2025. Proc Ling Soc Amer 10(1). 5940. https://doi.org/10.3765/plsa.v10i1.5940.

Gender assignment strategies in Spanish-English mixed noun phrases

Anna Knall*

Abstract. This study explores how Spanish-English bilinguals in California assign grammatical gender when code-switching between a Spanish determiner and an English noun. In bilingual speech, these moments of cross-linguistic contact offer a window into how speakers navigate competing grammatical systems. Using data from a semi-spontaneous picture description task, the analysis draws on 649 mixed noun phrase tokens to examine whether bilinguals rely on analogical gender, semantic cues, phonological endings, number, or determiner type when choosing between masculine and feminine determiners. While masculine determiners were strongly favored overall, the results show that analogical gender plays a meaningful role—feminine determiners were more likely to appear when the English noun had a grammatically feminine Spanish translation. Other factors did not reach statistical significance. These findings support the view that gender assignment in code-switching reflects structural sensitivity to the grammatical systems of both languages, as well as variability shaped by individual bilingual experience.

Keywords. code-switching; bilingualism; grammatical gender; mixed noun phrases; analogical gender.

1. Introduction. Code-switching (CS), the seamless alternation between languages in natural discourse, provides the ideal context for observing bilingual behavior when faced with cross-linguistic barriers (Valdés Kroff 2016). At the morphosyntactic level, there is a growing body of literature on code-switching within mixed noun phrases and the strategies that multilinguals employ when the grammars of their languages diverge (Cisneros et al. 2023). One point of interest involves language dyads that exhibit code-switches between determiners and nouns, typically when the first language has a grammatical gender classification system, while the second language does not. For example, when engaging in CS, a Spanish-English bilingual has a choice between masculine and feminine determiners to define a word in a specific sentence, as exemplified in (1a-b).

(1) Spanish (Cruz 2022: 585)

a.	Ya	estamos	en	**el**	plane.
	Already	1PL-be-PRS	in	the-M	plane.
	'We are already on the plane'				
b.	Ya	estamos	en	**la**	plane.
	Already	1PL-be-PRS	in	the-F	plane.
	'We are already on the plane'				

In other words, bilinguals make a conscious decision to use one of the gender-marking determiners, thus associating a noun with a morphological gender category that is non-existent in English. Across a series of studies, several unique gender assignment strategies have been

* I would like to thank Dr. Ernesto Gutiérrez Topete for granting me access to his corpus and Dr. Justin Davidson for his continued mentorship and for advising this project.
Author: Anna Knall, University of California, Berkeley (annaknall@berkeley.edu)

evidenced: semantic gender, translation equivalent, default, and phonological shape (Bellamy & Parafita Couto 2022).

In the context of gender assignment within mixed Spanish determiner-English noun phrases (DET-noun), this study aims to identify which of those four strategies bilinguals in California employ when applying Spanish gender to English nouns. Previous literature has focused on various communities, ranging from Southern Arizona to New Mexico and Northern Belize, with little emphasis on Californian code-switching communities. Additionally, there is a lack of consistency in grammatical gender assignment which could fluctuate depending on the identity of the community, the structure of the languages used in code-switching, and experimental designs. As such, more studies like the present one are necessary for a more holistic conclusion on what might account for the variation in gender assignment. This study builds on previous research methodologies that adopted natural speech analysis, using recordings from a corpus of puzzle activities that elicited spontaneous speech while encouraging the use of code-switched language.

I address the following questions, in part modeling those of Cisneros et al. (2023), but applying them to the Californian Spanish-English code-switching context: What strategies do California Spanish-English bilinguals prefer when assigning gender to English nouns in code-switched speech? More specifically, do they use semantic gender in accordance with an animate noun's biological referent? Regarding inanimate nouns, do they employ the translation equivalent, phonological shape or default gender assignment strategies?

2. Grammatical gender systems and code-switching: English and Spanish. In both English and Spanish, animate nouns might be categorized in terms of their biological gender, referred to as semantic gender. For example, the word 'mother' in English might use gendered pronouns such as 'she/her' because of its inherent feminine connotation; this would also be the case for the Spanish equivalent. However, the two languages diverge in their classification of inanimate objects. Spanish, similar to all languages in the Romance family, has a grammatical gender classification system for all nouns, animate and inanimate. Through binary categories, feminine (F) and masculine (M), the Spanish language requires that nouns utilize modifiers in accordance with their grammatical gender. For example, the word 'flower' (*flor*) belongs to the feminine classification and would be accompanied by feminine modifiers—such as the definite article (*la*) or the indefinite (*una*). Similarly, the word 'glass' (*vaso*) belongs to the masculine category and could be modified by the definite article (*el*) or the indefinite (*un*).

One indicator of gender is the word-ending phoneme: canonically, nouns ending in /o/ are masculine 99.87% of the time, as in '*el libro'* (the-M book), while those ending in /a/ are feminine 96% of the time, as in '*la mesa'* (the-F table) (Teschner & Russell 1984). Beyond these canonical patterns, additional phonological schemas exist; for example, nouns ending in <-z, -ión, -d, -dad, -tad, -umbre> often denote feminine gender, while those ending in <-r, -l, -n, -e> tend to be masculine (Trawick & Bero 2022). Although these rules are generally reliable, there are notable exceptions—particularly nouns ending in <-ma, -ta, -pa>, such as '*mapa'* (map) and '*problema*' (problem), which appear canonically feminine but are in fact masculine, and as such, take masculine modifiers. Conversely, some feminine nouns follow what seems like a masculine pattern, with a salient example being '*mano*' (hand) which ends in /o/.

Morphological characteristics such as grammatical gender are important to account for in code-switching (CS) research because they can highlight how bilinguals confront cross-linguistic barriers (Valdés Kroff 2016). In language dyads like English-Spanish, non-gendered and gendered respectively, inserting words from the former into the latter requires the assignment of

grammatical gender. Some morphosyntactic structures (i.e. mixed determiner-noun phrases) foster better environments for CS, whereas instances of mixed noun-adjective phrases are less common, if they occur at all (Otheguy & Lapidus 2003; Balam & Parafita 2019). This phenomenon, as Otheguy and Lapidus (2003) explain, might be a result of an avoidance strategy to maintain the syntactic order of attributive adjectives which are pre-nominal in English, but post-nominal in Spanish. Balam and Parafita (2019) support those findings through their analysis of sociolinguistic interviews from Northern Belize. In said interviews, participants avoided the production of mixed attributive adjective-noun phrases, highlighting that CS is more likely to occur in simplified contexts when cross-linguistic barriers have greater overlap. Therefore, since the syntactic placement of determiners in both Spanish and English is pre-nominal, CS is frequently observed in DET-noun structures.

3. Gender assignment strategies in mixed determiner-noun phrases. Factors mentioned above, such as a noun's semantic or grammatical gender, as well as phonological shape, might account for different assignment strategies in bilingual mixed noun phrases. Bellamy and Parafita (2022) explore four of these assignment strategies: semantic gender, translation equivalent, default, and phonological shape.

3.1. SEMANTIC GENDER ASSIGNMENT STRATEGY. The semantic gender strategy refers to the influence of a noun's biological or presupposed gender on the gender of a Spanish determiner that modifies an English word (Cruz 2022). In code-switched speech, this can be exemplified by attributing a gendered determiner to a noun that traditionally carries an inherent gender outside of a specific context, as in '*el* [the-M] uncle,' or when referring to someone specific, such as a biologically female pet (e.g., '*la* [the.F] dog').

When analyzing the Russian-English dyad, Chirsheva (2009) found that nouns denoting humans always had the right gender assignment in code-switched phrases. However, literature on Spanish-English CS has yielded varying results. Cruz (2021) found semantic gender to be an accurate predictor for gender assignment in mixed DET-noun phrases in this specific community. In a later study, Cruz (2022) corroborated those results through experimental tasks which used vocabulary from the same CESA corpus. Conversely, the results of Valdés Kroff (2016) using the Bangor Miami Corpus found that immediate indication of biological gender only applied to a subset of his data.

3.2. TRANSLATION EQUIVALENT ASSIGNMENT STRATEGY. Compared to the semantic gender strategy based on biological gender, translation equivalence, also called analogical gender, refers to using the determiner that corresponds in gender with the Spanish translation of the English noun. That means one would expect to observe the code-switched phrase '*la* [the.F] table' (*la mesa*) instead of '*el* [the.M] table,' which would not be in agreement with the Spanish grammatical gender.

Despite a majority of literature finding a preference for defaulting to a masculine determiner regardless of analogical gender, there is evidence for analogical gender as an accurate assignment strategy. According to Bellamy and Parafita Couto (2022), bilinguals whose first language has a grammatical gender classification tend to favor the translation equivalent strategy. A preference for feminine determiners also appears in the *New Mexico Spanish-English Bilingual Corpus*, specifically in instances where the code-switched noun was feminine in Spanish or ambiguous due to dialectal or contextual variance. One such example of dialect variation is the production of the word 'college' and the uncertainty on whether the speaker was thinking of the Spanish word '*colegio*' [M] or '*universidad*' [F] (Trawick & Bero 2022).

Furthermore, Trawick and Bero (2022) also concluded that despite many words exhibiting a masculine determiner in front of an analogically feminine word, the use of a feminine determiner increased the chances that the subsequent English word would be analogically feminine. The fact that there are no instances in which a feminine determiner was used for a noun with masculine analogical gender indicated a "co-activation of the bilingual lexico-semantic system" which increased the likelihood of applying the feminine gender in a mixed DET-noun phrase, even if the result is defaulting to masculine (Trawick & Bero 2022: 192). These findings are also in alignment with other literature (Valdés Kroff 2016; Beatty-Martinez & Dussias 2017) observing that a code-switch is less likely to occur after a feminine determiner; however, the determiner will almost categorically precede a feminine word whether produced directly in Spanish or in its analogical English translation. Because of this, masculine determiners do not provide reliable cues for either code-switches nor the analogical gender of a CS (Valdés Kroff 2016).

Another important aspect to note when considering the translation equivalent assignment strategy is the context in which data were examined (i.e. experimental versus natural speech). In some cases, empirical data gathered from experimental tasks (such as forced-switch or picture naming tasks) demonstrated a preference for the analogical gender assignment strategy (Cisneros et al. 2023; Fairchild & Van Halen 2017). Chirsheva (2009) attributes this to a possible increase in the metalinguistic awareness of the gender of nouns in less spontaneous, less naturalistic tasks. Not only might experimental tasks encourage participants to be more attentive to the gender of nouns, but they also yield significantly more CS than sociolinguistic interviews, showing that mixed DET-noun phrases are not as common in natural speech (Cisneros et al. 2023).

3.3. DEFAULT ASSIGNMENT STRATEGY. Another strategy for code-switching assignment is defaulting to a particular gender, a practice often shaped by the established norms and practices within a given code-switching community (Bellamy & Parafita Couto 2022; Trawick & Bero 2022). As further observed by Bellamy and Parafita (2022), the default strategy tends to be favored by habitual code-switchers, with a distinct inclination toward the masculine gender. Cruz (2021) aligns with this trend, highlighting the prevalence of masculine default strategies in code-switching assignments. Within the same community in Southern Arizona, the results from experimental tasks substantiated those earlier observations, identifying a prevailing preference for the masculine gender in DET-noun switches in constructions such as '*el* [the-M] wand' (Cruz 2022: 588). In line with the community-based argument for a default strategy, Cruz (2022) also suggested that this preference was closely linked to the unique code-switching norms within the Southern Arizona community. Additionally, the study found no solid connections between semantic, analogical gender, or phonological shape—the latter of which will be discussed in detail in the next section.

Balam, Lakshmanan, and Parafita's (2021) research into simultaneous Spanish-English bilingual children from Miami Dade, Florida also contributed to the growing body of evidence supporting an overwhelming preference for masculine defaults. One explanation they offer for the phenomenon is the possibility that bilinguals simplify cross-linguistic processes.

3.4. PHONOLOGICAL SHAPE ASSIGNMENT STRATEGY. The last prominent gender assignment strategy is based on the ending of a specific word that, when code-switched, resembles phonological gender schemas of the gendered language. In terms of Spanish-English CS, an example would be pairing the English word 'arena' (*estadio* in Spanish) with a feminine determiner because the ending <-a> is canonically feminine based on Spanish gender schemas.

Depending on the language dyad, there has been evidence of phonological shape as both an accurate and inaccurate assignment strategy. For example, in Russian-English bilinguals, Chirsheva (2009) found phonological cues to be the preference, especially in natural speech. Similarly, Parafita Couto et al. (2015) studied Spanish-Basque mixed DET-noun phrases, Basque being a grammatically genderless language with a high frequency of words ending in /-a/. Accordingly, the use of feminine determiners was observed at higher rates than masculine ones. Adopting an argument for cross-linguistic effects, this phenomenon was considered as a direct result from Spanish gender schemas (Parafita Couto et al. 2015, Trawick & Bero, 2022).

When exclusively investigating Spanish-English pairs, Cruz (2021) did not find any evidence of phonological shape-based assignment; however, the corpus utilized from South Arizona, CESA, did not present many examples of nouns that patterned similarly to phonologically Spanish-like feminine words. Conversely, Trawick & Bero (2022) found that in the *New Mexico Spanish-English Bilingual Corpus*, words with canonically feminine endings (e.g. soda) and those with masculine endings (e.g. bill) both had corresponding determiner genders about 90% of the time. Those nouns that did not resemble a Spanish ending or had ambiguous endings, often adopted masculine determiners (Trawick & Bero 2022). This could point to the idea that the masculine default strategy discussed above might not only be a community-driven strategy but also a result of phonological constraints of English words. Balam et al. (2021) also poses this open question, recognizing that the infrequent use of the feminine canonical <-a> ending in English could contribute to a higher rate of masculine gender assignment.

Overall, despite the conflicting evidence on preferences for a particular gender assignment strategy, previous literature in this review mostly agrees that these strategies differ among communities, language dyads, and investigation methodologies. For this reason, the present study extends existing research on the Spanish-English language dyad, drawing observations from a bilingual corpus that has not been analyzed for mixed DET-noun constructions.

4. Data. The data for this study is drawn from a corpus collected by Gutiérrez Topete (2023). In this study, the participants were Spanish-English bilingual speakers of ages 18 years or older at the time of their participation. More specifically, participants were considered bilingual if they could (1) hold a conversation on various topics in both English and Spanish, and (2) could read in both languages at a basic level (i.e. 4th grade level or higher). When self-assessing their proficiency, participants on average indicated an English dominance, but had overall high proficiency in both Spanish and English across all modes of communication—on a 0 to 6 scale, 4.5 to 5.5 and 5.6 to 5.8 respectively (Gutiérrez Topete 2023).

The present study analyzes tokens from the puzzle-task subcorpus of Gutiérrez Topete (2023). In this activity, participants were asked to spot the difference within an image that was designed to primarily include objects whose translation was a non-cognate in English and Spanish, and whose name began with a voiceless stop-consonant, which was of relevance to the original study. Each recording was between 10-20 minutes long and included samples such as "turkey" (*pavo*) and "peel" (*cáscara*). The data from this specific activity was chosen for the current study for two reasons: (1) it prompted participants to engage in code-switching more naturally, and (2) all participants described the same image which will yield significantly greater token counts for specific nouns. Ultimately, the data of 57 out of the 60 participants were analyzed, having removed those who did not produce any tokens of mixed DET-noun phrases.

4.1. VARIABLES. In the collection phase, I used Praat to annotate every example of a code-switch between a Spanish determiner and an English noun (Boersma & Weenink 2024). The independent variables (IVs) were: (1) the noun's grammatical gender based on its Spanish translation (masculine, feminine, and ambiguous), where ambiguous meant that the translation could be either masculine or feminine based on dialectal variation, (2) number (singular vs. plural), (3) phonological ending (feminine, masculine, or ambiguous for those English words that do not resemble Spanish endings), (4) semantic gender (biologically gendered referent yes vs. no), (5) animacy (animate vs. inanimate) and (6) determiner type (definite article, indefinite article, or adjective determiner). The difference between semantic gender and animacy was included because many of the "animate" nouns were animals that did not have a clear semantic gender that could be inferred solely based on the picture. The dependent variable (DV) was the binary gender of the determiner (feminine or masculine).

Initially, 880 tokens of mixed DET-noun phrases were collected from 57 recordings of the Gutiérrez Topete (2023) corpus. Before the statistical analysis, ambiguous tokens were removed from the dataset, resulting in 649 final tokens. One category of ambiguous tokens referred to words that have multiple translations with conflicting grammatical gender in Spanish; example (2) below shows the use of the word 'fridge' which could be translated as '*el* [the.M] *refrigerador*' or '*la* [the.F] *nevera*" (among other terms depending on the Spanish dialect). Another category was that of words that referred to animate nouns, but whose semantic gender was ambiguous; in (3), it is unclear how the speaker would determine if the 'crab' is biologically masculine or feminine. The objects in the picture that might have portrayed animals (e.g. a puppet that was a lion), were included as tokens and were coded as not having a semantic gender but being animate. After removing ambiguous tokens, a fixed-effects logistic regression was run in R (version 4.4.3), using the packages car, emmeans, and lmtest, to examine the effects of each independent variable as main effects on the dependent variable (R Core Team 2025).[1]

(2) en bajo **del fridge** se ve el perro. (Participant 2)
in bottom **of the-M fridge** 3SG-REFL-see-PRS the dog
'Under the fridge, the dog can be seen.'

(3) **el crab** está como entrando en la olla. (Participant 4)
the-M crab 3SG-be-PRS like enter-PROG in the-F pot
'The crab is like entering the pot.'

5. Results. When looking at the overall distribution of the analogical gender versus its correct production, there is a clear dominance of the masculine determiner (see Figure 1). Of the total 649 tokens, 62.7% (N = 407) had a masculine Spanish equivalent and 37.3% (N = 242) had feminine translations. However, those numbers shift for the gender for the determiner produced, where 80.4% (N = 522) were defined by a masculine determiner and only 19.6% (N = 127) by a feminine determiner. The change in the determiner distribution is significant enough to show an overall preference of producing masculine determiners.

The regression model identified analogical gender as a statistically significant predictor for the gender of the determiner produced in front of a noun (see Table 1). The beta coefficients in Table 1 indicate the log-odds of feminine determiner production. The intercept represents a

[1] A mixed-effects logistic regression model with speaker as a random effect was initially run, but it failed to converge. A fixed-effects model was accordingly used instead.

plural animate noun with feminine analogical gender, an adjective determiner, and with an ambiguous phonological shape.

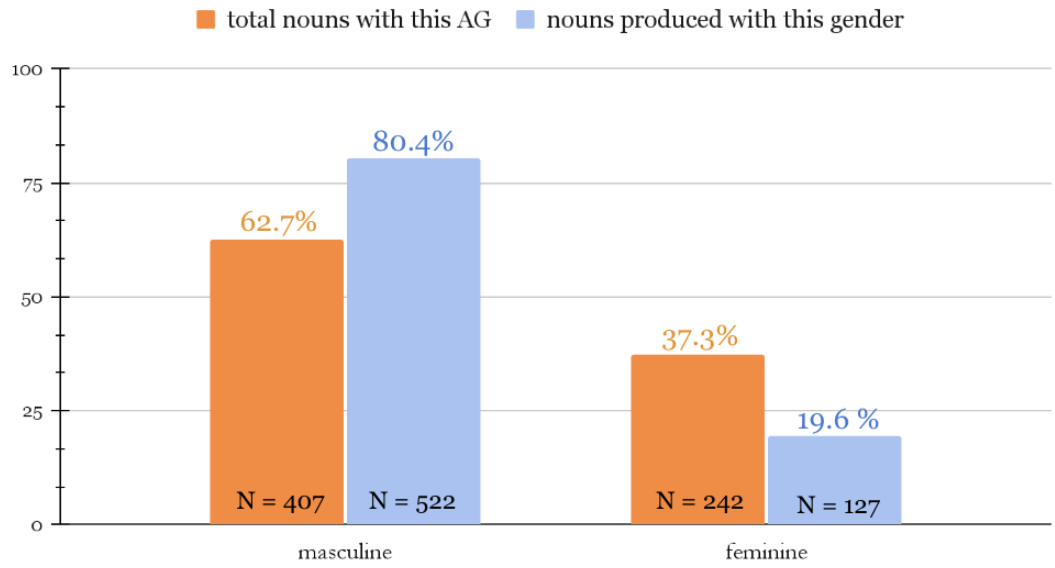

Figure 1. Distribution of total items by analogical gender (AG) and their produced gender

	β coefficient (in log-odds)	Standard Error	z value	p value
Intercept	-16.7556	542.1394	-0.031	0.9753
Analogical Gender (masculine)	-2.5055	0.2784	-8.998	<.001***
Semantic Gender (yes)	15.2020	542.1385	0.028	0.9776
Animacy (animate)	14.8708	542.1383	0.027	0.9781
Number (singular)	-0.3737	0.4034	-0.926	0.3543
Determiner (definite)	1.8856	1.0775	1.750	0.0801
Determiner (indefinite)	2.1984	1.0615	2.071	0.0384*
Phonological shape (feminine)	0.1079	0.4037	0.267	0.7892
Phonological shape (masculine)	-0.2920	0.2414	-1.210	0.2263

Table 1. Model predictions of the production of a feminine determiner

The beta coefficients in Table 1 indicate the log-odds of feminine determiner production. The intercept represents a plural animate noun with feminine analogical gender, an adjective determiner, and with an ambiguous phonological shape. Under these reference conditions, the predicted log-odds of producing feminine gender were −16.7556 ($p = .975$). For nouns with a masculine analogical gender, the model predicted a 2.5055 log-odds decrease relative to the intercept ($p < .001$), suggesting a significant reduction in the likelihood of producing feminine gender. The model estimated that indefinite determiners increased the likelihood of feminine gender production by 2.1984 log-odds ($p = .038$) and definite determiners by 1.8856 log-odds ($p = .080$) relative to the adjective baseline. However, a Tukey post-hoc test on determiner types did not yield significant differences, indicating that the observed statistical significance of

determiner type should be regarded as inconclusive. Other factors inferred from the regression table—including semantic gender, animacy, number, and phonological cues—were not significant predictors.

5.1. ANALOGICAL GENDER AS A MAIN EFFECT. Figure 2 visualizes the distribution of the nouns with masculine analogical gender (N = 407) and feminine analogical gender (N = 242) in terms of the gender of the determiner that modified them. Nouns with masculine analogical gender were overwhelmingly produced with masculine determiners 94.8% of the time (N = 386), while only 5.2% (N = 21) were produced with feminine determiners. In contrast, nouns with feminine analogical gender showed a more balanced distribution: 56.2% (N = 136) were produced with masculine determiners, and 43.8% (N = 106) with feminine determiners. Although feminine nouns were produced with masculine determiners more than the feminine equivalent, this asymmetry still suggests that while masculine analogical gender is strongly predictive of masculine gender production, feminine analogical gender is associated with greater variability in determiner choice.

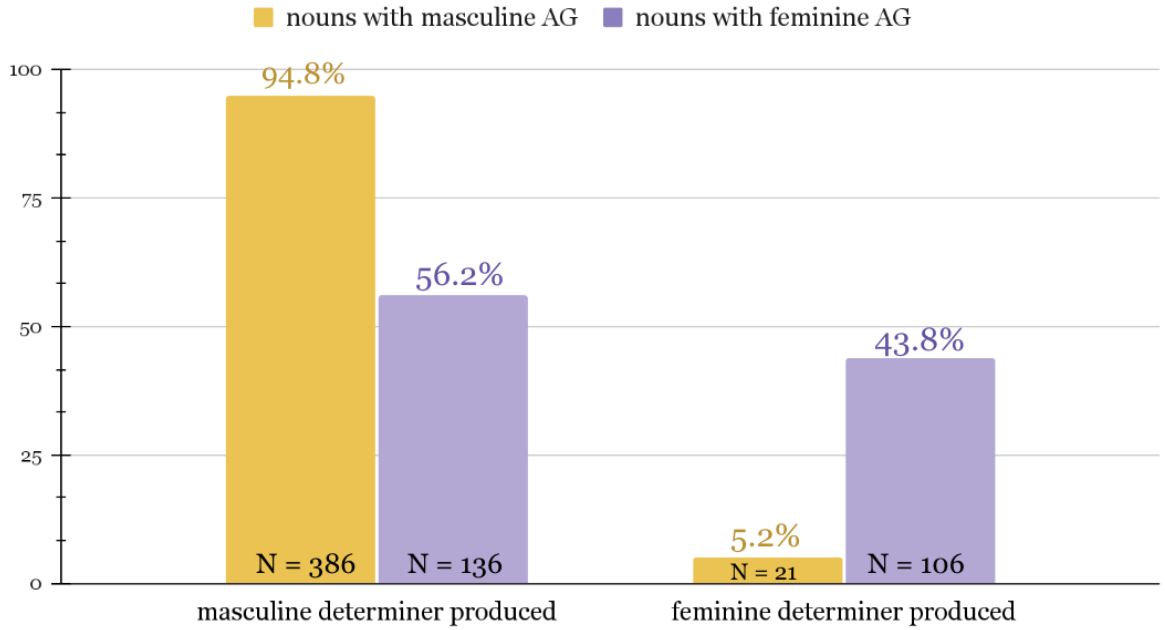

Figure 2. Distribution of produced determiner gender by analogical gender (AG)

Critically, the logistic regression (Table 1) is consistent with a significant effect of analogical gender on gender production ($p < .0001$). The relatively frequent use of feminine determiners with nouns that have feminine analogical gender—compared to the near-categorical masculine production for masculine items—indicates that the appearance of feminine gender in code-switched determiners is not random. Rather, it tends to align with Spanish translation equivalents that bear feminine grammatical gender, reinforcing the influence of analogical gender in bilingual code-switching contexts, even when English lexical items are embedded in Spanish speech.

6. Discussion. Following an analysis of all potential factors identified as relevant in previous literature (i.e., analogical gender, semantic gender, and phonological gender), along with two less commonly discussed variables (i.e., number and determiner type), the present study found a preference for a masculine-default strategy that is nonetheless sensitive to analogical gender.

To address the masculine-default portion of the findings, these results align with earlier literature on bilingual code-switching communities in Southern Arizona (Cruz 2021; Cruz 2022) and New Mexico (Bellamy & Parafita Couto 2022). Two explanations for the masculine default strategy proposed by previous studies are: (1) it is a norm established within the community; and (2) it is more common among habitual code-switchers. Since definitions of "community" can vary in scale, and this study's data were primarily collected from Spanish-English bilinguals in the Bay Area, we might consider that regional context as the relevant community. However, future work should explore how social networks or smaller speech communities may adopt particular code-switching strategies depending on the degree of interaction among their members.

While masculine determiners were more salient across both masculine and feminine nouns, the effect of analogical gender indicates that the production of a feminine determiner increases the likelihood that the English noun following it has a feminine translation equivalent in Spanish. This finding aligns with Trawick & Bero (2022), where analogical gender emerged as one of the most significant predictors, showing that an analogically feminine noun is more likely to be preceded by a feminine determiner. The current study utilized recordings of semi-spontaneous speech, in which participants described images and were encouraged—but not required—to code-switch. This design offers a middle ground between experimental and naturalistic approaches, providing a context rich in opportunities for mixed determiner-noun phrase production without introducing the heightened metalinguistic awareness that often accompanies forced code-switching. At the same time, as participants were prompted to code-switch naturally, there was considerable inter-speaker variation in their frequency of language alternation. Specifically, for mixed determiner-noun phrases, the number of tokens per participant ranged from 1 to 42. This variation suggests that an individual's own code-switching practices may play as important of a role as the task design itself. Measuring frequency may more clearly reflect how experimental tasks diverge from natural speech—tasks that explicitly force code-switches may lead to increased use of analogical gender due to greater awareness of the gender assignment process (Cisneros et al. 2023; Chirsheva 2009; Fairchild & Van Halen 2017). That said, it remains difficult to determine whether code-switching frequency drives the dominance of masculine determiners without comparing natural speech to forced tasks. Future studies should adopt a mixed-method approach—similar to Cisneros et al. (2023)—to compare gender assignment strategies used by the same individuals across different code-switching contexts.

Another explanation for the overall masculine skew could be a simplification of cross-linguistic processes. In the present study, some of the participants made gender agreement mistakes even when the DET-noun pair was completely in Spanish. This could be interpreted as a simplification strategy since speakers were prompted to code-switch, thus possibly applying strategies of cross-linguistic practices to their monolingual speech. On the other hand, gender disagreements could also be a result of an initial incorrect acquisition of the grammatical gender in Spanish. Montrul et al. (2013) found that Spanish native speakers in Mexico do not always present patterns of traditionally considered "perfect" gender assignment. For example, the accuracy of marking '*el gran calle*' (the-M big street) as grammatically incorrect was 92% — meaning that 8% of the native speakers responded that masculine gender is correct for '*calle*' (street) which is morphologically feminine (Montrul et al. 2013). This suggests that assignment strategies of bilinguals should be compared not to what is prescriptively correct by Spanish grammatical schemas, but by the gender patterns that monolinguals possess in natural speech.

In regards to semantic gender, the tokens collected in the present study were not ideal for establishing whether an animate noun had an evident biological gender. Despite the small sample of words that did have a biological referent, the variable was included in the model because previous studies on Spanish-English and Russian-English had found semantic gender to be significant in predicting the gender of the determiner (Chirsheva 2009; Cruz 2021; Cruz 2022). Still, in this dataset, semantic gender was not a significant variable, aligning with Valdés Kroff (2016) who similarly did not find it to have that influential of an impact.

When considering phonological shape as a factor, the present study found no significant evidence for gender assignment based on whether the English words had canonically feminine or masculine endings as per Spanish schemas. This could be a result of English words not patterning as phonologically Spanish-like feminine words, something also observed by Cruz (2021). In the present study, there were only 55 examples of words that could fall under the feminine category, whereas 309 were masculine and 285 ambiguous. Given that the highest number of words patterned similarly to masculine phonological shapes, as was previously hypothesized, the masculine default strategy could also be a result of the constraints of English word-structure.

Lastly, unlike previous research, the current study also ran statistical tests on whether the definiteness of the determiner preceding the noun or the number of the noun could increase the chance of the determiner produced to be feminine. Indefinite determiners marginally increased the odds of feminine gender selection relative to adjective determiners, but pairwise comparisons between the three levels were not statistically significant. This should be further addressed in future research.

7. Conclusion. In this paper, I have investigated the gender assignment strategies implemented by Spanish-English bilinguals from California, when code-switching between a Spanish determiner and an English noun. I conclude that in semi-spontaneous speech where participants were encouraged to code-switch naturally, bilinguals in this larger California community prefer masculine determiners, yet determiner choices are still influenced by analogical gender associations. Considering the near significance of determiner type, there is potential for future directions incorporating more linguistic factors, such as a detailed focus on the level of definiteness of the modifier or the word's syntactic role. Moreover, future research should also include sociolinguistic factors based on language dominance or code-switching frequency, given that code-switching and intense contact between Spanish and English could potentially result in language variation patterns and contact-induced change (Backus 2025). Presently, however, this study showcases CS as a natural feature of bilingual speech that exhibits unique grammatical constraints relative to either of its constituent languages individually. Generalizing gender assignment strategies in bilingualism will require more intense studying of diverse language pairs, highlighting the role of code-switching and typological variation in bilingual language use.

References

Backus, Ad. 2005. Codeswitching and language change: One thing leads to another?. *International Journal of Bilingualism*, 9(3-4). 307-340. https://doi.org/10.1177/13670069050090030101.

Balam, Osmer, Usha Lakshmanan & María del Carmen Parafita Couto. 2021. Gender assignment strategies among simultaneous Spanish/English bilingual children from Miami, Florida. *Studies in Hispanic and Lusophone Linguistics* 14(2). 241–280. https://doi.org/10.1515/shll-2021-2045.

Balam, Osmer & María del Carmen Parafita Couto. 2019. Adjectives in Spanish/English code-switching: Avoidance of grammatical gender in bi/multilingual speech. *Spanish in Context* 16(2). 194–216. https://doi.org/10.1075/sic.00034.bal.

Beatty-Martínez, Anne L. & Paola E. Dussias. 2017. Bilingual experience shapes language processing: Evidence from codeswitching." *Journal of Memory and Language* 95. 173–89, https://doi.org/10.1016/j.jml.2017.04.002.

Bellamy, Kate, & María del Carmen Parafita Couto. 2022. Gender assignment in mixed noun phrases: State of the art. In Dalila Ayoun (ed.), *The acquisition of gender: Crosslinguistic perspectives*, 13-48. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/sibil.63.02bel.

Boersma, Paul & David Weenink. 2024. *Praat: doing phonetics by computer* (version 6.4.11). [Computer program]. http://www.praat.org/

Chirsheva, Galina. 2009. Gender in Russian–English code-switching. *International Journal of Bilingualism* 13(1). 63–90. https://doi.org/10.1177/1367006909103529.

Cisneros, Mark, Eva Rodríguez-González, Kate Bellamy & María del Carmen Parafita Couto. 2023. Gender strategies in the perception and production of mixed nominal constructions by New Mexico Spanish-English bilinguals. *Isogloss. Open Journal of Romance Linguistics* 9(2)/6. 1–30. https://doi.org/10.5565/rev/isogloss.253.

Cruz, Abel. 2023. Linguistic factors modulating gender assignment in Spanish–English bilingual speech. *Bilingualism: Language and Cognition* 26(3). 580–591. https://doi.org/10.1017/S1366728922000839.

Cruz, Abel. 2021. A syntactic approach to gender assignment in Spanish–English bilingual speech. *Glossa: A journal of general linguistics* 6(1). 1–40. https://doi.org/10.16995/glossa.5878.

Fairchild, Sarah & Janet G. Van Hell. 2017. Determiner-noun code-switching in Spanish heritage speakers. *Bilingualism: Language and Cognition* 20(1). 150–161. https://doi.org/10.1017/S1366728915000619.

Gutiérrez Topete, Ernesto. 2023. *Methodological, linguistic, and social effects in language alternation: Evidence from voice onset time in Spanish-English bilinguals*. Berkeley, CA: UC Berkeley dissertation.

Montrul, Silvina, Justin Davidson, Israel de la Fuente & Rebecca Foote. 2013. Early language experience facilitates the processing of gender agreement in Spanish heritage speakers. *Bilingualism: Language and Cognition* 16(1). 1–21. https://doi.org/10.1017/S1366728913000114.

Otheguy, Ricardo, & Naomi Lapidus. 2003. An adaptive approach to noun gender in New York contact Spanish. In Rafael Núñez-Cedeño, Luis López & Richard Cameron (eds.), *A Romance Perspective on Language Knowledge and Use, Selected Papers from the 31st Linguistic Symposium on Romance Languages (LSRL)*. 209–32. Amsterdam/Philadelphia: John Benjamins Publishing Company. https://doi.org/10.1075/cilt.238

Parafita Couto, María del Carmen, Amaia Munarriz, Irantzu Epelde, Margaret Deuchar & Beñat Oyharçabal. 2015. Gender conflict resolution in Spanish-Basque mixed DPs. *Bilingualism: Language and Cognition* 18. 304–323. https://doi.org/10.1017/S136672891400011X.

R Core Team. 2024. R: A language and environment for statistical computing (version 4.3.1). Vienna: R Foundation for Statistical Computing. https://www.r-project.org/

Teschner, Richard V. & William M. Russell. 1984. The gender patterns of Spanish nouns: An inverse dictionary-based analysis. *Hispanic Linguistics* 1. 115–132.
Trawick, Sonya & Trevor Bero. 2022. Gender assignment: Monolingual constraints contribute to a bilingual outcome. *International Journal of Bilingualism* 26(2). 181–197. https://doi.org/10.1177/13670069211053466.
Valdés Kroff, Jorge R. 2016. Mixed NPs in Spanish-English bilingual speech: Using a corpus-based approach to inform models of sentence processing. In Rosa E. Guzzardo Tamargo, Catherine M. Mazak & María del Carmen Parafita Couto (eds.), *Spanish-English code-switching in the Caribbean and the US*, 281–300. Amsterdam: John Benjamins. https://doi.org/10.1075/ihll.11.12val.