English or another language. A total of 44 Spanish-English bilingual speakers from the University of Florida (17 men and 27 women, mean age=19.8 years, standard deviation=0.8) participated voluntarily or for class credit in the present investigation. All bilingual participants were born in the United States and, following Silva-Corvalan's criteria for classifying generation in HSs, twenty-seven bilingual participants were 2nd generation HSs since they were born in the US but not their parents and seventeen bilingual participants were 3rd generation HSs since they were born in the US as well as their parents. Monolingual and bilingual speakers were not significantly different in mean age ($p= 0.06$). As mentioned above, we use the Multilingual Naming Test (Gollan *et al.* 2012) to divide the bilingual group into two linguistic groups: a more dominant Spanish-speaking bilingual group (+dominant)



and a less dominant Spanish-speaking bilingual group (-dominant). The MINT consists of an offline Picture Naming Task (PNT) presented on a computer using a PowerPoint presentation, including 68 pictures. Bilingual speakers are asked to identify in Spanish the images in the pictures as fast as possible. As instructed, the experimenter would provide participants with up to 6 seconds to name each image, and then proceed to show the next picture. Results on The Multilingual Naming

Table 2. Mean averages and p-values for MSs and HSs (both groups) results in the LEAP, LBQ and DELE

LBQ and LEAP	MSs	HSs	p-value
Age	20.8	19.8	0.06
Self-reported dominance	Spanish	English	n/a
Spa. frequency of use (writing) 1-8 scale*	1.02	2.7	<0.001***
Spa. frequency of use (speaking) 1-8 scale*	1.05	2.1	<0.001***
Spa. frequency of use(reading) 1-8 scale*	1.7	3.6	<0.001***
Only LEAP			
Spa. self-rated prof. (speaking) 1-10 scale	9.2	8.1	<0.001***
Spa. self-rated prof. (comprehen.) 1-10 scale	9.2	8.7	<0.001***
Spa. self-rated prof. (reading) 1-10 scale	9.2	8.3	<0.001***
Exposure/usage to Spanish (family) 1-10 scale	9.96	8.4	<0.001***
Exposure/usage to Spa. (friends) 1-10 scale	9.5	4.9	<0.001***
Exposure to Spanish (watching TV) 1-10 scale	8.02	3.8	<0.001***
Exposure to Spa. (music/radio) 1-10 scale	8.1	4.6	<0.001***
Standard proficiency measures			
Prof. In Spa. (DELE) (out of 50)	46.55	24.45	<0.001***

*1=frequency,8=infrequency (See section 4.1. for further details)

Task (MINT; Gollan *et al.* 2012) showed that, among the HSs, 24 were more dominant (D+ bilinguals) in Spanish (based on a score higher than 34 in the MINT), than the other 20 HSs (D- bilinguals) in Spanish (who scored lower than 34). Table 2 below shows the results of the background questionnaires as well as the standard proficiency measures.

As we can observe, heritage speakers self-reported to have lower proficiency in Spanish as well as less exposure and usage to the minority language than the monolingual speakers group. The DELE results showed that heritage speakers scored lower in the proficiency test than monolingual speakers of Spanish. A two-sample t-test indicated that the mean rates for the monolingual groups and both heritage speakers' groups were statistically significant. The subsequent table shows the results of each bilingual group.

Table 3. Mean averages and p-values for the two bilingual groups' results in the LEAP, LBQ, MINT and MELICET

LBQ and LEAP	HSs (D+ bilinguals)	HSs (D- bilinguals)	p-value
Age	19.9	19.75	0.463
LBQ			
Spa. frequency of use (writing) 1-8 scale*	2.04	3.55	<0.001***
Spa. frequency of use (speaking) 1-8 scale*	1.58	2.8	<0.001***
Spa. frequency of use(reading) 1-8 scale*	2.8	4.7	<0.001***
Eng. frequency of use (writing) 1-8 scale*	1.45	1.25	0.32
Eng. frequency of use (speaking) 1-8 scale*	1	1.1	0.16
Eng. frequency of use (reading) 1-8 scale*	1.54	1.4	0.54
LEAP			
Spa. self-rated prof. (speaking) 1-10 scale	8.7	7.4	<0.001***
Spa. self-rated prof. (comprehen.) 1-10 scale	9.5	7.8	<0.001***
Spa. self-rated prof. (reading) 1-10 scale	9.1	7.4	<0.001***
Eng. self-rated prof. (speaking) 1-10	9.7	9.8	0.49

LBQ and LEAP	HSs (D+ bilinguals)	HSs (D- bilinguals)	p-value
Eng. self-rated prof. (comprehend.) 1-10 scale	9.9	9.95	1
Eng. Self-rated prof. (reading)	9.75	9.95	0.18
Exposure/usage to Spanish (family) 1-10 scale	9.25	7.4	0.01*
Exposure/usage to Spa. (friends) 1-10 scale	5.6	2.2	<0.001***
Exposure to Spanish (watching TV) 1-10 scale	5.2	2.15	<0.001***
Exposure to Spa. (music/radio) 1-10 scale	6.25	2.8	<0.001***
LBQ and LEAP			
Exposure to Eng. (family) 1-10 scale	1.91	5.55	<0.001***
Exposure to Eng. (friends) 1-10 scale	7.3	9.8	<0.001***
Exposure to Eng (watching TV) 1-10 scale	7.8	9.2	0.03*
Exposure to Eng. (music/radio) 1-10 scale	7.75	9.15	0.02*
Objective proficiency measures			
Prof. In Spa. (DELE) (out of 50)	25.3	23.35	0.07
Prof. In Eng. (MELICET) (out of 50)	44.41	44.55	0.62
Dominance in Spa. (MINT) (out of 68)	45.6	28.3	<0.001***

*1=frequency, 8=infrequency (See section 4.1. for further details)

As we can observe in Table 3, scores from the DELE were numerically higher for the more dominant HSs group (HS + dominant). However, results from the t-test show that these values were not statistically significant. Results from the MINT were statistically significant and we can observe that more/less dominance in Spanish correlated with self-reported frequency of use of Spanish (speaking, writing, reading), self-reported proficiency (speaking, comprehension, reading) and self-reported exposure to the Spanish in different contexts (family, friends, media).

4.5. Materials and Design

The materials for the Picture Naming Task (PNT) and Picture Description Task (PDT) consisted of color drawings depicting the stimuli selected for the experiment. The PDT was an experimental design similar to the one used in Navarrete, Basagni, Alario, and Costa (2006) where the referent noun is not uttered, just the agreeing element (in our experiment, the adjective). In experiment 2, 36 fillers were blurred using an online photo editor (LunaPic photo editor 2018). We piloted the stimuli and images with 10 bilingual speakers of Spanish to ensure that the pictures chosen for the experiments were representative of the thing they depicted, and to avoid dialectal variation and exclude (and replace) items that did not produce the target response.

The stimuli consisted of two experimental lists, with each list containing 44 experimental items, 88 fillers, and 10 practice pictures, resulting in a total of 142 pictures. The order of the lists was counterbalanced. If participants saw List 1 in experiment 1, then List 2 was used for Experiment 2 and if List 2 was administered in Experiment 1, then List 1 was used for Experiment 2. Half of the experimental items were feminine, and half masculine and half of the stimuli were transparent/canonical nouns and the other half were opaque/non-canonical nouns. The lists were balanced in the number of opaque and transparent endings. The lists were controlled for noun frequency, letter length, noun gender-correlated endings (opaque and transparent), gender, imageability, and concreteness. Two-tailed t-tests corroborated that there were no significant differences between the two lists of experimental items regarding frequency ($p = .98$), word length ($p = .69$), imageability ($p = .85$), and concreteness ($p = .32$). Likewise, we carried out two-tailed t-tests with the lists of fillers. Lexical frequency was obtained from the ESPAL database (Duchon, Perea, Sebastián-Gallés, Martí and Carreiras 2013; <http://www.bcbl.eu/databases/espal/index.php>). High-frequency nouns were defined by a frequency of more than 30 per million occurrences and low-frequency nouns had fewer than 30 per million occurrences. A

Table 4. Experimental stimuli and lexical frequency

Transparent/Canonical				Opaque/Non-canonical			
HF	Freq.	LF	Freq.	HF	Freq.	LF	Freq.
Arco	35.373525	Barba	13.9421	Árbol	49.10445	Coliflor	0.47113
Banco	118.62657	Bota	3.43111	Carne	78.869	Maiz	0.55565
Bolsa	54.780714	Búho	2.35239	Corazón	148.457	Buzón	1.36464
Brazo	48.779529	Casco	22.7863	Cruz	105.8834	Tenedor	1.904
Caballo	82.492738	Codo	6.62827	Flor	33.14136	Pincel	2.37838
Caja	51.49907	Corbata	5.20189	Frente	319.5867	Cicatriz	2.726
Casa	403.59025	Cuchillo	17.1068	Fuente	113.0543	Nube	2.96323
Castillo	82.856643	Flecha	9.60775	Juez	82.11909	Guante	3.69429
Corona	52.204136	Fresa	1.99498	Leche	34.96738	Imán	3.88274
Cuello	36.163069	Hoja	25.3629	León	102.42	Jabón	3.93148
Estrella	55.183609	Hombro	17.4642	Muerte	320.9773	Lápiz	5.44558
Fuego	119.10094	Hueso	13.6042	Noche	273.38	Barril	7.60302
Lengua	119.43885	Huevo	12.5287	Paquete	32.90092	Botón	7.82721
Libro	210.86676	Bufanda	1.03323	Pie	108.265	Pez	12.7497
Lluvia	39.090556	Pera	3.54158	Puente	63.37797	Serpiente	13.022
Luna	67.64	Pierna	15.7227	Raíz	43.36969	Nuez	15.0436
Mesa	100.24611	Plato	18.5332	Red	124.29	Miel	16.6292
Ojo	39.57	Pluma	17.9158	Sangre	105.1459	Nariz	20.2812
Playa	39.084058	Pulpo	1.44587	Sol	146.865	Llave	21.8668
Puerta	158.97779	Trofeo	10.699	Torre	62.35449	Reloj	26.1459
Toro	31.074896	Uña	1.65707	Traje	31.71173	Nieve	30.2724
Vestido	31.224357	Vaso	15.0501	Tren	40.39672	Sal	30.7077

two-tailed t-test corroborated that there was a significant difference between HF and LF items in each list, verifying that there was not an overlap in frequency between HF and LF nouns ($p = .001$).

4.6. Procedure

Participants filled out and signed a consent form at the beginning of the session. Then, the two background questionnaires were administered, followed by the experimental tasks. Pictures from both experimental tasks were presented using E-Prime 2.0 (SP1, Psychology Software Tools, Pittsburgh, PA). Naming times were recorded using an Audio-Technica AT1200 microphone connected to a response box. Correct and incorrect responses were verified manually. Each trial began with a 500 ms. fixation cross (+) in the middle of the screen immediately followed by the picture.

Each participant saw a different list in each experiment.

In experiment 1, participants were instructed to name the pictures with the corresponding definite article *el*-_{DET.M.SG} or *la*-_{DET.F.SG} (the) as quickly as possible, using only one word.

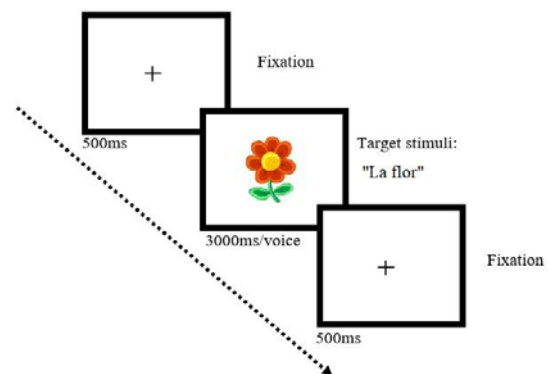


Figure 1. Sample of Experiment 1 picture naming task testing an opaque feminine noun

Experiment 2 was a picture description task. Participants were presented with a picture stimulus on the screen and were instructed to use *borroso/borrosa* ('blurry'-ADJ.M.SG /'blurry'-ADJ.F.SG) or *claro/clara* ('clear'-ADJ.M.SG /'clear'-ADJ.F.SG) to describe what appeared on the screen. Participants were instructed to use only the adjective.

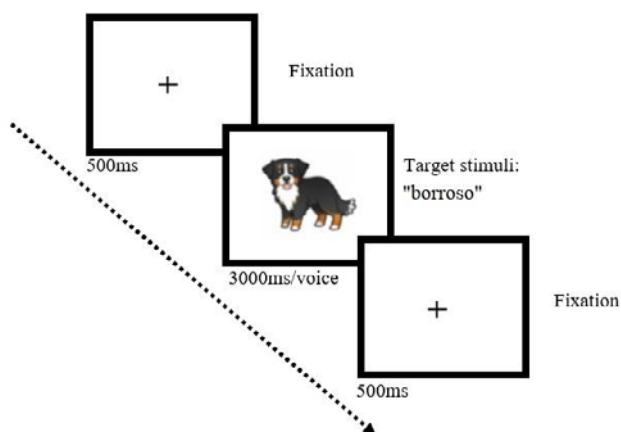


Figure 2. Sample of Experiment 2 picture description task with a filler item (blurred)

After completing the two experimental tasks, monolingual participants completed the vocabulary test and the DELE. HSs participants completed the vocabulary test, the MELICET and the MINT test.

5. EXPERIMENT 1 RESULTS: PICTURE NAMING TASK

In this experiment, reaction time (RTs) analyses were conducted on correct responses only. A response was accepted as correct if participants could name the picture that appeared on the screen with the correct article *el* or *la* before the picture disappeared from the screen (3000 ms.). Responses with *un* or *una* ('a'/'an') were also accepted. Correct responses after self-correction were not accepted. Answers that were not clear due to lack of audibility or unclear pronunciation were discarded.

5.1. Monolingual group data

From a total of 1495 possible trials, accuracy data consisted of 1275 data points. There

were no inaccurate trials for monolingual speakers, that is, the noun was always named with the corresponding article. Skipped trials or non-identifiable responses consisted of 178 data points (12%) (these were excluded from the analysis). From a total of 1495 possible trials, RTs data consisted of 1243 data points. Data points outside the 3000 ms. window, non-recorded responses, skipped trials and inaccurate or non-identifiable responses constituted 209 data points (14%) (those were excluded from the RTs analysis).

5.2. Bilingual group data

For D+ bilinguals, from a total of 1057 trials, accuracy data for this group consisted of 843 data points. Skipped trials or non-identifiable responses constituted 178 data points (16%) (those were not included in the accuracy analysis). From a total of 1057, RTs data consisted of 692 data points. Data points outside the 3000 ms. window, non-recorded responses, and skipped trials and inaccurate or non-identifiable responses constituted 365 data points (34.5 %) (those were not included in the RTs analysis).

For D- bilinguals, from a total of 880 trials, accuracy data for this group consisted of 471 data points. Skipped trials or non-identifiable responses constituted 311 data points (35%) (those were excluded from the accuracy analysis). From a total of 880, RTs data consisted of 398 data points. Data points outside the 3000 ms. window, non-recorded responses, skipped trials and inaccurate or non-identifiable responses constituted 481 data points (54 %) (those were excluded from the RTs analysis).

The independent variables reported in both analyses are noun canonicity (transparent and opaque) (within-subject variables) and group (between-subject variable): Monolingual speakers, more Spanish-dominant bilinguals (D+ bilinguals) and less Spanish-dominant bilinguals (D- bilinguals). The interaction between frequency and canonicity was also reported. A repeated measures ANOVA was carried out to obtain the statistical significance of the data. The following section outlines the results of the RT analysis.

5.3. Results: Monolingual, D+ and D-Bilinguals

As observed in Figure 3, monolinguals were numerically faster naming canonical/transparent nouns with both high- and low-frequency nouns. However, for the bilingual group, they seemed to name transparent nouns slightly slower than opaque nouns with high-frequency nouns. Regarding low-frequency nouns, both bilingual groups were numerically faster naming transparent nouns. This difference was more visible for the D- group. Below, Figure 3 shows the mean picture naming times and standard error (SE) for Experiment 1 RTs analysis.

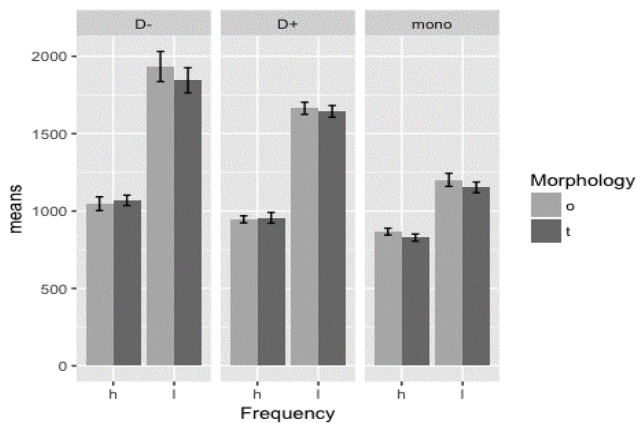


Figure 3. Experiment 1 mean RTs and SE split by frequency and canonicity

To assess the statistical significance of the results, a repeated-measures ANOVA was conducted using R (R core team 2015, version 3.1.3) with a 2 x 2 x 3 factorial design with the within-subjects factors of Frequency (High versus Low) and Canonicity (Opaque versus Transparent) and the between-subjects factor of Group (Monolinguals, Bilinguals D+, Bilinguals D-). Below the results of this model are presented.

As the output shows, a main effect of group ($F(2,70) = 89.645, p = 0.001$), canonicity ($F(1,70) = 6.27, p = 0.001$) and frequency ($F(1,70) = 532.791, p = 0.001$) was found. However, the interaction of canonicity with group or frequency was not significant.

The groups were explored individually and a repeated-measures ANOVA was conducted with a 2 x 2 factorial design with the within-

Table 5. Experiment 1 three-groups 2 x 2 x 3 repeated-measures ANOVA output

Effect	DFn	DFd	F	p
Group	2	70	89.645	<0.001***
Frequency	1	70	532.791	<0.001***
Canonicity	1	70	6.27	<0.001***
Group:Canonicity	2	70	0.791	0.457
Freq:Canoni	1	70	1.372	0.245
Group:Freq:Canoni	2	70	1.186	0.311

subjects factors of Frequency (High versus Low) and Canonicity (Opaque versus Transparent). For the monolingual group, the output showed a main effect of canonicity ($F(1,32) = 5.741, p = 0.001$), that is, transparent nouns were named faster than opaque nouns. There was no significant interaction between frequency and canonicity for this group. However, for the bilingual group, results from the repeated-measures ANOVA analysis showed that canonicity was not significant in any of the bilingual groups.

For the accuracy results, numerically, monolinguals scored at ceiling. Among the bilingual group, both bilingual groups were on average more accurate with high-frequency transparent nouns. The larger magnitude difference between opaque and transparent nouns were seen with low-frequency nouns.

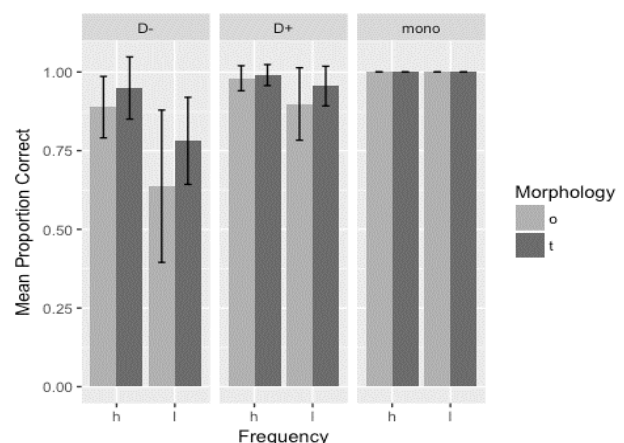


Figure 4. Experiment 1 mean accuracy and SE split by frequency and canonicity

To assess the statistical significance of the results, a repeated-measures ANOVA was conducted using R (R core team, 2015, version

3.1.3) with a 2 x 2 x 3 factorial design with the within-subjects factors of frequency (high versus low) and canonicity (opaque and transparent) and the between-subjects factor of group (Monolinguals, D+ bilinguals, D- bilinguals). The results of this model are presented in Table 6.

Table 6. Experiment 1 three-groups 2 x 2 x 3 repeated-measures ANOVA output

Effect	DFn	DFd	F	p
Group	2	74	84.946	<0.001***
Frequency	1	74	43.229	<0.001***
Canonicity	1	74	15.393	<0.001***
Group:Canonicity	2	74	9.550	<0.001***
Frequency:Canonicity	1	74	5.750	<0.001***
Group:Freq:Canon.	2	74	2.564	0.083

A main effect for group ($F(2,74) = 84.946$, $p = 0.001$), frequency ($F(1,74) = 43.229$, $p = 0.001$) and canonicity ($F(1,74) = 15.393$, $p = 0.001$) was found. Furthermore, the interaction between group and canonicity ($F(2,74) = 9.550$, $p = 0.001$) and the interaction between canonicity and frequency was significant ($F(1,74) = 5.750$, $p = 0.001$). Given these interactions, separate analyses per group were conducted as well as a bilingual comparison analysis.

A repeated-measures ANOVA was conducted with a 2 x 2 factorial design with the within-subjects factors of Frequency (High versus Low) and Canonicity (Opaque versus Transparent). Monolinguals scored at ceiling which translated to null results for our ANOVA model. The statistical analysis conducted (repeated-measures ANOVA) with both bilingual groups found that canonicity ($F(1,42) = 15.289$, $p = 0.001$) was significant as well as the frequency and canonicity interaction ($F(1,42) = 5.712$, $p = 0.001$) indicating, based on the direction of the results of Figure 4, effects of transparency in high- and low- frequency nouns. Further statistical analysis with each subgroup found a main effect of canonicity in D+ bilinguals ($p = .035$) and D- bilinguals ($p = .001$) as well as frequency effects in D+ bilinguals ($p = .001$) and D- bilinguals ($p = .001$) indicating that both bilingual populations were individually more accurate with canonical/transparent nouns

than with non-canonical/opaque nouns but also more accurate with high-frequency than low-frequency nouns.

6. EXPERIMENT 2: PICTURE DESCRIPTION TASK

In experiment 2, RTs analyses were also conducted on correct responses. A response was accepted as correct if participants could describe the picture that appeared on the screen using *claro/a* (clear-_{ADJ.SING.MASC/FEM}). Accuracy analyses were conducted on responses in which participants produced only *claro/a* (clear-_{ADJ.SING.MASC/FEM}). Accurate responses that were outside the RTs frame (3000 ms.) or were not recorded by the program were also coded for the accuracy analysis. Responses that were not clear due to lack of audibility or unclear pronunciation were discarded.

6.1. Monolingual data

From a total of 1495 possible trials, accuracy data consisted of 1289 data points (adjective matches and mismatches with the correct referent). Inaccurate trials consisted of 17 data points. Skipped trials or non-identifiable responses consisted of 206 data points (13%) from the total possible trials (those were excluded from the accuracy analysis). From a total of 1495 RTs data consisted of 1237 data points. Data points outside the 3000 ms. window, non-recorded responses, skipped trials and inaccurate or non-identifiable responses constituted 258 data points (17%) (those were excluded from the RTs analysis).

6.2. Bilingual data

For D+ bilinguals, from a total of 1054 data points, accuracy data for this group consisted of 895 data points (adjective matches and mismatches with the correct referent). Inaccurate data consisted of 104 data points. Skipped trials or non-identifiable responses constituted 159 data points (15%) (those were excluded from the accuracy analysis). From a total of 1056 data points, RTs data

consisted of 755 data points. Data points outside the 3000 ms. window, non-recorded responses, skipped trials and inaccurate or non-identifiable responses constituted 301 data points (28%) (those were excluded from the RTs analysis).

For D- bilinguals, from a total of 880 data points, accuracy data for this group consisted of 639 data points (adjective matches and mismatches with the correct referent). Inaccurate responses constituted 169 data points. Skipped trials or non-identifiable responses constituted 241 data points (28%) (those were excluded from the accuracy analysis). From a total of 880 data points, RTs data consisted of 416 data points. Data points outside the 3000 ms. window, non-recorded responses, skipped trials and inaccurate or non-identifiable responses constituted 435 data points (49%) (those were excluded from the RTs analysis).

The independent variables investigated in both analyses were canonicity (opaque/transparent) and frequency (within-subject variables) and group (between-subject variable): monolingual speakers, more Spanish- dominant HSs (D+ bilinguals), and less Spanish-dominant HSs (D- bilinguals).

6.3. Results: Monolingual, D+ and D- Bilingual Group

Looking at the average times for canonicity and frequency together, all groups named more slowly adjectives on low- and high-frequency opaque noun trials than adjectives on low- and high- frequency transparent trials. Among all groups, less dominant bilinguals produced the slowest naming times with low-frequency opaque nouns. The greater magnitude difference between opaque and transparent noun trials was observed amongst monolingual speakers with low-frequency nouns.

To assess the statistical significance of the RTs results, a repeated-measures ANOVA was conducted using R (R core team 2015, version 3.1.3).

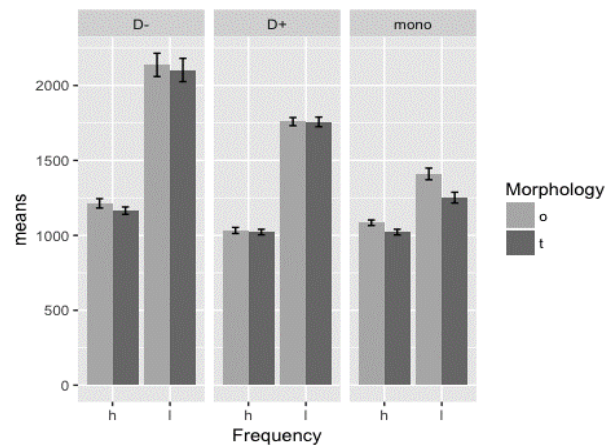


Figure 5. Experiment 2 mean RTs and SE split by frequency and canonicity

Table 7. Experiment 2 three-groups $2 \times 2 \times 3$ repeated-measures ANOVA output

Effect	DFn	DFd	F	p
Group	2	73	43.071	<0.001 ***
Frequency	1	73	545.040	<0.001 ***
Canonicity	1	73	32.000	<0.001 ***
Group:Canonicity	2	73	8.909	<0.001***
Freq:Canon.	1	73	5.906	<0.001***
Group:Freq:Canon.	2	73	2.610	0.08

As the output shows and based on the numerical patterns of Figure 5, canonicity was significant ($F(1,73) = 32, p = 0.001$), Frequency and also the interaction between frequency and canonicity ($F(1,3) = 5.906, p = 0.001$) indicating a different effect of canonicity depending on the noun frequency. Furthermore, the significant interaction between group and canonicity ($F(2,73) = 8.909, p = 0.001$) also indicates a difference in canonicity effect for each group. Due to these results, we conducted separate ANOVA models for each subgroup as well as one for the bilingual group together.

The monolingual group results indicated that canonicity ($F(1,33) = 33.33, p = 0.001$) was significant as well as the interaction of canonicity and frequency ($F(1,33) = 10.418, p = 0.001$), indicating that the effect of canonicity was different depending on the lexical frequency of the noun. Because of this interaction, a pairwise test was run to study all pairs. Results revealed that monolingual participants were faster with low-frequency transparent nouns than low-frequency

opaque nouns, revealing a positive effect of transparency with low-frequency nouns. No significant effect of canonicity for the high-frequency nouns was found.

For the bilingual group analysis, statistical results showed that canonicity was significant along with the interaction between group and canonicity ($F(1,40) = 12.383, p = .001$), indicating a different canonicity effect within each bilingual group. As a result, separate ANOVAs were conducted for each bilingual group and follow-up for any interactions. Results showed that canonicity was not significant for the more Spanish-dominant group. However, for the less Spanish-dominant group, there was a main effect of canonicity. Based on the numerical patterns, the canonicity effect indicated a positive effect of canonical/transparent nouns, that is, adjectives on transparent noun trials were named faster than on opaque noun trials by the less Spanish-dominant group.

As per the accuracy results, mean accuracy showed that all groups were more accurate with high-frequency transparent nouns. For low-frequency nouns, monolingual speakers were also numerically more accurate with transparent nouns. However, both bilingual groups were numerically less accurate with low-frequency transparent nouns.

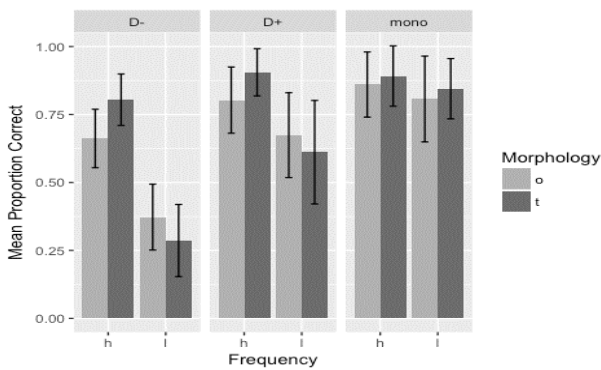


Figure 6. Experiment 2 mean accuracy and SE split by frequency and canonicity

To assess the statistical significance of the results, a repeated-measures ANOVA was conducted using R (R core team, 2015, version 3.1.3) with a 2 x 2 x 3 factorial design with the within-subjects factors of frequency (High versus Low) and canonicity (Opaque versus

Transparent) and the between-subjects factor of group (Monolinguals, D+ bilinguals, D- bilinguals). The results of this model are presented below.

Table 8. Experiment 2 three-groups 2 x 2 x 3 repeated measures ANOVA output

Effect	DFn	DFd	F	p	ges
Group	2	75	129.414	<0.001***	0.506
Frequency	1	75	130.006	<0.001***	0.359
Canonicity	1	75	6.059	<0.001***	0.012
group:Canonicity	2	75	0.150	0.860	0.0006
Freq:Canon.	1	75	14.813	<0.001***	0.042
Group:Freq.:Canon.	2	75	6.852	<0.001***	0.039

The output showed a main effect of Frequency ($F(1,75) = 130.006, p = 0.001$) as well as canonicity ($F(1,75) = 6.059, p = 0.001$) and an interaction between frequency and canonicity ($F(1,75) = 14.813, p = 0.001$) indicating a differential effect of canonicity based on the lexical frequency of the noun referent. Furthermore, the 3-way interaction between frequency, canonicity, and group was significant ($F(2,75) = 6.85, p = 0.001$). Separate analyses per subgroup were conducted to further explore these interactions.

Further analysis indicated no effect of canonicity or canonicity and frequency interaction in the monolingual group. The statistical analysis of both bilingual groups indicated an interaction of frequency and canonicity. After carrying out a pairwise comparison for each bilingual subgroup the results showed that the more Spanish-dominant bilingual group was more accurate with high-frequency nouns, revealing a positive effect of transparency, although no significant effect of Canonicity was found in the low-frequency nouns. Similar to the more Spanish-dominant bilinguals, the less Spanish-dominant group was more accurate with high-frequency transparent nouns, however, for low-frequency items, this group indicated a trend to be less accurate with transparent nouns.

7. DISCUSSION

The present investigation examined the role of noun canonicity in the lexical retrieval

of gender features in oral production. Gender correlated endings and lexical frequency were manipulated in our study to provide further evidence on the role of predictive nominal endings in monolingual speakers and HSs of Spanish that differ in their amount of exposure to Spanish while growing up.

Regarding research question 1, results were partially compatible with the Reliable Cue Hypothesis (Bates *et al.* 1996; Taft and Meunier 1998; Caffarra *et al.* 2015) since there was a main effect of canonicity, a facilitatory effect of noun transparency, in the Experiment 1 RTs analysis and an interaction of canonicity and frequency in Experiment 2 RTs analysis of monolingual speakers of Spanish. Furthermore, bilingual speakers showed a main effect of canonicity in the accuracy results of both experiments. Therefore, noun phonological cues were a reliable cue when accessing gender, however, these cues were not the most reliable information to access gender since gender could also be retrieved vis-à-vis an abstract lexico-syntactic route in line with the predictions of the Two-Route Hypothesis (Gollan and Frost 2001). According to this hypothesis, a main effect of noun frequency (noun lexico-syntactic information) and canonicity or an interaction between canonicity and frequency was predicted, possibly resulting in a reduced frequency effect in transparent nouns. As expected, a main effect of both frequency and canonicity for monolinguals was observed in the Experiment 1 RTs analysis, indicating faster naming times for transparent nouns and for high-frequency nouns. In Experiment 2, an interaction of canonicity was observed for monolinguals in the RTs analysis and further analyses indicated that monolingual speakers were faster with adjectives in low-frequency transparent noun trials than with low-frequency opaque trials indicating a facilitatory effect of noun transparency. For bilingual speakers, both groups had a main effect of canonicity in Experiment 1 accuracy analysis; that is, heritage speakers were more accurate with transparent nouns. In Experiment 2, an interaction with frequency and canonicity was observed and further analyses indicated that both bilingual groups were more accurate with high-frequency transparent nouns (D+ bilinguals reaching significance). These results clearly indicate

that participants, both monolingual speakers and heritage speakers of Spanish, make use of both the lexico-syntactic and also word-form information in line with studies that found evidence of these two routes to access gender (Caffarra *et al.* 2014) and corroborating the Two-Route hypothesis proposed by Gollan and Frost (2001).

Regarding research question 2, results indicated transparency effects in both bilingual groups' accuracy analysis in both experiments, in line with previous offline studies (Hur *et al.* 2020; Alarcon 2011; Montrul *et al.* 2013). Unlike previous online studies (Montrul *et al.* 2014), results found a lack of canonicity effects for the bilingual group in the RTs analysis of Experiment 1 and for the more Spanish-dominant HSs in Experiment 2. Monolingual speakers showed a canonicity effect in both experiments RTs analysis indicating faster reaction times with nouns overtly marked for gender. As suggested above, bilingual speakers' divided frequency of use of both languages (Gollan 2008) or/and the lack of noun phonological information in English nouns might lead to a weaker activation of overt gender-correlated cues in Spanish as seen by the lack of canonicity effects in the RTs analysis in both bilingual groups. Regarding the unanticipated results of D- bilinguals in Experiment 2 RTs analysis, it is worth mentioning that noun-adjectival agreement was cognitively costlier for all groups but, particularly, for the D- bilinguals. For this group, data points outside the 3000 ms. window, non-recorded responses, skipped trials, inaccurate or non-identifiable responses constituted 49% of the experimental stimuli (versus 28% for D+ bilinguals and 17% for monolinguals). For this reason, the different outcome between the HSs groups in Experiment 2 could be derived from inappropriate experimental methodology. Despite this methodological inadequacy, the bilingual speakers' error rates analysis clearly indicated that, similar to monolingual speakers of Spanish, both heritage speakers' groups seem to be utilizing overt gender marking cues.

Regarding research question 3, an alternative view of the role of noun phonological cues was proposed based on the studies carried out with regular and irregular verbal morphology in English, the Dual Mechanism Model (Pinker

1990, 1999). In line with previous studies testing this hypothesis on gender agreement in Portuguese (Resende and Mota 2017), the present investigation did not fully meet the predictions of the Dual Mechanism Model. The present investigation found that participants were faster with low-frequency transparent nouns than with low-frequency opaque nouns, in line with the predictions of the model. However, no evidence was found with the high-frequency opaque noun trials.

8. CONCLUSION

In the present investigation monolinguals speakers and Spanish-English bilinguals that differ in Spanish dominance named pictures that represent canonical and non-canonical nouns (opaque and transparent) for the purpose of determining how gender-correlated endings are accessed and utilized across different linguistic populations. Results indicated that predictive nominal endings are reliable cues to access gender (Bates *et al.* 1996; Taft and Meunier 1998), however, not the most reliable cue since speakers utilized both lexico-syntactic and the word-form information to access gender as predicted by the Two-route Hypothesis (Gollan and Frost 2001).

It can be deduced from the findings that, even though both groups seem to utilize gender-correlated endings when accessing gender, predictive nominal endings might be used differently by monolingual speakers and HSs of Spanish as seen by the lack of canonicity effects in the RTs analysis in the bilingual group. These findings are valuable in that they provide information that contributes to our understanding of the cognitive processes in HSs adult bilingualism. Particularly, bilingual speakers of Spanish that, despite having acquired Spanish from childhood and in a naturalistic way, purportedly differ from the monolingual norm in their linguistic performance. The present investigation does not reveal reliable differences between the two HSs bilingual groups in terms of whether dominance impacts the use of gender-marking phonemes. Part of this outcome might be due to some methodological shortcomings, such as the range of low-frequency nouns included. Moreover, the sample size of heritage

speakers that was split by dominance might have affected the results of the present investigation. Future research should include a larger group of heritage speakers in each dominance group and, additionally, exclude low-frequency items in the stimuli to exclusively observe canonicity effects.

ACKNOWLEDGMENTS

I deeply thank Dr. Valdés Kroff for his invaluable assistance in this research. Moreover, this research would not have been possible without the help of Dr. Ana Margarita Haché, Dr. Francisco Cruz, Rafaela Carrasco, and Teresa Ureña from Centro de Desarrollo Curricular, the Spanish department from Pontificia Universidad Católica Madre y Maestra (Santiago de los Caballeros, Dominican Republic) who provided me with the support and the facilities to collect the monolingual data in the Dominican Republic.

REFERENCES

- Alarcón, I. V. (2011). Spanish gender agreement under complete and incomplete acquisition: Early and late bilinguals' linguistic behavior within the noun phrase. *Bilingualism: Language and Cognition*, 14(3), 332-350. <https://doi.org/10.1017/S1366728910000222>
- Alarcos, E. (1999). *Gramática de la lengua española*. Espasa Calpe.
- Badecker, W., Miozzo, M., & Zanuttini, R. (1995). The two-stage model of lexical retrieval: Evidence from a case of anomia with selective preservation of grammatical gender. *Cognition*, 57(2), 193-216. [https://doi.org/10.1016/0010-0277\(95\)00663-J](https://doi.org/10.1016/0010-0277(95)00663-J)
- Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Perception & Psychophysics*, 58(7), 992-1004. <https://doi.org/10.3758/BF03206827>
- Biran, M., & Friedmann, N. (2012). The representation of lexical-syntactic information: Evidence from syntactic and lexical retrieval impairments in aphasia. *Cortex*, 48(9), 1103-1127. <https://doi.org/10.1016/j.cortex.2011.05.024>

- Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form* (Vol. 9). John Benjamins Publishing. <https://doi.org/10.1075/tsl.9>
- Bybee, J. (1995). Regular morphology and the lexicon. *Language and cognitive processes*, 10(5), 425-455. <https://doi.org/10.1080/01690969508407111>
- Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: evidence from the tip-of-the tongue phenomenon. *Cognition*, 64(3), 309-343. [https://doi.org/10.1016/S0010-0277\(97\)00031-0](https://doi.org/10.1016/S0010-0277(97)00031-0)
- Carroll, S. (1989). Language acquisition studies and a feasible theory of grammar. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 34(4), 399-418. <https://doi.org/10.1017/S0008413100024300>
- Caffarra, S., Janssen, N., & Barber, H. A. (2014). Two sides of gender: ERP evidence for the presence of two routes during gender agreement processing. *Neuropsychologia*, 63, 124-134. <https://doi.org/10.1016/j.neuropsychologia.2014.08.016>
- Caffarra, S., Siyanova-Chanturia, A., Pesciarelli, F., Vespignani, F., & Cacciari, C. (2015). Is the noun ending a cue to grammatical gender processing? An ERP study on sentences in Italian. *Psychophysiology*, 52(8), 1019-1030. <https://doi.org/10.1111/psyp.12429>
- Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop Shopping for Spanish Word Properties. *Behavior Research Methods*, 45, 1246-1258. <https://doi.org/10.3758/s13428-013-0326-1>
- Gollan, T. H., & Frost, R. (2001). Two routes to grammatical gender: Evidence from Hebrew. *Journal of Psycholinguistic Research*, 30(6), 627-651. <https://doi.org/10.1023/A:1014235223566>
- Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of memory and language*, 58(3), 787-814. <https://doi.org/10.1016/j.jml.2007.07.001>
- Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & cognition*, 33(7), 1220-1234. <https://doi.org/10.3758/BF03193224>
- Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, 140 (2), 186. <https://doi.org/10.1037/a0022256>
- Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish-English bilinguals. *Bilingualism: Language and Cognition*, 15(3), 594-615. <https://doi.org/10.1017/S1366728911000332>
- Gonon, M. H., Bruckert, R., & Michel, F. (1989). Lexicalization in an anomic patient. *Neuropsychologia*, 27(4), 391-407. [https://doi.org/10.1016/0028-3932\(89\)90047-X](https://doi.org/10.1016/0028-3932(89)90047-X)
- Goodglass, H., Kaplan, E., & Weintraub, S. (1983). *Boston naming test*. Lea & Febiger.
- Green, J. N. (1988). Spanish. In M. Harris, & N. Vincent (Eds.). *The Romance languages* (pp. 79-130). Oxford University Press.
- Grosjean, F. (2008). *Studying bilinguals*. Oxford University Press.
- Harris, J. W. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, 22(1), 27-62.
- Hur, E., Lopez Otero, J. C., & Sanchez, L. (2020). Gender Agreement and Assignment in Spanish Heritage Speakers: Does Frequency Matter? *Languages*, 5(4), 48. <https://doi.org/10.3390/languages5040048>
- Kulke, F., & Blanken, G. (2001). Phonological and syntactic influences on semantic misnamings in aphasia. *Aphasiology*, 15(1), 3-15. <https://doi.org/10.1080/02687040042000070>
- Jackendoff, R. (2002). English particle constructions, the lexicon, and the autonomy of syntax. *Verb-particle explorations*, 67-94. <https://doi.org/10.1515/9783110902341.67>

- Lloret, M. R., & Viaplana, J. (1998). Variació morfofonològica. Variants morfològiques. *Caplletra. Revista Internacional de Filologia*, 25, 43-62.
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. [https://doi.org/10.1044/1092-4388\(2007/067\)](https://doi.org/10.1044/1092-4388(2007/067))
- Martínez-Gibson, E. A. (2011). A comparative study on gender agreement errors in the spoken Spanish of heritage speakers and second language learners. *PortaLinguarum: revista internacional de didáctica de las lenguas extranjeras*, (15), 177-193. <https://doi.org/10.30827/Digibug.31991>
- Marslen-Wilson, W. D. (2007). Morphological processes in language comprehension. In G. Gaskell (Ed.). *Oxford handbook of psycholinguistics* (pp. 175-193). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198568971.013.0011>
- Meyer, A. S., & Bock, K. (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation? *Memory & Cognition*, 20(6), 715-726. <https://doi.org/10.3758/BF03202721>
- Miozzo, M., & Caramazza, A. (1997). The retrieval of lexical-syntactic features in tip-of-the-tongue states. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 17, 1-14.
- Montrul, S., Davidson, J., De La Fuente, I., & Foote, R. (2014). Early language experience facilitates the processing of gender agreement in Spanish heritage speakers. *Bilingualism: Language and Cognition*, 17(1), 118-138. <https://doi.org/10.1017/S1366728913000114>
- Montrul, S., De La Fuente, I., Davidson, J., & Foote, R. (2013). The role of experience in the acquisition and production of diminutives and gender in Spanish: Evidence from L2 learners and heritage speakers. *Second Language Research*, 29(1), 87-118. <https://doi.org/10.1177/0267658312458268>
- Montrul, S. A. (2012). Is the heritage language like a second language? *Eurosla Yearbook*, 12(1), 1-29. <https://doi.org/10.1075/eurosla.12.03mon>
- Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, 58(3), 503-553. <https://doi.org/10.1111/j.1467-9922.2008.00449.x>
- Montrul, S., & Potowski, K. (2007). Command of gender agreement in school-age Spanish-English bilingual children. *International Journal of Bilingualism*, 11(3), 301-328. <https://doi.org/10.1177/13670069070110030301>
- Navarrete, E., Basagni, B., Alario, F. X., & Costa, A. (2006). Does word frequency affect lexical selection in speech production? *Quarterly Journal of Experimental Psychology*, 59(10), 1681-1690. <https://doi.org/10.1080/17470210600750558>
- O'Rourke, P. L., & Van Petten, C. (2011). Morphological agreement at a distance: Dissociation between early and late components of the event-related brain potential. *Brain research*, 1392, 62-79. <https://doi.org/10.1016/j.brainres.2011.03.071>
- Pinker, S. (1991). Rules of language. *Science*, 253(5019), 530-535. <https://doi.org/10.1126/science.1857983>
- Pinker, S. (1999). How the mind works. *Annals of the New York Academy of Sciences*, 882(1), 119-127. <https://doi.org/10.1111/j.1749-6632.1999.tb08538.x>
- Resende, N. C. A., & Mota, M. B. (2017). A behavioral study to investigate the processing routes of grammatical gender in Brazilian Portuguese. *Revista de Estudos da Linguagem*, 25(3), 1367-1395. <https://doi.org/10.17851/2237-2083.25.3.1367-1395>
- Rosinski, R. R., Golinkoff, R. M., and Kukish, K. S. (1975). Automatic semantic processing in a picture-word interference task. *Child Development*, 247-253. <https://doi.org/10.2307/1128859>
- Sá-Leite, A. R., Luna, K., Tomaz, Â., Fraga, I., & Comesaña, M. (2022). The mechanisms underlying grammatical gender selection in language production: A meta-analysis of the gender congruency effect. *Cognition*, 224, 105060. <https://doi.org/10.1016/j.cognition.2022.105060>

- Sá-Leite, A. R., Fraga, I., & Comesaña, M. (2019). Grammatical gender processing in bilinguals: An analytic review. *Psychonomic Bulletin & Review*, 26(4), 1148-1173. <https://doi.org/10.3758/s13423-019-01596-8>
- Silva-Corvalán, C. (1994). *Language Contact and Change: Spanish in Los Angeles*. Clarendon Press/Oxford University Press.
- Schmid, M. S., Köpke, B., Keijzer, M., & Weilemar, L. (Eds.). (2004). *First language attrition: Interdisciplinary perspectives on methodological issues* (Vol. 28). John Benjamins Publishing. <https://doi.org/10.1075/sibil.28>
- Taft, M., & Meunier, F. (1998). Lexical representation of gender: A quasiregular domain. *Journal of Psycholinguistic research*, 27(1), 23-45. <https://doi.org/10.1023/A:1023270723066>
- Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and cognition*, 4(2), 105-122. <https://doi.org/10.1017/S1366728901000220>
- Urrutia, M., Domínguez, A., & Álvarez, C. J. (2009). Gender activation in transparent and opaque words. *Psicothema*, 21(1), 1-8.
- Valdés, G. (2005). Bilingualism, heritage language learners, and SLA research: Opportunities lost or seized? *The Modern Language Journal*, 89(3), 410-426. <https://doi.org/10.1111/j.1540-4781.2005.00314.x>
- Valdés, G. (2014). Heritage language students: Profiles and possibilities. In T. Wiley et al. (eds.). *Handbook of heritage, community, and native American languages in the United States* (pp. 41-49). Routledge.
- Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological science*, 8(4), 314-317. <https://doi.org/10.1111/j.1467-9280.1997.tb00444.x>