B O R E A L I S
An International Journal of Hispanic Linguistics

Gender Assignment and Agreement in L2 Spanish: Evidence from a Medium-Scale Learner Corpus

Maura Cruz Enríquez & François Pichette
Université TÉLUQ

ABSTRACT. This study investigates gender assignment and agreement accuracy in the written productions of French-speaking learners of Spanish across three proficiency levels. Drawing on a medium-scale learner corpus, we coded all noun phrases for gender assignment (based on determiner inflection), for noun-adjective agreement, and for determiner-adjective agreement, and we examined the impact of various linguistic and learner-related predictors using Bayesian mixed-effects models. Although the overall error rate was relatively low, likely due to task type and familiar vocabulary, the models revealed robust effects of proficiency level and of underlying grammatical and lexical factors. Regarding gender assignment, accuracy was significantly lower for nouns with non-prototypical or ambiguous gender markers, for feminine nouns, and when Spanish and French differed in grammatical gender. Moreover, lower accuracy was observed with certain types of determiners. Noun-adjective agreement was influenced by the same factors, except for non-prototypical gender markings, which did not have a significant effect. In addition, less accuracy was observed with prenominal adjectives. Determiner–adjective agreement, in turn, only showed lower accuracy with feminine nouns, but the results of the statistical model should be interpreted with caution, due to high Pareto *k* values. Nevertheless, descriptive data confirm the relevance of distinguishing between noun–adjective and determiner–adjective agreement and highlight the need for larger corpora with a greater number of errors to model this phenomenon more conclusively. Overall, these findings contribute to a better understanding of gender processing in L2, demonstrate the value of medium-sized corpus analysis in second language acquisition research, and lay the groundwork for future research exploring crosslinguistic combinations beyond Spanish and French.

Keywords: gender assignment; gender agreement; L2 Spanish; medium-sized corpus analyses, usage-based approaches to SLA

RESUMEN. Este estudio investiga la precisión en la asignación y la concordancia de género en las producciones escritas de aprendientes francófonos de español en tres niveles de competencia. A partir de un corpus de aprendientes de tamaño medio, se codificaron todas las frases nominales en función de la asignación de género (reflejado en el determinante), así como de la concordancia entre el sustantivo y el adjetivo y entre el determinante y el adjetivo, y se examinó el impacto de diversos predictores lingüísticos y relacionados con los aprendientes mediante modelos bayesianos de efectos mixtos. Aunque la tasa general de errores fue relativamente baja —probablemente debido al tipo de tarea y al uso de vocabulario familiar—, los modelos revelaron efectos robustos del nivel de competencia y de factores gramaticales y léxicos subyacentes. La precisión en la asignación de género fue significativamente menor en los sustantivos con marcas de género no prototípicas o ambiguas, en los sustantivos femeninos y en los casos en los que el género gramatical en español no coincidía con el género en francés. Además, se observó una menor precisión en

 Borealis: An International Journal of Hispanic Linguistics, 2025, 14 / 3. pp. 155-189. https://doi.org/10.7557/1.14.3.8133

ciertos tipos de determinantes. La concordancia sustantivo-adjetivo se vio influenciada por los mismos factores, con excepción de las marcas de género no prototípicas, que no tuvieron un efecto significativo. Además, se observó una menor precisión con los adjetivos prenominales. La concordancia determinante-adjetivo, por su parte, solo mostró menor precisión con sustantivos femeninos, pero los resultados del modelo estadístico deben interpretarse con cautela debido a valores elevados de Pareto *k*. Aun así, los datos descriptivos demuestran la pertinencia de distinguir entre concordancia sustantivo-adjetivo y concordancia determinante-adjetivo, y ponen de relieve la necesidad de contar con corpus más amplios que incluyan un mayor número de errores para modelizar este fenómeno de manera más concluyente. En conjunto, estos resultados contribuyen a una mejor comprensión del procesamiento del género en L2, demuestran la utilidad del análisis de corpus de tamaño medio en la adquisición de lenguas segundas, y sientan las bases para futuras investigaciones sobre combinaciones de lenguas más allá del español y el francés.

Palabras clave: asignación de género; concordancia de género; español L2; análisis de corpus de tamaño medio, enfoques basados en el uso en la ALS

1. Introduction

In recent years, scholars in the field of Second Language Acquisition (SLA) have emphasized the need for medium-sized and large-scale corpus analyses to investigate key linguistic phenomena (Ellis, 2019). Corpus-based approaches offer several advantages: they provide ecologically valid data, allow researchers to track patterns of accuracy and errors across various linguistic contexts, and facilitate the study of multiple predictive variables simultaneously. Beyond its research value, corpus analysis has also been increasingly recognized for its pedagogical potential. As Rojo (2021) argues, bringing second language (L2) corpus findings into the classroom can foster data-driven learning and help both instructors and learners focus on actual usage patterns, including systematic errors and variation in learner language. By grounding instructional practices in empirical evidence, corpus-based approaches can support the development of more targeted and effective teaching materials (see, e.g., Joan Casademont et al., 2022). In this context, analysing learner corpora not only advances theoretical understanding but also contributes to applied goals in language pedagogy.

The increasing advancements in Natural Language Processing (NLP) have made it possible to analyse vast amounts of data efficiently, providing insights into patterns that traditional experimental methods may overlook. In recent years, there has been an increasing call for the use of automated analysis tools on L2 corpora to investigate a range of linguistic phenomena, including grammatical gender (Gudmestad et al., 2019).

Despite these calls, analyses of medium-sized or large L2 corpora remain few. In the case of L2 Spanish, we are not aware of any such study examining gender assignment or gender agreement accuracy. To address this gap, the present study examines the acquisition of grammatical gender in L2 Spanish by Francophone learners through the analysis of a medium-sized corpus of semi-spontaneous productions across different proficiency levels. Our corpus, which consists of more than 350,000 words, enables us to examine well-established predictors of gender assignment and agreement, such as defaulting to the masculine gender, L1-L2 gender congruency, and the role of prototypical vs. non-prototypical gender markings (Muñoz-Liceras et al., 2008; Pérez-Leroux et al., 2023).

This paper presents the results of our corpus analysis, discussing how Francophone learners of Spanish navigate grammatical gender assignment and agreement and how their

errors align with patterns predicted by existing SLA theories. Given the size of our dataset, manual annotation and analysis would prove impractical. Thus, we leverage computational tools from NLP, particularly Python and the SpaCy library (Honnibal et al., 2020), to automate key aspects of data processing, such as tokenization, grammatical gender classification, and error detection. By combining a corpus-based methodology with NLP techniques, we contribute to the ongoing discussion on the role of usage-based approaches in SLA research and highlight the potential of medium-sized and large-scale corpus analysis for studying morphosyntactic acquisition in L2.

2. Gender in Spanish and French

Grammatical gender is a feature of language that lies at the interface of syntax and the lexicon. Although it appears in the inflections of modifying determiners and adjectives associated with a noun in languages such as French and Spanish, this paper adopts the lexical hypothesis (Caramazza & Miozzo, 1997), according to which grammatical gender is a lexical property of the noun and therefore a component of word knowledge. Several researchers have adopted this view (e.g., Ayoun, 2007; Ecke, 2022; Klassen et al., 2023). Consequently, when processing a word in a second language, the activation of its first-language equivalent includes its lexical gender, which makes it prone to crosslinguistic transfer.

In French, grammatical gender is a core feature of the language, with every noun classified as either masculine or feminine. Unlike languages that offer more transparent clues, French often provides limited morphological indicators for gender, making it necessary for speakers to memorize the gender of many nouns individually. French word endings are estimated to provide grammatical gender cues in roughly 80% (Lyster 2006) to 85% (Tucker et al., 1977) of instances. However, very few word endings are strongly associated with a specific gender, while most others are not clearly marked for gender. After excluding nouns with an inherent biological (or semantic) gender, Pichette (2023) identified only 18 two-phoneme endings in the Lexique 3 database (New et al., 2004) that showed a gender bias exceeding 90 percent. For example, suffixes, like *-uche* or *-anse*, reliably signal feminine gender, while others, like *-age* or *-ment*, typically indicate masculine. Most common nouns, especially monosyllabic ones, show no clear pattern. Gender agreement affects articles, adjectives, pronouns, and sometimes even verb forms (in compound tenses with *être*), making mastery of gender crucial for grammatical accuracy.

The Spanish grammatical gender is fundamentally similar to that of French. It also consistently marks gender agreement across articles, adjectives, and pronouns. However, the Spanish system tends to be more transparent and predictable than the French one. The most notable rule is that "nouns ending in '-o' often have masculine gender (99.9%) and those ending in '-a' generally have feminine gender (96.3%)" (Wu & Schiller, 2023). Despite a handful of exceptions among common words, such as *el día* (masculine) and *la mano* (feminine), that rule reflects a regular pattern that learners and language users rely on. As in French, Spanish speakers tend to default to masculine in cases of uncertainty or when referring to mixed-gender groups.

Another relevant factor in gender assignment is the degree of gender congruency between the learner's L1 and the target L2. Authors have been estimating the percentage of nouns in Spanish that contain gender cues to be in the same range as for French, with

80% mentioned by Teschner and Russell (1984) or 70-80% by Harris (1991). However, such estimates should be interpreted with caution. Many nouns have multiple possible translations across the two languages, and both Spanish and French contain polysemous nouns, making it difficult to establish precise rates of gender congruency. We consider that, for the purposes of this study, what matters more than general estimates is what actually occurs in the L2 corpus under analysis.

3. Studies on grammatical gender acquisition

The acquisition of grammatical gender depends on a variety of factors, and three of those are of particular importance for the present study. The first one, related to the language being learned, is the impact of gender cues, which we find in Spanish almost exclusively at the end of nouns. The second one is related to the first language spoken by the learner, in this case, French. The gender that has already been assigned to concepts in French are expected to impact the assignment of gender in Spanish. The third one is the use of a default gender, which is masculine for a vast majority of languages, including French and Spanish. An overview of those factors will be given in this section.

3.1. L2 Word endings

Whether in L1 or L2, gender appears to be learned mostly by mastering correspondence rules between affixes and gender, based on the strength of such associations. In gendered languages, morphological and phonological properties of words offer cues to their grammatical gender, which are most often located at the end of words (Bordag et al., 2006; Kempe & Brooks, 2008). These cues contribute to what Edmonds and Gudmestad (2021) define as the overall "transparency" of a language's gender-marking system, meaning the reliability of gender indicators on both nouns and modifiers. Transparency, therefore, is a property of the language system itself, determined by the consistency and reliability of the mappings between word forms and grammatical categories (Edmonds & Gudmestad, 2021). Word-final cues have been shown to exert a strong influence on gender assignment (Seigneuric et al., 2007), though languages differ significantly in how predictably noun gender can be inferred from form. Research indicates that success in using these linguistic cues depends heavily on the transparency of the system: the more regular and transparent the gender-marking cues, the more efficiently gender is acquired. This pattern is well-supported by evidence showing that higher degrees of regularity and transparency facilitate both first language (L1) gender acquisition (Velnić, 2020) and L2 gender acquisition (Edmonds & Gudmestad, 2021; Gómez Carrero & Ogneva, 2024).

Importantly, as is the case with lexical knowledge in general, gender assignment depends on the learner's previous amount of input and interaction in the L2. As learners' proficiency in the language increases, the more experience they gain at encountering, noticing, and processing word endings, making them increasingly better at mastering those ending-gender correspondence rules. This has been largely supported by research in SLA, among which we find many studies conducted in L2 Spanish with speakers of L1 English (Foote, 2015; Franceschina, 2005; Grüter et al., 2012). Of special interest to us are studies on L2 Spanish gender with speakers of L1 French (Berdasco Muñoz, 2013; Valenzuela et al., 2004; White et al., 2004). Those studies unanimously suggested that learners' success in grammatical gender assignment is tightly linked to their proficiency level, with more

proficient learners showing greater accuracy and sensitivity to gender features encountered in Spanish nouns.

Consequently, regarding the use of word cues in the form of word-ending regularities, we expect learners of Spanish who are below native-like proficiency to show shortcomings in gender assignment. Errors are expected to stem from imperfect knowledge of exceptions to the gender paradigms in the target language, in which case learners would overgeneralize gender rules in Spanish. Errors can also be expected to stem from L1 transfer, in which case the gender of L1 equivalents would be applied to the Spanish nouns, as we explain in the next section.

3.2. L1 impact on L2 gender acquisition and use

Crosslinguistic influence is a well-documented phenomenon in SLA. With respect to gender (as is the case with many other language features), transfer can be seen as working on two different levels. Among the researchers which have made this point is Sabourin, Stowe, and de Haan (2006), who use the terms "deep" and "surface" transfer to refer to these two levels. The first one has to do with the claim that the very existence of a certain feature (such as gender) in one's L1 facilitates the acquisition of this feature in another language, while the second one refers to congruency between gendered systems.

Ellis, Conradie and Huddlestone (2012) found that L1 speakers of Italian had an advantage over speakers of genderless L1s when marking gender in L2 German, and they attributed this advantage as evidence of "deep transfer", since German and Italian are both gendered languages, while their gender systems are not congruent. In a recent study, Gómez Carrero and Ogneva (2024) found that speakers of both L1 English and L1 Russian (which has three genders) could acquire gender in L2 Spanish (a language with two genders), but that for speakers of Russian this process was facilitated. However, findings on this topic remain contradictory. For example, Ragnhildstveit (2017) did not find any such facilitation when comparing gender marking in the writing of learners of Norwegian from five different L1 backgrounds, including genderless languages as well as ones with a similar and a with different number of genders than Norwegian. The studies which have taken up this issue are far too long to list here, but see, for example, Gómez Carrero and Ogneva (2024) for a recent overview of the pertinent theoretical issues and a bibliography on this topic with respect to L2 Spanish.

Surface transfer, in turn, refers to crosslinguistic effects of specific words on L2 gender assignment. The extent of gender assignment congruence between translation equivalents differs across language pairs (Sotiropoulou, 2022). As expected, L2 learners benefit more from positive transfer when the gender of nouns aligns across their languages. White and her colleagues (2004), for example, show that L1 French learners of Spanish exhibit higher accuracy rates in gender assignment when dealing with gender-congruent nouns, but perform less successfully with nouns whose translation equivalents have mismatched gender.

Crosslinguistic influence in gender assignment is particularly well documented for cognates (see, e.g., Vanhove, 2017, for a study on Dutch speakers assigning gender in L2 German). Indeed, positive transfer appears strongest in the case of cognates (Dewaele & Véronique, 2001; Lemhöfer et al., 2008, 2010). Evidence of gender transfer from L1 to L2 also emerges from studies examining gender assignment even in L2 genderless languages like English (e.g., Leśniewska & Pichette, 2014; Morales et al., 2014).

Numerous psycholinguistic studies on lexical processing in adult bilinguals with various L1-L2 combinations further support the existence of crosslinguistic influence on L2 gender. These studies typically reveal a gender congruency effect, where nouns sharing the same gender across the bilinguals' languages lead to more accurate and/or faster processing (see, e.g., Bordag et al., 2006; Lemhöfer et al., 2008; Paolieri et al., 2020; Salamoura & Williams, 2007).[1] Similar gender congruency effects have also been observed in successive bilingual children (Lemmerth & Hopp, 2019).

For reasons of perceived similar distance, speakers of French will therefore rightly assume that the Spanish gender system is similar to that of French, which should cause both deep and surface transfer to occur. For the latter, we expect rules of French to interfere with Spanish ones.

3.3. Default masculine grammatical gender in L1 and L2 acquisition

Masculine appears to be the default grammatical gender for a vast majority of languages (Bellamy & Parafita Couto, 2022), including both Spanish and French. This phenomenon has been demonstrated in L1 development for French (e.g., Karmiloff-Smith, 1979) and for Spanish (e. g., Pérez-Pereira, 1991). It has also been observed when those languages are learned by speakers of other languages, both in French (Bartning, 2000; Holmes & de la Bâtie, 1999), and Spanish (Grüter et al., 2012; McCarthy, 2008; Montrul et al., 2008; White et al., 2004).

For L1 Spanish, Dominguez, Cuetos & Segui (1999) obtained shorter reaction times by L1 speakers on lexical decision tasks for masculine than for feminine nouns, and Eddington and Hualde (2008) also showed that native speakers of Spanish tend to incorrectly overextend masculine agreement compared to feminine agreement. Gómez Carrero and Ogneva (2024) noted the use of masculine by default in L2 Spanish by both native speakers of English (which is non-gendered), as evidenced by higher gender accuracy for masculine determiner phrases than for feminine ones.

For French, Tucker, Lambert, and Rigault (1977) found that when French-speaking children were uncertain about a noun's gender, they overwhelmingly assigned masculine, suggesting it acts as the default gender. In L2 French, Matthews (2016) demonstrated that English-speaking learners also tended to default to masculine gender when unsure, reflecting a similar bias than that observed in L1 speakers.

3.4. Gender acquisition studies and corpus analysis

Gender Assignment refers to the process by which a grammatical gender is attributed to a noun, either based on semantic properties (like biological sex) or formal properties (such as word endings). According to Corbett (1991), gender assignment involves the rules or principles a language uses to associate nouns with a particular gender category. For example, in Spanish, *coche* ("car") is assigned the masculine gender (*el coche*) despite not carrying any biological meaning. In contrast, Gender Agreement is the process by which related elements in a sentence (such as determiners, adjectives, or past participles) match the gender of the noun they modify. In Spanish, adjectives and articles must reflect the noun's gender.

[1] For recent reviews, see Johanessen et al., (2024) and Sá-Leite Fraga & Comesaña (2019). See also Costa, Kovacic, Franck, & Caramazza (2003) for contrary evidence.

Studies on L2 French have found that indefinite determiners (*un/une*) tend to be more difficult to acquire than definite ones (*le/la*). For instance, Bartning (2000) compared gender assignment accuracy across definite, indefinite, and possessive (*mon/ma*, *ton/ta*, *son/sa*) determiners, and found that indefinite determiners posed the greatest difficulty, although the difference between definite and indefinite forms was less pronounced among advanced learners.

In Spanish, both definite (*el libro* 'the book', *la casa* 'the house') and indefinite (*un libro* 'a book', *una casa* 'a house') determiners provide gender information about the noun, as they do in French. Plural first-, second-, and third-person possessive determiners also provide gender information (*nuestro libro* 'our book', *nuestra casa* 'our house'), whereas singular first-, second-, and third-person possessive determiners do not (*mi libro*$_M$ 'my book', *mi casa*$_F$ 'my house'). While some studies on L2 Spanish have explored the role of determiners and morphophonological cues in gender assignment (see, e.g., Afonso et al., 2014), none to our knowledge have directly compared learners' accuracy with definite versus indefinite determiners. Further research is therefore needed to address this comparison explicitly.

The analysis of written corpora presents challenges. The first challenge is that it may be problematic to tease apart gender assignment and gender agreement, and to distinguish gender assignment errors from gender agreement errors. In other words, it may be impossible to determine if an error stems from the wrong gender being assigned to a noun, or if incorrect inflections on the other sentence constituents result from insufficient command of grammar. For example, the noun phrase **El*$_M$ *coche*$_M$ *blanca*$_F$ ("the white car") would show correct assignment, but incorrect agreement[2], since the word *coche* is correctly assigned a masculine determiner (*el*) but incorrectly assigned a feminine adjective (*blanca*). On the contrary, the sentence **La coche es rápida* manifests incorrect gender assignment, possibly under the influence of the French word *automobile* or *voiture* (both of which are feminine).

However in a phrase such as *Equipos$_M$ olímpicas$_F$ ("Olympic teams"), it is not clear whether (1) gender was correctly assigned to the noun *equipos* with incorrect gender agreement for the adjective *olímpicas*, or if (2) despite its correct form, the noun was wrongly assigned in the person's mind the feminine gender (which is the gender in L1 French), reflected in the adjective *olímpicas*, which would be inflected accordingly. Among researchers in Spanish L2 acquisition, Franceschina (2005) and White and her colleagues (2004) suggest that in such cases, gender features on nouns are usually correctly represented, and that errors occur at the level of agreement morphology (i.e., adjectives and determiners). Errors in L2 Spanish do not always lend themselves to such easy classification, and those issues -as well as how we intend to address them- will be discussed in the method section.

The second issue to consider in our analyses is that the accuracy of gender marking might be influenced by the amount of processing effort required. Even when learners know the correct gender of a noun, they may still fail to mark it accurately on other sentence elements which should agree with them, particularly under conditions of strain such as time pressure or when agreement must be maintained across greater distances within the

[2] As explained by Franceschina (2005, p. 97), gender assignment in Spanish extends beyond word ending rules, and can also be based to a lesser extent on assignment rules of a semantic, grammatical, etymological, and phonological nature.

sentence. Research shows that error rates increase as the linear distance between the noun and its modifier grows (Bruhn de Garavito & White, 2002; Edmonds et al., 2020; Edmonds & Gudmestad, 2018; Keating, 2009; Lichtman, 2009). A higher error rate is thus found with increased linear distance between the noun and its modifier, making it easier to correctly inflect the adjective when it immediately follows or precedes the noun. For example, the adjective *listos* ('ready') is more likely to be inflected correctly in attributive position (1) than in predicative position (2), and the latter is more likely to be inflected correctly than in example (3), where the linear distance between adjective and noun is greater.

(1) Los equipos$_{M}$ *listos*$_{M}$.
'The ready teams.'

(2) Los equipos$_{M}$ estaban *listos*$_{M}$.
'The teams were ready.'

(3) Los equipos$_{M}$ de todo el mundo estaban *listos*$_{M}$.
'Teams from all over the world were ready.'

Managing these longer dependencies places a greater cognitive load on the learner, often making successful gender agreement too demanding to achieve. Consequently, when examining gender agreement in our analysis, we will take linear distance into consideration.

Bartning (2000) examined the acquisition of adjectival agreement in L2 French by advanced and pre-advanced learners, within the framework of Pienemann's Processability Theory (Pienemann & Keßler, 2013; Pienemann & Lenzing, 2014). The study focused on three syntactic positions for French adjectives: attributive anteposition (AP; *mon petit chien* 'my little dog'), attributive postposition (PP; *le journal suédois* 'the Swedish newspaper'), and the predicative position (Pred; *ils sont trop froids* 'they are too cold'). In attributive positions, the adjective and noun are adjacent (as in Spanish: *la casa amarilla*), while in predicative constructions, the adjective appears after a copula, a structure that is similar in both French and Spanish (e.g., *la casa es amarilla*). Her results showed that prenominal attributive adjectives posed the greatest difficulty for L2 learners, suggesting that syntactic position plays a key role in learners' ability to process and produce gender agreement accurately.

While French and Spanish share similar rules for adjective-noun agreement, French allows —and often requires— more frequent use of adjectives in prenominal position (*la petite maison*, 'the little house') than Spanish (Lichtman, 2009). This may help explain why relatively little research has focused on prenominal adjective agreement in L2 Spanish, despite the challenges it might entail.

4. Method

4.1. Research questions and predictions

The present study aims to investigate the production of grammatical gender in L2 Spanish by L1 French learners, drawing on a medium-sized corpus of written texts produced across a range of proficiency levels. Based on the considerations presented in the previous section, we formulated the following research questions:

(1) How do specific linguistic factors such as L2 noun gender, L2 gender transparency, L1-L2 gender congruency, and determiner type affect learners' accuracy in gender assignment, and how do these effects vary across proficiency levels?
(2) How does adjective-noun gender agreement accuracy vary as a function of adjective position (prenominal vs. postnominal), linear distance between the noun and the adjective, L2 noun gender, L2 gender transparency, and L1-L2 gender congruency, and to what extent are these effects modulated by the learners' proficiency level?
(3) How does determiner-adjective gender agreement accuracy vary as a function of adjective position (prenominal vs. postnominal), linear distance between the noun and the adjective, the noun's gender, gender marking transparency, and L1-L2 gender congruency, and to what extent are these effects modulated by learners' proficiency level?

Based on the Lexical hypothesis and on the second language acquisition research listed above, we expect the prototypical categories of the variables (e.g., masculine nouns, prototypical ending, L1-L2 gender congruence, short linear distance, postposition of adjective) to facilitate gender assignment and agreement. On the contrary, we expect their non-prototypical variants to be associated with more errors in gender assignment and agreement.

4.2. Corpus assembly

The texts included in the corpus were written by 219 students enrolled in L2 Spanish language courses at a university in Quebec between 2020 and 2025[3]. Student metadata is available thanks to a mandatory questionnaire completed by all language learners, which includes information such as age, gender, and the languages they speak. Sixty-five percent (n = 142) of the students were women, and 35 % were men (n = 77). The average age was 36.5 years, with ages ranging from 22 to 65 years. All students were native speakers of French and had no significant exposure to other gendered languages. They were distributed across three proficiency levels: Beginner (n = 77), Intermediate (n = 62), and Advanced (n = 80). The Beginner level consists of students enrolled in Spanish courses equivalent to levels A1 and A2, according to the Common European Framework of Reference for Languages (CEFR). Intermediate level students were enrolled in B1 and B2 level courses. At the Advanced level, students were enrolled in a C1 level Spanish writing course.

To achieve the highest possible word count, all available student submissions were included. In this way we managed to create a corpus that contained more than 350,000

[3] The learner corpus used in this study is not yet publicly available because it contains non-anonymized linguistic productions from participants. The anonymization process is currently underway. Once completed, the anonymized corpus will be made publicly available in the same GitHub repository that hosts the Python and R code used in the project. Researchers interested in accessing the corpus for replication or related analyses may also contact the authors directly to request access under appropriate ethical and confidentiality conditions.

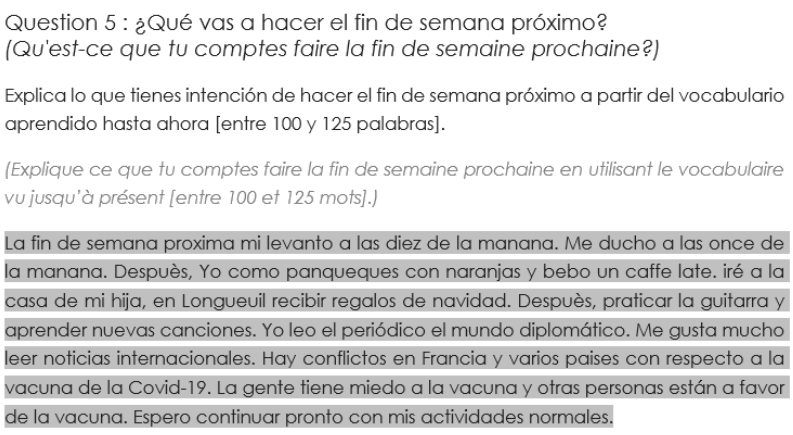
words. Tasks were designed to elicit free or guided language output appropriate to each participant's proficiency level.

It should be noted that, since these were production tasks focused on the course content and were also evaluative, participants were expected to pay close attention to what they wrote and try to make as few mistakes as possible. Nevertheless, the main objective of this study is not to quantify overall learner accuracy, but rather to examine how different linguistic and cognitive variables contribute to gender assignment and agreement errors at each proficiency level. The variability and richness of the corpus provide a unique opportunity to analyse the interaction of these variables in a large and ecologically valid dataset. Moreover, in academic writing situations, one can expect a phenomenon of avoidance on the part of students for items and rules they deem challenging, which also leads to fewer errors than if it were spontaneous production, a phenomenon that has been attested for a long time (see Schachter, 1974).

At the Beginner level, the texts come from a personal diary task in which learners were instructed to write about themselves and their favourite public figures. These productions include both descriptive and narrative passages. Participants were required to complete a Diary in which they talked about their personal lives, their family members, their favourite character, their favourite book, among other topics. In these tasks, they had to put into practice the vocabulary, grammar, and structures learned in the Spanish course. The texts were descriptive and narrative. Participants were given unlimited time to complete the task, but the online writing platform provided no spelling or grammar correction features.

In the Diary, instructions for some questions appeared in Spanish and French. We decided to include in the analysis all the NPs that appeared in the students' response, even those which had appeared in the instructions, since in most cases they did not offer clues about the gender of the nouns, such as *Describe tu casa* ("Describe your house"). Furthermore, we observed that, although a clue to the gender of the noun was sometimes provided in the instructions, this did not prevent assignment errors from occurring. Figure 1 shows an example of the question instructions, and the output collected. Note that the participant makes an assignment error even though the same noun phrase (*el fin de semana*) appears in the Spanish instructions.

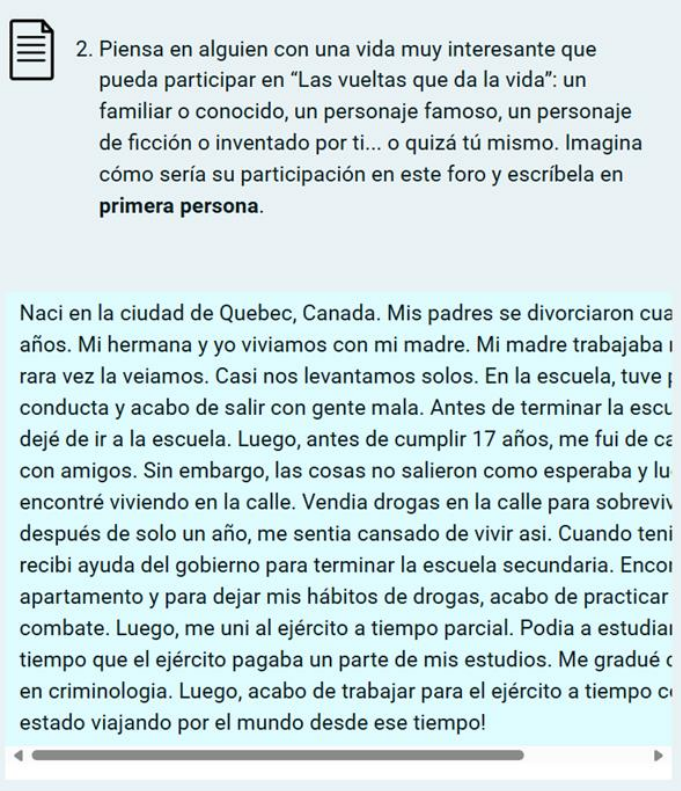
Figure 1. Example of the question instructions and the output collected for the Beginners task



At the Intermediate level, the corpus includes a narrative (see Figure 2), and two opinion essays. The three texts were written as part of a time-limited online exam. The

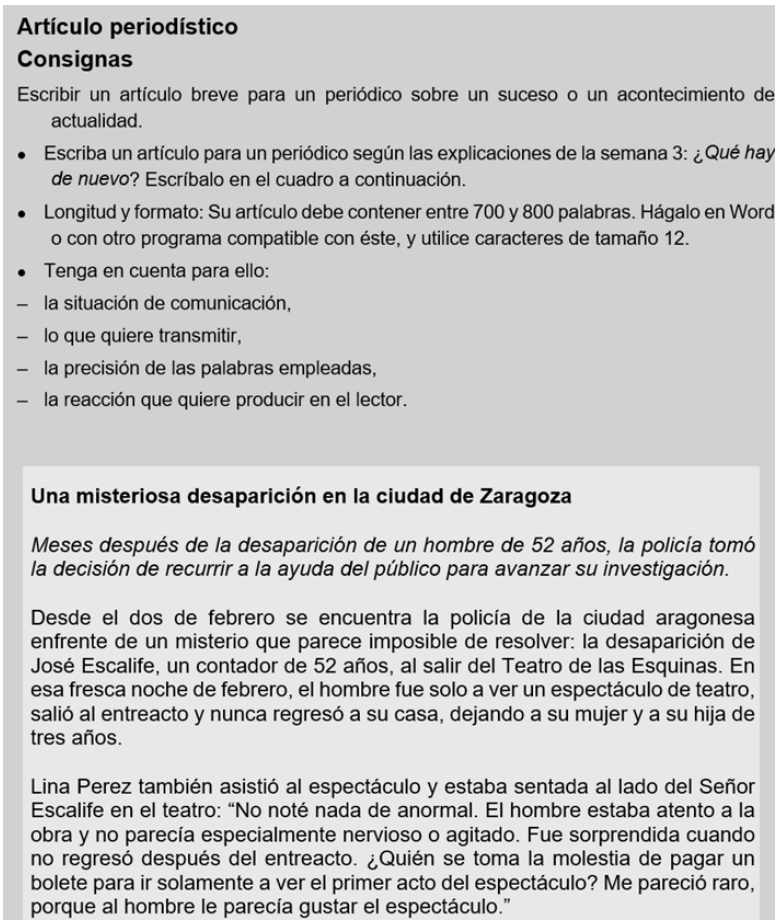
tasks required learners to produce texts based on prompts provided during the exam. As with the Beginner levels, the writing environment did not include any automatic correction tools.

Figure 2. Example of the question instructions and the output collected for an Intermediate task



At the Advanced level, the range of writing tasks was more diverse and cognitively demanding. Learners produced opinion pieces, narrative texts set in the past, and newspaper articles (see Figure 3) inspired by or responding to previously assigned readings, among others. These tasks were designed to elicit a variety of discourse types and more complex syntactic structures. Unlike the Intermediate level, participants were given unlimited time to complete their assignments. However, the online writing platform provided no access to spelling or grammar correction features.

Figure 3. Example of the question instructions and the output collected for an Advanced task



Artículo periodístico

Consignas

Escribir un artículo breve para un periódico sobre un suceso o un acontecimiento de actualidad.

- Escriba un artículo para un periódico según las explicaciones de la semana 3: ¿*Qué hay de nuevo*? Escríbalo en el cuadro a continuación.
- Longitud y formato: Su artículo debe contener entre 700 y 800 palabras. Hágalo en Word o con otro programa compatible con éste, y utilice caracteres de tamaño 12.
- Tenga en cuenta para ello:

– la situación de comunicación,
– lo que quiere transmitir,
– la precisión de las palabras empleadas,
– la reacción que quiere producir en el lector.

Una misteriosa desaparición en la ciudad de Zaragoza

Meses después de la desaparición de un hombre de 52 años, la policía tomó la decisión de recurrir a la ayuda del público para avanzar su investigación.

Desde el dos de febrero se encuentra la policía de la ciudad aragonesa enfrente de un misterio que parece imposible de resolver: la desaparición de José Escalife, un contador de 52 años, al salir del Teatro de las Esquinas. En esa fresca noche de febrero, el hombre fue solo a ver un espectáculo de teatro, salió al entreacto y nunca regresó a su casa, dejando a su mujer y a su hija de tres años.

Lina Perez también asistió al espectáculo y estaba sentada al lado del Señor Escalife en el teatro: "No noté nada de anormal. El hombre estaba atento a la obra y no parecía especialmente nervioso o agitado. Fue sorprendida cuando no regresó después del entreacto. ¿Quién se toma la molestia de pagar un bolete para ir solamente a ver el primer acto del espectáculo? Me pareció raro, porque al hombre le parecía gustar el espectáculo."

4.3. Corpus description

A total of 45,295 noun phrases (NPs) were extracted from the corpus. Note that this figure represents the total number of noun occurrences rather than distinct lexical items; many nouns recur, with some appearing more frequently than others. Of these, 36,231 NPs (approx. 80% of all NPs) were retained for analysis, of which 20,094 contained feminine nouns and 16,137 were masculine. In terms of proficiency distribution, 19,463 NPs were produced by Advanced learners, 10,674 by Beginners, and 6,094 by Intermediate learners. These NPs met the following requirements:

- The nouns exhibit a clear masculine or feminine gender. For example, we excluded 37 NPs that contained nouns with ambiguous grammatical gender in Spanish (e.g., *el/la mar*, *el/la azúcar*, *el/la interrogante*, *el/la internet*, *el/la pijama*).
- The NPs allowed for the examination of gender assignment and/or gender agreement. For example, we retained *la casa*, but not *mi casa*, since the latter does not allow the gender assignment to be evaluated because the determiner *mi* does not have grammatical gender. Likewise, we retained *casa pequeña* and *la cama es cómoda*, but not *casa grande* or *la cama es grande*, since the last two do not allow for the assessment of gender agreement because the adjective *grande* has no overt grammatical gender.

- The nouns displayed grammatical gender with no semantic motivation. We excluded nouns with semantic gender (e.g., *la mujer*, *el hombre*). This decision was based on previous research indicating that semantic gender provides a strong cue for accurate gender agreement in Spanish (Alarcón, 2009; Berdasco Muñoz, 2013).
- Regarding crosslinguistic gender comparison, a total of 81.79% of the nouns in the dataset shared the same grammatical gender in both Spanish and French. About 84% (n = 16,900) of feminine nouns are also feminine in French, while 79% (n = 12,783) of masculine nouns are also masculine in French. Table 1 shows the distribution of grammatical gender in Spanish and in their French equivalents.

Table 1. Distribution of grammatical gender in Spanish and in their French equivalents

Spanish Gender	French Gender	Example	Count	Percent
Fem	Fem	*casa-maison*	**16,900**	84.0
Fem	Masc	*cama-lit*	2,958	14.7
Fem	Both	*tarde-après-midi*	263	1.31
Masc	Masc	*nombre-nom*	**12,783**	79.0
Masc	Fem	*armario-armoire*	2,373	14.7
Masc	Both	*año-an/année*	1,014	6.27
Total			36,291	100

In terms of morphological gender marking, Table 2 presents the distribution of gender markers (Protot_mark) by noun gender. As shown in the table, approximately 58% of both feminine and masculine nouns display prototypical gender marking (e.g., *-a* for feminine and *-o* for masculine). It is also noteworthy that a very small proportion of feminine nouns of only 0.33% bear a non-prototypical gender marker.

Table 2. Distribution of gender markers in Spanish nouns

Noun_gen	Protot_mark	Example	Count	Percent
Fem	Protot	*casa*	11,656	57.9
Fem	Amb	*calle*	8,399	41.7
Fem	No_protot	*mano*	66	0.33
Masc	Protot	*año*	9,450	58.4
Masc	Amb	*lugar*	4,924	30.4
Masc	No_protot	*tema*	1,796	11.1

So far, the dataset reveals that most of the nouns share the same grammatical gender in both Spanish and French, and that the majority also display prototypical gender markings. Only a minority of less than 30% show a gender mismatch between the two languages. Furthermore, just 5% of the nouns (n = 1,862) carry a non-prototypical gender marker, of which only 0.33% are feminine nouns[4]. These findings suggest that the potential

[4] Please note that the raw figures from our corpus are slightly higher than the actual percentages found in the Spanish lexicon. Based on an analysis of 7,700 different Spanish nouns, (Beck, 2016) found that only 3%

for gender assignment errors should not be particularly high, especially considering the well-documented tendency to default to masculine gender in ambiguous contexts. This is further reinforced by the fact that, particularly at the Beginner and Intermediate levels, the texts rely primarily on vocabulary taught in the course, meaning that participants are likely to be familiar with most of the words. Additionally, these words tend to be high-frequency items, especially at the Beginner level. Nonetheless, as noted earlier, our goal is to determine whether a medium-scale corpus analysis can identify errors that align with the same underlying factors observed in previous studies.

4.4. Data processing and analysis

The data were processed using Python (Python Software Foundation, 2025) and several NLP libraries, including SpaCy for part-of-speech tagging, morphological analysis, and syntactic function. Custom scripts were developed to automate the identification and classification of noun phrases, determine gender agreement patterns, and extract relevant lexical and grammatical features. The analysis code used in this study is fully available in a public GitHub repository (https://github.com/MauraTeluq20/L2-Spanish-gender-assignment).

While most processing steps were automated, a double manual review was conducted to correct tagging errors. This manual revision was carried out by both authors, who are fluent in Spanish and French, to ensure consistency and accuracy. The rate of tagging errors was around 4 %. The most frequent errors involved incorrect identification of noun phrases or misclassification of word categories—for example, adjectives erroneously tagged as nouns when appearing after determiners. Further details on the challenges encountered during the use of SpaCy for automatic annotation, as well as examples of error types, can be found in Cruz Enríquez et al. (2025) and in a forthcoming publication currently in preparation by Cruz Enríquez and colleagues (2025).

1. Gender assignment vs. agreement

Gender assignment was operationalized as the gender marking on the determiner (e.g., *la casa* vs. *el casa*), whereas gender agreement was assessed by whether the adjective matched the grammatical gender of the noun as marked by the determiner (e.g., *la casa roja* vs. **la casa rojo*).

As we said earlier, errors are not always easy to classify. A particularly interesting case is a phrase like **el casa roja*. This construction suggests that the learner assigned the masculine gender (*el*) to the noun *casa (*which is incorrect, since *casa* is feminine) but still produced a feminine adjective (*roja*). This may indicate that the learner recognized the inherent gender of *casa*, possibly through lexical knowledge or phonological cues, but failed to apply it consistently across all components of the noun phrase. Such mismatches provide evidence that learners may process gender assignment and agreement through partially independent mechanisms.

Similarly, phrases with no determiner, such as *casa roja* and **casa rojo,* fall into a distinct category. In *casa roja*, the adjective matches the noun's inherent gender, which might reflect implicit gender knowledge even in the absence of explicit assignment via a determiner. In contrast, **casa rojo* reveals a lack of agreement and may indicate that,

deviate from the standard gender rules. The discrepancy in our data is due to the fact that exception words tend to occur more frequently in our corpus than regular ones.

without a determiner to guide gender assignment, the learner struggled to retrieve or apply the correct gender form.

In our analysis, we first examine the gender of the adjective in relation to that of the noun, across all noun phrases that contain both elements, to identify potential mismatches and investigate how these are influenced by the predictive variables. This initial comparison focuses on whether the adjective reflects the inherent gender of the noun (*Noun_adj_agreement*), regardless of whether a determiner is present.

However, to assess gender agreement more precisely, we restrict our second analysis to noun phrases that include both a determiner and an adjective. This allows us to isolate cases in which gender assignment is overtly encoded through the determiner, ensuring a clearer interpretation of agreement patterns (*Det_adj_agreement*).

This two-pronged approach allows us to address different aspects of gender processing. The broader *Noun_adj_agreement* variable captures implicit knowledge or lexical retrieval (especially relevant when determiners are absent), while the more restrictive *Det_adj_agreement* variable provides a cleaner test of morphosyntactic agreement once gender has been explicitly assigned.

2. Predictive variables

For analyzing gender assignment, the dependent variable was labeled as *Noun_det_assignment*, and coded as correct (1) or error (0) at the noun-determiner level (e.g., ***una*** *mesa* vs. **un mesa*). The following independent variables were included:

- Target gender in Spanish: In our dataset it was labeled as **Noun_gen** (Masc or Fem). It indicates the grammatical gender (Masc or Fem) of each Spanish noun, to assess possible default masculine usage.
- Gender congruency between L1 and L2, to test for transfer effects: In our dataset it was labeled as **Compare_L2_L1**. It compares the grammatical gender of each noun in Spanish L2 with that of its French L1 equivalent. It has three possible values: *same* (e.g., *la casa* vs. *la maison*), *different* (e.g., *el vestido* vs. *la robe*), and *both* (e.g., el año vs. l'an$_{M}$, l'année$_{F}$), allowing us to evaluate possible positive or negative transfer from the L1.
- Gender marking type, based on morphological cues: In our dataset it was labeled as **Protot_mark**. It categorizes each Spanish noun based on its morphological ending, with three possible values: *prototypical* (e.g., *la cama*, *el caso*) *non-prototypical* (e.g., *la mano*, *el idioma*), and *ambiguous* (e.g., *camión*).
- Determiner type, based on the grammatical category of the determiner, was labeled as **Det_type** in our dataset. This variable classifies determiners into five categories: definite (e.g., *el coche*, *la casa*), indefinite (e.g., *un coche*, *una casa*), possessive (e.g., *nuestro coche*, *vuestra casa*), demonstrative (e.g., *este coche*, *esa casa*), and "Otro", which includes the form *otro* (along with its inflected variants *otra, otros, otras*). The latter was assigned its own category because it poses classification challenges due to its hybrid grammatical behavior. Although it shares morphosyntactic properties with adjectives (e.g., gender and number agreement), it also displays features characteristic of quantifiers and indefinite determiners. For instance, *otro* can function as a determiner introducing preverbal subjects (e.g., *Otra lengua romance es el italiano*), a property typical of determiners. It also admits pronominal uses (e.g., *Quiero otro*) and partitive constructions (e.g., *Otro de los*

problemas que tenemos), which align with quantifier-like behavior (RAE, 2024a). Given this multifunctional profile, *otro* was grouped in a separate category in our analysis to reflect its distinctive syntactic and semantic features.

For analyzing gender agreement, the dependent variables were labeled as Noun_adj_agreement and Det_adj_agreement, also coded as correct (1) or error (0) at the noun-adjective level (e.g., *mesa pequeña**a*** vs. ***mesa pequeño**o*) and at the determiner-adjective level (e.g., ***el/la** mesa pequeñ**o/a*** vs. ***el/la** mesa pequeñ**a/o***), respectively. The dependent variables shown above were included (except for Determiner type), to which the following were added:

- Distance between noun and adjective: It is a numerical variable indicating the number of intervening words between the adjective and the noun. It ranges from 0, when the adjective is adjacent to the noun, to 10, which is the maximum distance observed in the corpus. In our dataset it was labeled as **Adj_Noun_Distance**.
- Noun-adjective order: In our dataset it was labeled as **Pre_noun**. It is a binary variable that indicates whether the adjective appears before or after the noun. It has two possible values: *Yes*, when the adjective is in prenominal position, and *No*, when it is in a postnominal position.

3. Determining the French equivalent of Spanish nouns

To analyze the potential effects of shared or divergent grammatical gender across languages, it was necessary to identify the appropriate French equivalent for each Spanish noun in the corpus. In many cases, such as *casa* ("house"), which almost invariably translates to *maison* in French, the equivalence was straightforward. However, for other nouns —particularly polysemous and homonymous ones— a manual review of the corpus was required to consider the context of use and accurately determine the French equivalent before determining its grammatical gender.

A representative example is the word *bolsa*, which can correspond to either *le sac* (when referring to a physical bag), *la sacoche* in the case of a purse, or *la bourse* (when referring to the stock market). Similarly, homonymous nouns such as *corte* —which may mean either *la corte* (*la cour* in French, *the court* in English) or *el corte* (*la coupe* in French, *the cut* in English)—required contextual analysis to establish the intended meaning and correct translation.

Additional challenges arose from semantic and formal variation between the two languages. In some instances, automatic translations had to be corrected, especially when a Spanish noun had multiple possible synonyms in French. In such cases, different criteria guided the selection process:

- Analogical criterion: When multiple French synonyms were interchangeable in the corpus contexts, we selected the one most similar in form to the Spanish noun. For example, for *suscripción* (feminine), we chose *souscription* (feminine) rather than *abonnement* (masculine), allowing us to classify this as a case of gender agreement across languages.
- Semantic criterion: In other cases, the decision was based on meaning. For instance, the word *idioma* was matched with *langue*, despite *idiome* being closer in form, as the latter has a different and more restricted meaning in French, and it has such low frequency that it is often not known by native speakers. To validate this choice, we consulted ten participants (during an informal virtual meeting unrelated to the

writing tasks and held afterward) and asked which French word they associated with *idioma*. All of them answered *langue*.

4. Data analysis

In conducting our data analysis, we adhered to established practices in the field of SLA, drawing on methodological recommendations designed to enhance the robustness and clarity of quantitative findings (Gries, 2021; Gries & Ellis, 2015; Norris, 2015; Norris et al., 2015). This included applying best practices for the treatment of predictor variables (Pasta, 2009).

For the analysis, separate multilevel logistic regression models were fitted using Bayesian[5] estimation with a Bernoulli family and logit link. The first model focused on gender assignment, predicting whether the gender assigned to a noun was target-like. The second model targeted gender agreement, predicting whether the adjective agreed in gender with the noun. The third model also targeted gender agreement, but it served to predict whether the adjective agreed in gender with the determiner.

In all three models, fixed effects included learner-related and linguistic predictors: proficiency level (Level), the comparison between the gender of the noun in the L1 and L2 (Compare_L2_L1), and the type of gender marking on the noun (Protot_mark, with levels such as prototypical, ambiguous, or non-prototypical). The assignment model also included the type of determiner as a variable. The agreement models additionally included variables related to morphosyntactic processing: the linear distance between the adjective and the noun, and adjective position (prenominal or postnominal).

All three models were implemented using the brms package (Bürkner, 2017) in R (R Core Team, 2022)[6], and all continuous predictors were standardized prior to analysis. The reference levels for categorical predictors were set to reflect the most canonical, unmarked conditions or those which are considered to facilitate gender assignment and agreement in L2 (e.g., masculine gender, postnominal adjective, prototypical gender marking, same gender in the L2 and the L1, and definite determiners). The variable representing proficiency level was treated as an ordinal predictor, in line with expert recommendations (e.g., Pasta, 2009), to appropriately capture the assumed progression in learners' performance across levels.

The priors were the default weakly informative priors. The fixed effects were assigned flat (noninformative) priors, while Student-t (3, 0, 2.5) priors were used for the intercept and group-level standard deviations. Model convergence was verified using Rhat statistics and effective sample size metrics. All model parameters converged successfully (Rhat = 1.00). A Markov Chain Monte Carlo (MCMC) sampling was run with 2,000 iterations, 1,000 warmups and four chains. To assess the model's predictive performance, a Leave-One-Out cross-validation (LOO) analysis was conducted, implemented via the loo package (Vehtari et al., 2017).

[5] For further information on Bayesian statistics with R, McElreath (2020) provides an accessible and comprehensive introduction.

[6] While Python also allows for this type of analysis—for example, with the *bambi* package—we chose R because of its simpler syntax and more user-friendly integration of statistical modeling tools. R offers also better documentation and more extensive community support for linguistic data analysis, which makes it more accessible for researchers in linguistics who may not have a strong background in computer science or software development.

5. Results

5.1. Gender assignment (Research Question 1)

We fitted a Bayesian mixed-effects logistic regression model to examine gender assignment. The model included random intercepts for participants (*Subject_ID*) and for lexical items (*Noun_lemma*), to account for subject- and item-level variability. It showed satisfactory predictive accuracy (LOOIC: 3356), with all Pareto *k* diagnostics indicating good reliability ($k < 0.7$). The results, based on 32,151 observations, showed high overall accuracy, with an intercept estimate of 7.01 (95% CI [6.50, 7.57]) on the logit scale, corresponding to an estimated probability of approximately 99%. This intercept represents the predicted accuracy for masculine nouns with a definite determiner and congruent gender between L1 and L2 at the reference proficiency level (Beginner). Table 3 presents the results for fixed effects for the gender assignment model. For each predictor, we report the estimated coefficient, its standard error (SE), and the 95% highest posterior density interval (HPD CI).

Table 3. Fixed effects for the gender assignment model (Bayesian logistic regression)

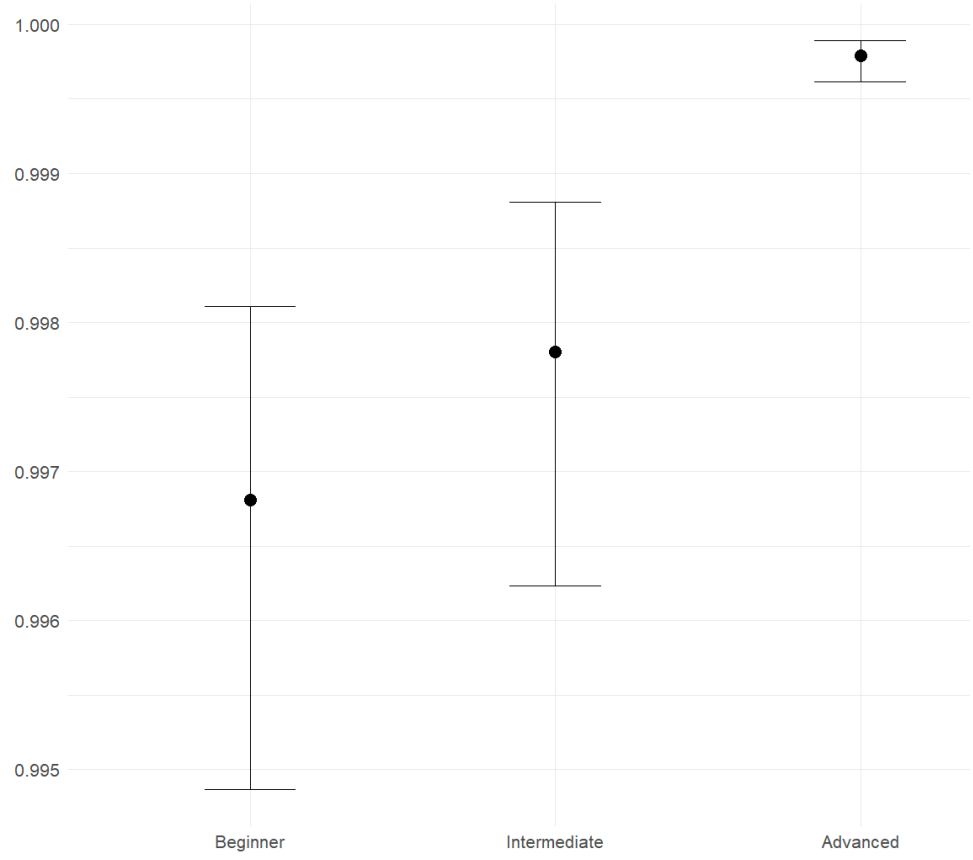
Predictor	**Est.**	**SE**	**95% HPD CI**
Intercept	7.01	0.27	[6.50, 7.57]
Level.L	1.96	0.18	[1.62, 2.32]
Level.Q	0.75	0.19	[0.37, 1.13]
Protot_markAmb	-0.79	0.17	[-1.11, -0.47]
Protot_markNo_protot	-2.49	0.39	[-3.25, -1.71]
Compare_L2_L1both	0.10	0.64	[–1.17, 1.39]
Compare_L2_L1different	-1.45	0.18	[-1.81, -1.10]
Noun_genFem	-0.89	0.17	[-1.23, -0.56]
Det_typeCuant	-0.85	0.22	[-1.27, -0.42]
Det_typeDem	-1.21	0.23	[-1.65, -0.76]
Det_typeIndef	-0.14	0.16	[-0.45, 0.17]
Det_typeOtro	-0.84	0.39	[-1.58, -0.06]
Det_typePoss	-0.46	0.42	[-1.23, 0.38]

Random intercepts for both subjects and items were included in the model to account for individual variation in performance and lexical variability across nouns. The standard deviation of the random intercepts was substantial for both participants (sd = 0.94) and noun (sd = 1.30), indicating that variation across individuals and lexical items is not negligible. Item-level variation may reflect differences in lexical frequency, or morphophonological complexity, while participant-level variation may reflect unmeasured differences in cognitive resources. Importantly, despite this variability, the fixed effects of proficiency and predictive variables remained robust, which strengthens the interpretation that these factors systematically influence gender assignment accuracy.

The **proficiency** variable showed a nonlinear effect on gender agreement accuracy, as revealed by both a significant linear and quadratic component (Level.L = 1.96, 95% CI [1.62, 2.32]; Level.Q = 0.75, 95% CI [0.37, 1.13]). Estimated probabilities indicated an increase in accuracy from Beginner (96.6%) to Intermediate (97.8%), and a further rise to

Advanced (99.8%). All pairwise comparisons between levels were significant, with odds ratios ranging from 0.06 to 0.63, indicating a large developmental effect. The contrast between Beginner and Advanced was particularly pronounced (odds ratio = 0.06, 95% HPD [0.0356, 0.0985]), while the difference between Intermediate and Advanced also remained substantial (odds ratio = 0.10, 95% HPD [0.0540, 0.1646]). The contrast between Beginner and Intermediate is less substantial (odds ratio = 0.63, 95% HPD [0.3408, 0.9613]). These results support a less marked improvement in the Intermediate level, followed by larger gains at the highest proficiency level, consistent with a non-linear progression. Figure 4 displays the effect of proficiency level on gender assignment. Each line represents a different proficiency level (Beginner, Intermediate, Advanced). Vertical bars show 95% credible intervals around the estimates, while each point indicates the estimated probability of correct assignment. The shorter intervals for the Advanced group suggest greater consistency or precision in this group's responses.

Figure 4. Effect of proficiency on gender assignment



The graph in Figure 5 illustrates the effect of the other predictors, which are described in detail below the figure.

Figure 5. Effect of predictors on gender assignment by proficiency level

Morphological prototypicality (see Figure 5 A): The presence of a non-prototypical gender marker (e.g., *idioma*$_{M}$) had a strong negative effect on accuracy (Est. -2.47, 95% CI [-3.25, -1.71]), and even ambiguous markers showed a significant decrease in accuracy (Est. -0.79, 95% CI [-1.11, -0.47]) compared to prototypical ones. All pairwise contrasts were statistically significant, especially between prototypical and non-prototypical nouns, with odds ratios up to 12:1 between prototypical and non-prototypical items (odds ratio = 11.98; 95% HPD CI: 4.55–23.4).

L2–L1 gender congruency (see Figure 5 B): A mismatch in grammatical gender between Spanish and French (*cama*$_{F}$ vs. *lit*$_{M}$) significantly reduced accuracy (Est. -1.45, 95% CI [-1.81, -1.10]). The odds of correct agreement were over four times higher in congruent than in incongruent cases (odds ratio 4.26; 95% HPD CI 2.938–5.93). No significant difference was found between fully matched pairs (e.g., *casa*$_{F}$ vs. *maison*$_{F}$) and cases with both gender options in French (e.g., *après-midi*$_{M\ or\ F}$, *an*$_{M}$*/année*$_{F}$).

Determiner type (see Figure 5 C): Accuracy varied across determiner types. Definite determiners (e.g., ***el*** *coche*, ***la*** *casa*) yielded the highest accuracy, followed by indefinite (e.g., ***un*** *coche*, ***una*** *casa*). Using definite determiners (*Det_typeDef*) as the reference, demonstratives (Est. -1.21, CI [-1.65, -0.76]), quantifiers (Est. -0.85, CI [-1.27, -0.42]), and

"other" (*otro*) determiners (Est. -0.84, CI [-1.58, -0.06]) all significantly reduced accuracy. Indefinite and possessive determiners showed more variable results, with wider intervals. Pairwise contrasts revealed significant differences between definite and demonstrative determiners (e.g., ***aquel*** *coche*, ***aquella*** *casa*) (odds ratio = 3.36; 95% HPD CI: 1.979–4.948), as well as between definite and quantifiers (e.g., ***algunos*** *coches*, ***varias*** *casas*) (odds ratio = 2.34; 95% HPD CI: 1.451–3.467), and between definite and other type (e.g., ***otros*** *coches*, ***otras*** *casas*) (odds ratio = 2.36; 95% HPD CI: 0.834–4.381). No reliable differences were observed between definite and indefinite determiners.

Noun gender (see Figure 5 D): Feminine nouns were associated with significantly lower accuracy than masculine ones (Est. -0.89, 95% CI [-1.23, -0.56]), indicating that feminine gender assignment remains more challenging.

These findings support previous observations in the literature that gender assignment improves with proficiency and is strongly influenced by the existence of masculine by default, by the presence of prototypical cues and by crosslinguistic consistency. Additionally, the type of determiner was shown to affect performance: while no significant difference was found between definite and indefinite determiners, other types —such as demonstratives and quantifiers— were associated with lower accuracy. This highlights the relevance of including determiner type in analyses of gender assignment in L2 Spanish.

5.2. Gender agreement: Noun_adjective (Research Question 2)

We fitted a Bayesian mixed-effects logistic regression model to examine gender agreement between nouns and adjectives. The model included random intercepts for participants (*Subject_ID*) and for lexical items (*Noun_lemma* and *Adj_lemma*), to account for subject- and item-level variability. The model showed satisfactory predictive accuracy (LOOIC: 741.5), with all Pareto *k* diagnostics indicating good reliability ($k < 0.7$). The dependent variable was whether the adjective matched the gender of the noun (Noun_adj_binary). Random intercepts were included for subject, noun lemma and adjective lemma. The results, based on 6,331 observations, showed a high overall probability of gender agreement of approximately 99%, with an intercept estimate of 8.53 (95% CI [6.78, 10.80]) on the logit scale. This intercept represents the predicted accuracy for masculine nouns with a definite determiner and congruent gender between L1 and L2 at the reference proficiency level (Beginner). Table 4 shows the results for fixed effects for the gender agreement model and presents the results for fixed effects for the gender assignment model. For each predictor, we report the estimated coefficient, its standard error (SE), and the 95% highest posterior density interval (HPD CI).

Table 4. Fixed effects for the gender agreement model (Bayesian logistic regression)

Predictor	**Est.**	**SE**	**95% HPD CI**
Intercept	8.53	1.02	[6.78, 10.80]
Level.L	1.92	0.55	[0.89, 3.03]
Level.Q	1.08	0.61	[-0.16, 2.28]
Protot_markAmb	-0.85	0.39	[-1.65, -0.10]
Protot_markNo_protot	-1.60	0.98	[-3.60, 0.24]
Compare_L2_L1both	-0.45	1.29	[-2.99, 2.07]
Compare_L2_L1different	-1.07	0.44	[-1.97, -0.21]
Noun_genFem	-1.64	0.44	[-2.55, -0.84]
Adj_Noun_Distance	-0.26	0.33	[-0.87, 0.41]
Pre_nounYES	-1.38	0.54	[-2.49, -0.39]

Concerning random effects, the standard deviation of the random intercepts was substantial for participants (sd = 2.32), noun (sd = 1.53) and adjective (sd = 1.14), indicating that variation across individuals and lexical items is not negligible. Once again, despite this variability, the fixed effects of proficiency and predictive variables remained robust, which strengthens the interpretation that these factors systematically influence gender assignment accuracy.

The model revealed a strong linear effect of **proficiency** level (Level.L: Est. 1.92, 95% CI [0.89, 3.03]) and a weak quadratic trend (Level.Q = 1.03, 95% CI [–0.17, 2.20]), indicating that accuracy improved steadily across levels, with a sharper increase from intermediate to advanced learners. Pairwise comparisons confirmed this pattern: there was little difference between Beginners and Intermediate learners, whereas advanced learners showed substantially higher odds of producing correct gender agreement (odds ratios = 0.07–0.09, HPD < 1). Figure 6 displays the effect of proficiency on gender agreement.

Figure 6. Effect of proficiency on gender agreement

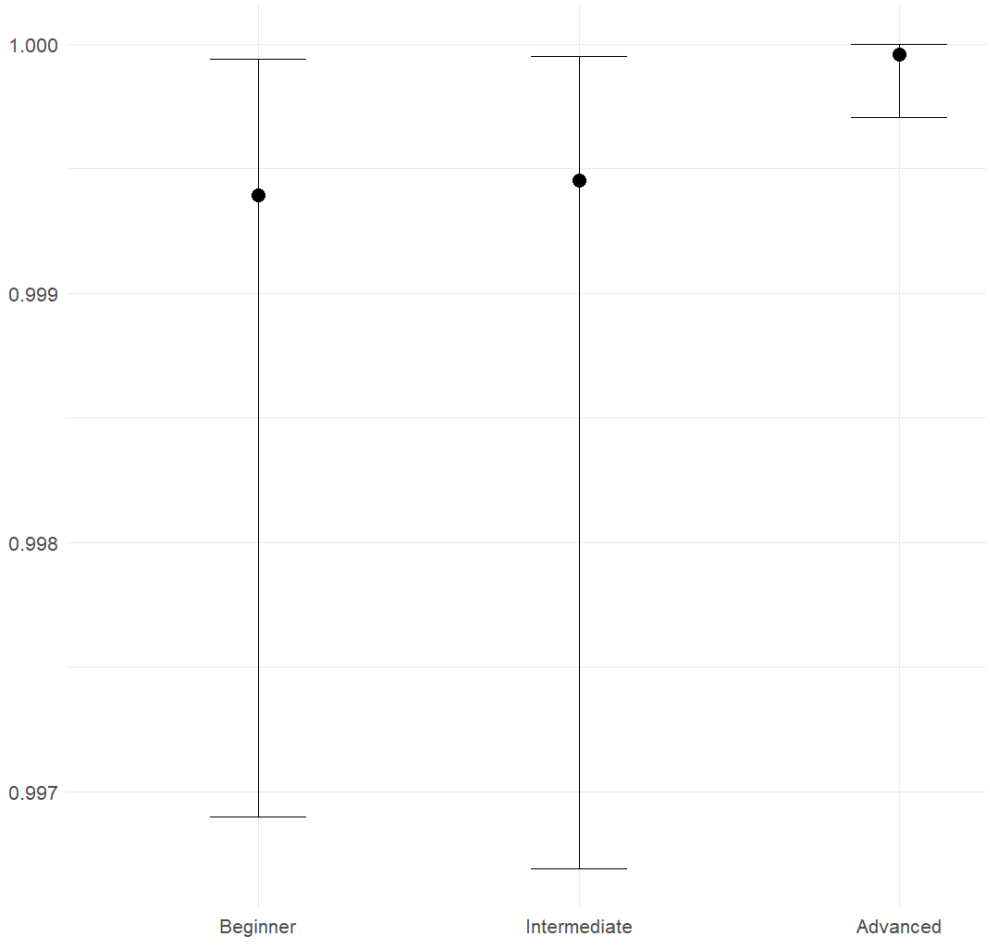
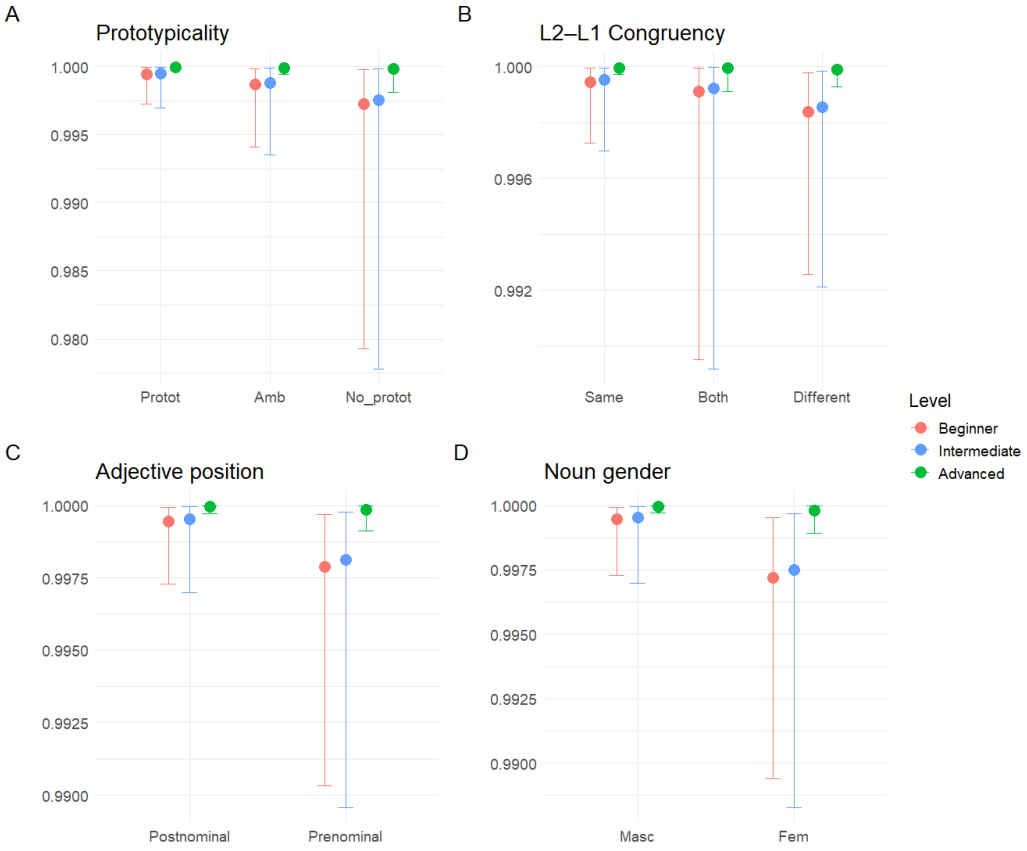


Figure 7 illustrates the effect of the other predictors, which are described in detail below the figure.

Figure 7. Effect of predictors on gender agreement by proficiency level



Morphological prototypicality (see Figure 7 A): Ambiguous gender marking on the noun (e.g, *camión*M, *calle*F) reduced accuracy compared to prototypical forms (Est. -0.85, 95% CI [-1.65, -0.10]), while nouns with non prototypical marking (e.g, *idioma*M, *mano*F) showed a similar negative trend, though with wider uncertainty (Est. -1.60, 95% CI [-3.60, 0.24]).

L2–L1 gender congruency (see Figure 7 B): Once again, a mismatch in grammatical gender between Spanish and French (*cama*F vs. *lit*M) significantly reduced accuracy (Est. -1.01, 95% CI [-1.85, -0.24]), while cases with both gender options in French (e.g, *après-midi*M *or* F, *an* M */année*F) did not differ significantly (Est. -0.45, 95% CI [-2.99, 2.07]) from fully matched pairs (e.g, *casa* F vs. *maison* F).

Adjective position (see Figure 7 C): adjectives in prenominal position were more likely to result in errors (Est. -0.26, 95% CI [-0.87, -0.41]), suggesting that canonical postnominal position may facilitate correct agreement. Additionally, greater linear distance between the noun and the adjective tended to decrease accuracy, although the credible interval included zero, therefore not yielding a significant effect (Est. -0.26, 95% CI [-0.87, 0.41]).

Noun gender (see Figure 7 D): Feminine nouns were also associated with lower agreement rates than masculine ones (Est. -1.65, 95% CI [−2.55, −0.84]), consistent with previous findings showing masculine defaulting in L2 production.

5.3. Gender agreement: Determiner_Adjective (Research Question 3)

To further examine gender agreement patterns, we analyzed a subset of noun phrases containing both a determiner and an adjective, focusing on whether these two elements agreed in gender. *Det_adj_agreement* was assessed by whether the adjective matched the grammatical gender of the noun as marked by the determiner (Yes = same gender: e.g., *la casa roja*, **el casa rojo*; No = mismatch: e.g., **la casa rojo, *el casa roja*).

The analysis revealed a clear asymmetry depending on whether the determiner had been correctly assigned. As shown in Table 5, when the determiner matched the grammatical gender of the noun (correct assignment), agreement with the adjective was nearly categorical (98.5 % for feminine nouns and 99.6 % for masculine nouns). In contrast, when the determiner was incorrectly assigned, agreement between the determiner and the adjective dropped dramatically, with percentages close to chance (47.1% for feminine nouns and 51.9 % for masculine nouns). These results suggest that learners generally apply gender agreement rules successfully when gender assignment is accurate. However, when they assign the wrong gender to the noun, their use of adjective agreement becomes unstable. This phenomenon likely reflects conflicting representations of gender, where the adjective may align with either the noun's actual gender or the erroneous one marked by the determiner.

Table 5. Distribution of determiner–adjective agreement by noun gender and accuracy of gender assignment

Noun gender	**Noun_det assignment**	**Det_adj agreement**	**Example**	**n**	**%**
Masculine	Correct	Yes	*El año pasado*	2207	99.6
Masculine	Correct	No	**El año pasada*	8	0.4
Masculine	Error	No	**La año pasado*	14	**51.9**
Masculine	Error	Yes	**La año pasada*	13	**48.1**
Feminine	Correct	Yes	*La calle ancha*	2084	98.5
Feminine	Correct	No	**La calle ancho*	31	1.5
Feminine	Error	No	**El calle ancha*	16	**47.1**
Feminine	Error	Yes	**El calle ancho*	18	**52.9**

A Bayesian multilevel logistic regression model was fitted to examine the predictors of determiner-adjective gender agreement (*Det_adj_agreement*) in L2 Spanish. The model included linguistic predictors (*Protot_mark*, *Compare_L2_L1*, *Noun_gen*, *Adj_Noun_Distance*, *Pre_noun*) and learner proficiency level (*Level*), with random intercepts for *Subject_ID*, *Noun_lemma*, and *Adj_lemma*.

The results show that the model converged well (all Rhat = 1.00), and the highest variance was observed at the participant level, indicating substantial individual variability. Regarding random effects, variability in random intercepts was notable for participants (sd = 3.68), nouns (sd = 2.16), and adjectives (sd = 2.28), suggesting that individual differences and lexical item variation are not negligible. Predicted probabilities were close to 1 across all proficiency levels, indicating that determiner–adjective agreement is overall highly accurate in the data. Among the fixed effects, *Level* (Level.L: Est. 3.36, 95% CI [1.36,

6.02]) and *Noun_gen* were the only reliable predictors: the linear trend for *Level* suggests that higher proficiency is associated with more consistent agreement, while *Noun_gen* shows that masculine nouns yield higher agreement accuracy than feminine ones (Est. -1.86, 95% CI [-3.55, -0.60]). Other predictors, including *Protot_mark*, *Compare_L2_L1*, and *Adj_Noun_Distance*, had wide credible intervals crossing zero, suggesting limited or uncertain effects.

However, due to the extremely unbalanced distribution of agreement outcomes (i.e., very few errors), model comparison using leave-one-out cross-validation (LOO) revealed poor reliability, with 67% of Pareto *k* values exceeding 0.7, suggesting that the model's pointwise predictive accuracy is unstable and that individual observations have disproportionate influence on the posterior estimates.

These results suggest that, while learners almost always maintain internal agreement between determiners and adjectives, future research should examine this phenomenon in larger and more diverse L2 Spanish corpora, ideally containing a higher number of gender assignment errors to allow for more reliable statistical modeling.

6. Discussion

The present study examined the accuracy of gender assignment and gender agreement in Spanish among L2 learners across different proficiency levels. Overall, the number of errors was relatively low across models in our corpus, which is likely related to the nature of the activities and the characteristics of the lexical items used. As described earlier, most of the nouns in our corpus share the same grammatical gender in both Spanish and French and display prototypical gender markings, conditions that are known to facilitate accurate gender assignment. Only a minority of the nouns (less than 30%) show a gender mismatch between the two languages, and only 5% carry a non-prototypical gender marker, of which only 0.33% are feminine. These patterns reduce the likelihood of assignment errors, especially considering the learners' well-documented tendency to default to the masculine form in ambiguous cases or with unfamiliar words, as discussed earlier. Furthermore, the texts produced by Beginner and Intermediate learners mostly relied on vocabulary taught in class, typically involving high-frequency and familiar nouns, which further limited the potential for errors. Still, occasional improvisation was observed among participants, particularly at more advanced levels, which may explain some of the individual variation in accuracy.

In the case of gender agreement, as predicted by earlier literature on avoidance in L2 production, we observed that some learners, mainly those at the Beginner level, actively avoided using adjectives. This was possibly a strategy to reduce the risk of producing errors. Moreover, only 31% of the adjectives in the corpus occurred in prenominal position, suggesting a general preference for more canonical postnominal structures. This pattern may reflect sensitivity to syntactic complexity, and it supports the idea that learner output is shaped not only by knowledge of grammatical rules but also by strategic choices aimed at minimizing difficulty. These findings point to the importance of considering both the linguistic properties of the input and learner behavior when analyzing performance. They also highlight the value of corpus-based approaches in identifying subtle patterns of avoidance and variability in learner data.

6.1. Noun_det assignment and Noun_adj_agreement

Regarding predictive variables, the results of the models for gender assignment and gender agreement (noun-adjective) highlight both shared and distinct patterns in how specific linguistic and learner-related variables impact gender assignment and agreement accuracy. Across both models, proficiency level emerged as a robust predictor, with a strong linear increase in accuracy as learners Advanced (Level.L: assignment = 1.96; agreement = 1.92), and a smaller but still positive nonlinear component in the assignment model (Level.Q = assignment = 0.75; agreement = 1.08). This indicates that although all groups performed relatively well overall, learners at higher levels showed significantly greater accuracy, especially between Intermediate and Advanced levels.

The grammatical gender of the noun itself had a consistent impact. Feminine nouns were significantly more error-prone in both models (assignment = -0.89; agreement = -1.64), which is consistent with the frequent default to the masculine form in uncertain or ambiguous contexts. This asymmetry underscores the influence of default strategies in learner production.

A very important factor was the match between L1 and L2 gender. Learners performed better when the gender of the noun was the same in both Spanish and their L1 (French), confirming that positive "surface" transfer can facilitate acquisition and further highlights the role of crosslinguistic influence. The models indicate a significant decline in accuracy when the L2 gender differs from the learner's L1 (assignment = -1.45; agreement = -1.07), suggesting that transfer effects remain persistent even at higher proficiency levels. In contrast, no meaningful difference was found when the gender is ambiguous in the L1 (e.g., *après-midi*: 'afternoon', *jour/journée*: 'day'), which might reflect greater reliance on L2 cues in such cases.

The type of gender marking on the noun also played a crucial role. Compared to prototypically marked nouns, those with ambiguous gender markers led to a moderate decrease in performance (assignment = -0.79; agreement = -0.85), while non-prototypically marked nouns resulted in a much larger drop in assignment accuracy (-2.49) and showed a similar negative trend in agreement (-1.60), though with wider uncertainty. This finding supports previous work suggesting that prototypical morphological cues facilitate gender processing, particularly at earlier stages of acquisition. Interestingly, this effect was not significant in the agreement model, suggesting that the morphological form of the noun may play a less central role during adjective agreement than in the assignment of gender.

Syntactic configurations appear to influence agreement accuracy. Adjectives in prenominal position (Est. -0.68) were associated with more agreement errors, as well as the linear distance between noun and adjective (Est. -1.38), but the latter showed only a marginal effect. This pattern suggests that learners may rely on linear or default strategies when producing adjective-noun sequences and that placing the adjective before the noun probably increases processing demands, possibly due to the need to project or anticipate morphosyntactic features.

6.2. Noun_det assignment and Det_adj_agreement

The analysis of determiner-adjective gender agreement (*Det_adj_agreement*) provides additional insights into the challenges learners face when maintaining grammatical consistency across elements within the noun phrase. When we focus specifically on NPs that contain both determiners and adjectives (e.g., *la casa roja*), an interesting pattern

emerges. We observed that noun-adjective agreement is present in approximately 99% of the phrases where gender assignment is accurate—that is, when the determiner matches the gender of the noun. However, in cases where the determiner does not reflect the correct gender of the noun, the gender of the adjective appears to be assigned in a more arbitrary fashion.

Our data raise important questions about the mechanisms underlying gender agreement when the gender of the noun is uncertain or unknown. In such contexts, it seems plausible that gender agreement may be deprioritized, and that speakers may assign gender to adjectives in a way that is less systematic.

This result may indicate that learners rely on different mechanisms for gender assignment and agreement, possibly involving separate stages of lexical retrieval and morphosyntactic encoding. The limited number of errors observed, however, constrains the statistical power of the analysis, preventing strong conclusions about the predictive role of learner level or lexical factors. Future research using larger corpora or targeted elicitation tasks will be essential to confirm whether these tendencies reflect a systematic compensatory strategy or simply random variation in performance. In other words, further analysis is needed to reach more robust conclusions regarding the mechanisms underlying gender assignment and agreement, and to determine whether these should be treated as two distinct phenomena.

6.3. Summary, limits of the study and future research

Together, these results show that while learners generally succeed in both gender assignment and agreement, their performance is shaped by a complex interaction of factors: morphological transparency, crosslinguistic similarity, syntactic structure, and general proficiency. Our results are in line with the bulk of research that prototypical categories of our variables are associated with better gender assignment and agreement than their non-prototypical counterparts.

This study thus provides empirical support for theoretical accounts of gender assignment and agreement in L2, reinforcing patterns that have been observed in experimental tasks and small-scale production studies. By analyzing a medium-scale learner corpus, it extends previous findings and illustrates the added value of corpus-based approaches in capturing a wider range of learner strategies. The ability to examine patterns across thousands of authentic noun phrases allows for more robust generalizations and for testing the relative impact of multiple interacting variables.

Despite these contributions, the present analysis was limited to the most widely studied predictors in the literature. Future research should consider additional item-specific variables such as the absolute frequency of the noun, its most common collocational patterns, whether is a mass noun or a count noun, or its typical number (singular vs. plural). These lexical and distributional properties may further modulate learners' performance. Expanding the scope of this study to include such factors would provide a more nuanced understanding of the interface between lexical knowledge, morphosyntactic encoding, and error patterns in L2 production.

The corpus analysed in this study has the advantage of identifying the major challenges learners of Spanish face in gender assignment and agreement, as well as pointing the most problematic nouns and structures studied at each proficiency level, allowing for the development of instructional materials aimed at preventing the most frequent errors. This

approach effectively bridges corpus analysis with L2 teaching. However, it also has its limitations. Future research should explore corpora that include non-evaluative texts where participants are more inclined to take risks and less likely to avoid making errors. This approach could provide a more comprehensive understanding of how learners navigate gender agreement in less controlled, more naturalistic settings.

Finally, future research should also explore other language combinations to examine the robustness of transfer effects across different typological pairings.

7. Conclusions

This study examined gender assignment and agreement in advanced L2 Spanish learners, focusing on the interplay between the gender marked on determiners and on adjectives. While accuracy was high overall, certain features that are well-documented in the literature also emerged as consistent sources of difficulty in our corpus, such as feminine gender, non-prototypical morphology, and gender mismatches between L1 and L2. Overall, the results suggest that gender assignment and agreement may be related but are not fully explained by a single process.

The study also highlights the value of medium-scale corpus analysis for uncovering systematic patterns in L2 gender processing. One of the key advantages of corpus analysis within the context of L2 courses is that it allows instructors to identify and focus on the types of errors learners actually make. This enables the incorporation of concrete examples into classroom instruction, emphasizing frequent noun phrases that nonetheless trigger more errors.

Overall, more extensive corpora, ideally using richer datasets that capture a wider range of learner performance, are needed to reach firmer conclusions about the mechanisms of gender assignment and agreement, and to determine whether these should be treated as separate phenomena.

Maura Cruz Enríquez
Département Sciences humaines, Lettres et Communication
Université TÉLUQ
5800, rue Saint-Denis, Montréal (Québec), H2S 3L5, Canada
Maura.CruzEnriquez@teluq.ca
ORCID: 0000-0001-8151-7383

François Pichette
Département Sciences humaines, Lettres et Communication
Université TÉLUQ
455, rue du Parvis, Québec (Québec), G1K 9H6, Canada
Francois.Pichette@teluq.ca
ORCID: 0000-0002-6203-050X

References

Afonso, O., Domínguez, A., Álvarez, C. J., & Morales, D. (2014). Sublexical and Lexico-Syntactic Factors in Gender Access in Spanish. *Journal of Psycholinguistic Research*, *43*(1), 13-25. Scopus. https://doi.org/10.1007/s10936-012-9236-0

Alarcón, I. (2009). The Processing of Gender Agreement in L1 and L2 Spanish: Evidence from Reaction Time Data. *Hispania*, *92*(4), 814-828.

Ayoun, D. (2007). The second language acquisition of grammatical gender and agreement. In D. Ayoun (Ed.), *French Applied Linguistics* (p. 130-170). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.16.10ayo

Bartning, I. (2000). Gender agreement in L2 French: Pre-advanced vs advanced learners. *Studia Linguistica*, *54*(2), 225-237. https://doi.org/10.1111/1467-9582.00062

Beck, L. (2016, *Answer to « What percentage of nouns ending in -a are masculine and what percentage of nouns ending in -o are feminine? »* [Billet]. Spanish Language Stack Exchange. https://spanish.stackexchange.com/a/17436

Bellamy, K., & Parafita Couto, M. C. (2022). Gender assignment in mixed noun phrases. In D. Ayoun (Ed.), *The Acquisition of Gender: Crosslinguistic perspectives* (p. 13-48). John Benjamins Publishing Company. https://doi.org/10.1075/sibil.63.02bel

Berdasco Muñoz, E. (2013). *Gender Assignment in L2 Spanish Interlanguage: Comparison Among English, German, and French L1 Adult Learners*. http://hdl.handle.net/10657/1021

Bordag, D., Opitz, A., & Pechmann, T. (2006). Gender processing in first and second languages: The role of noun termination. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(5), 1090-1101. https://doi.org/10.1037/0278-7393.32.5.1090

Bruhn de Garavito, J., & White, L. (2002). The second language acquisition of Spanish DPs: The status of grammatical features. *The Acquisition of Spanish Morphosyntax: The L1/L2 Connection*, 153-178. https://doi.org/10.1007/978-94-010-0291-2_6

Bürkner, P.-C. (2017). brms: An *R* Package for Bayesian Multilevel Models Using *Stan*. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

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

Corbett, G. G. (1991). *Gender*. Cambridge University Press. https://doi.org/10.1017/CBO9781139166119

Costa, A., Kovacic, D., Franck, J., & Caramazza, A. (2003). On the autonomy of the grammatical gender systems of the two languages of a bilingual. *Bilingualism: Language and Cognition*, *6*(3), 181-200. https://doi.org/10.1017/S1366728903001123

Cruz Enríquez, M., Couture-Matte, R., & Pichette, F. (2025). Aplicación de Python al análisis automático de corpus: dos estudios de caso en la enseñanza de lenguas. In P. Román-Graván, M. Puig-Gutiérrez, C. Siles-Rojas, & S. Salas-Ruiz (Eds.), *Investigación e innovación educativa. Experiencias en la era digital* (p. 239-254). Dykinson. https://doi.org/10.14679/4025

Dewaele, J.-M., & Véronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition*, *4*(3), 275-297. https://doi.org/10.1017/S136672890100044X

Dominguez, A., Cuetos, F., & Segui, J. (1999). The Processing of Grammatical Gender and Number in Spanish. *Journal of Psycholinguistic Research*, *28*(5), 485-498. https://doi.org/10.1023/A:1023216326448

Ecke, P. (2022). Grammatical gender and article use in beginning learners of German. In D. Ayoun (Ed.), *The Acquisition of Gender. Crosslinguistic perspectives* (. 127-156). John Benjamins Publishing Company. https://doi.org/10.1075/sibil.63.06eck

Eddington, D., & Hualde, J. (2008). El Abundante Agua: Hermaphroditic Spanish Nouns. *Studies in Hispanic and Lusophone Linguistics*, *1*, 5-31. https://doi.org/10.1515/shll-2008-1004

Edmonds, A., & Gudmestad, A. (2018). Gender marking in written L2 French: Before, during, and after residence abroad. *Study Abroad Research in Second Language Acquisition and International Education*, *3*(1), 58-83. https://doi.org/10.1075/sar.16018.edm

Edmonds, A., & Gudmestad, A. (2021). The role of transparency in grammatical gender marking among stay abroad learners of Spanish and French. In M. Howard (Ed.), *Study Abroad and the Second Language Learner: Expectations, Experiences and Development* (pp. 233-253). Bloomsbury Academic. https://doi.org/10.5040/9781350104228.ch-012

Edmonds, A., Gudmestad, A., & Metzger, T. (2020). A Longitudinal Study of Grammatical-Gender Marking in French as an Additional Language. *Applied Linguistics*, *41*(5), 733-755. https://doi.org/10.1093/applin/amz036

Ellis, C., Conradie, S., & Huddlestone, K. (2012). The acquisition of grammatical gender in L2 German by learners with Afrikaans, English or Italian as their L1. *Stellenbosch Papers in Linguistics*, *41*, 17-27. https://doi.org/10.5774/41-0-131

Ellis, N. C. (2019). *Essentials of a Theory of Language Cognition*. https://doi.org/10.1038/s41562-017-0163

Foote, R. (2015). The production of gender agreement in native and L2 Spanish: The role of morphophonological form. *Second Language Research*, *31*(3), 343-373. https://doi.org/10.1177/0267658314565691

Franceschina, F. (2005). Fossilized Second Language Grammars. In *Lald.38*. John Benjamins. https://benjamins.com/catalog/lald.38. https://doi.org/10.1075/lald.38

Gómez Carrero, T., & Ogneva, A. (2024). Who does it better?: The acquisition of Spanish grammatical gender by L1 English and L1 Russian adults. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, *37*(1), 143-171. https://doi.org/10.1075/resla.21031.gom

Gries, S. Th. (2021). (Generalized Linear) Mixed-Effects Modeling: A Learner Corpus Example. *Language Learning*, *71*(3), 757-798. https://doi.org/10.1111/lang.12448

Gries, S. Th., & Ellis, N. C. (2015). Statistical Measures for Usage-Based Linguistics. *Language Learning*, *65*(S1), 228-255. https://doi.org/10.1111/lang.12119

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second language research*, *28*(2), 191-215. https://doi.org/10.1177/0267658312437990

Gudmestad, A., Edmonds, A., & Metzger, T. (2019). Using Variationism and Learner Corpus Research to Investigate Grammatical Gender Marking in Additional Language Spanish. *Language Learning*, *69*(4), 911-942. https://doi.org/10.1111/lang.12363

Harris, J. W. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, *22*(1), 27-62.

Holmes, V. M., & de la Bâtie, D. (1999). Assignment of grammatical gender by native speakers and foreign learners of French. *Applied Psycholinguistics*, *20*(4), 479-506. https://doi.org/10.1017/S0142716499004026

Honnibal, M., Montani, I., Van Landeghem, S., & Boyd, A. (2020). *spaCy Industrial-strength Natural Language Processing in Python*. https://spacy.io/

Joan Casademont, A., Gagné, N., & Viladrich Castellanas, È. (2022). Allers-retours entre recherche et pratique : Analyse de besoins et capsules de microapprentissage en apprentissage d'une langue tierce ou additionnelle. *Médiations & Médiatisations*, *12*, 8-33. https://doi.org/10.52358/mm.vi12.288

Johannessen, J. B., Lundquist, B., Rodina, Y., Tengesdal, E., Kaldhol, N. H., Türker, E., & Fyndanis, V. (2024). Cross-linguistic effects in grammatical gender assignment and predictive processing in L1 Greek, L1 Russian, and L1 Turkish speakers of Norwegian as a second language. *Second Language Research*, *41*(1), 217-259. https://doi.org/10.1177/02676583241227709

Karmiloff-Smith, A. (1979). *A Functional Approach to Child Language*. Cambridge University Press & Assessment. https://www.cambridge.org/universitypress/subjects/languages-linguistics/psycholinguistics-and-neurolinguistics/functional-approach-child-language-study-determiners-and-reference

Keating, G. D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning*, *59*(3), 503-535. https://doi.org/10.1111/j.1467-9922.2009.00516.x

Kempe, V., & Brooks, P. J. (2008). Second Language Learning of Complex Inflectional Systems. *Language Learning*, *58*(4), 703-746. https://doi.org/10.1111/j.1467-9922.2008.00477.x

Klassen, R., Lundquist, B., & Westergaard, M. (2023). L1 Grammatical Gender Variation through the Representation in the Lexicon. *Journal of Psycholinguistic Research*, *52*(2), 359-380. https://doi.org/10.1007/s10936-022-09867-7

Lemhöfer, K., Schriefers, H., & Hanique, I. (2010). Native language effects in learning second-language grammatical gender: A training study. *Acta Psychologica*, *135*(2), 150-158. https://doi.org/10.1016/j.actpsy.2010.06.001

Lemhöfer, K., Spalek, K., & Schriefers, H. (2008). Cross-language effects of grammatical gender in bilingual word recognition and production. *Journal of Memory and Language*, *59*(3), 312-330. https://doi.org/10.1016/j.jml.2008.06.005

Lemmerth, N., & Hopp, H. (2019). Gender processing in simultaneous and successive bilingual children: Cross-linguistic lexical and syntactic influences. *Language Acquisition*, *26*(1), 21-45. https://doi.org/10.1080/10489223.2017.1391815

Leśniewska, J., & Pichette, F. (2014). Songs vs. stories: Impact of input sources on ESL vocabulary acquisition by preliterate children. *International Journal of Bilingual Education and Bilingualism*, *19*, 1-17. https://doi.org/10.1080/13670050.2014.960360

Lichtman, K. (2009). *Acquisition of Attributive and Predicative Adjective Agreement in L2 Spanish*. https://www.academia.edu/64312637/Acquisition_of_Attributive_and_Predicative_Adjective_Agreement_in_L2_Spanish

Lyster, R. (2006). Predictability in French gender attribution: A corpus analysis. *Journal of French Language Studies*, *16*(1), 69-92. https://doi.org/10.1017/S0959269506002304

McCarthy, C. (2008). Morphological variability in the comprehension of agreement: An argument for representation over computation. *Second Language Research*, *24*(4), 459-486. https://doi.org/10.1177/0267658308095737

McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and STAN* (Chapman and Hall/CRC. https://doi.org/10.1201/9780429029608

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

Morales, L., Paolieri, D., Cubelli, R., & Bajo, M. T. (2014). Transfer of Spanish grammatical gender to English: Evidence from immersed and non-immersed bilinguals. *Bilingualism: Language and Cognition*, *17*(4), 700-708. https://doi.org/10.1017/S1366728914000017

Muñoz-Liceras, J., Fernández Fuertes, R., Perales, S., Pérez-Tattam, R., & Spradlin, K. (2008). Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional–lexical mixings. *Lingua*, *118*, 827-851. https://doi.org/10.1016/j.lingua.2007.05.006

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 516-524. https://doi.org/10.3758/BF03195598

Norris, J. M. (2015). Statistical Significance Testing in Second Language Research: Basic Problems and Suggestions for Reform. *Language Learning*, *65*(S1), 97-126. https://doi.org/10.1111/lang.12114

Norris, J. M., Ross, S. J., & Schoonen, R. (2015). Improving Second Language Quantitative Research. *Language Learning*, *65*(S1), 1-8. https://doi.org/10.1111/lang.12110

Paolieri, D., Demestre, J., Guasch, M., Bajo, T., & Ferré, P. (2020). The gender congruency effect in Catalan–Spanish bilinguals: Behavioral and electrophysiological evidence. *Bilingualism (Cambridge, England)*, *23*(5), 1045-1055. https://doi.org/10.1017/S1366728920000073

Pasta, D. (2009). *Learning When to Be Discrete: Continuous vs. Categorical Predictors*. SAS Global Forum, Paper *248*, 1-10. https://support.sas.com/resources/papers/proceedings09/248-2009.pdf

Pérez-Leroux, A. T., López, Y. Á., Barreto, M., Cuza, A., Marinescu, I., Yang, J., & Colantoni, L. (2023). The Phonetic and Morphosyntactic Dimensions of Grammatical Gender in Spanish Heritage Language Acquisition. *Heritage Language Journal*, *20*(1), 1-37. https://doi.org/10.1163/15507076-bja10017

Pérez-Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language*, *18*(3), 571-590. https://doi.org/10.1017/S0305000900011259

Pichette, F. (2023, June 14). *Gender assignment in L2 French by speakers of German and Polish*. XV Congreso Internacional de Lingüística General (XV CILG 2023), Madrid.

Pienemann, M., & Keßler, J. U. (2013). Processability theory. In S. M. Gass & A. Mackey (Eds.), *The Routledge Handbook of Second Language Acquisition* (pp. 228-246). https://doi.org/10.4324/9780203808184

Pienemann, M., & Lenzing, A. (2014). Processability theory. In B. VanPatten & J. Wil liams (Eds.), *Theories in Second Language Acquisition: An Introduction* (2nd ed., pp. 159-179). https://doi.org/10.4324/9780203628942

Python Software Foundation. (2025). *Python (Version 3.13.4)* [Logiciel]. https://www.python.org/

R Core Team. (2022). *R: The R Project for Statistical Computing* [Logiciel]. https://www.r-project.org/

RAE. (2024a, January 19). *Adjetivos, determinantes y pronombres (II). Cierto, determinado, semejante, otro, demás | Nueva gramática de la lengua española*. «Nueva gramática de la lengua española». https://www.rae.es/gramática/sintaxis/adjetivos-determinantes-y-pronombres-ii-cierto-determinado-semejante-otro-demás

RAE. (2024b, January19). *Posición del adjetivo en el grupo nominal (I). Distinciones fundamentales | Nueva gramática de la lengua española*. «Nueva gramática de la lengua española». https://www.rae.es/gramática/sintaxis/posición-del-adjetivo-en-el-grupo-nominal-i-distinciones-fundamentales

Ragnhildstveit, S. (2017). Gender Assignment and L1 Transfer in Norwegian Second Language Learners' Written Performance. In A. Golden, S. Jarvis, & K. Tenfjord (Eds.), *Crosslinguistic and Distinctive Patterns of Language Learning. Findings and Insights from a Learner Corpus* (pp. 110-154). Multilingual Matters. https://doi.org/10.2307/jj.22730646.10

Rojo, G. (2021). *Introducción a la lingüística de corpus en español*. Routledge. https://doi.org/10.4324/9781003119760

Sabourin, L., Stowe, L., & Haan, G. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, *221*, 1-29. https://doi.org/10.1191/0267658306sr259oa

Salamoura, A., & Williams, J. N. (2007). The representation of grammatical gender in the bilingual lexicon: Evidence from Greek and German. *Bilingualism: Language and Cognition*, *10*(3), 257-275. https://doi.org/10.1017/S1366728907003069

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

Schachter, J. (1974). An Error in Error Analysis. *Language Learning*, *24*(2), 205-214. https://doi.org/10.1111/j.1467-1770.1974.tb00502.x

Seigneuric, A., Zagar, D., Meunier, F., & Spinelli, E. (2007). The relation between language and cognition in 3- to 9-year-olds: The acquisition of grammatical gender in French. *Journal of Experimental Child Psychology*, *96*(3), 229-246. https://doi.org/10.1016/j.jecp.2006.12.003

Sotiropoulou, M. S. (2022). The chance of matching gender between French, Greek, and Spanish nouns. In E. Babatsouli (Ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2022* (pp. 61-67). University of Louisiana at Lafayette.

Teschner, R. V., & Russell, W. M. (1984). The gender patterns of Spanish nouns: An inverse derivational view. *Hispanic Linguistics*, *1*(2), 115-132.

Tucker, G. R., Lambert, W. E., & Rigault, A. (1977). *The French speaker's skill with grammatical gender: An example of rule-governed behavior*. Mouton. https://doi.org/10.1515/9783110805413

Valenzuela, E., Kozlowska, M., & Leung, Y.-K. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics*, *25*, 105-133. https://doi.org/10.1017/S0142716404001067

Vanhove, J. (2017). The Influence of Standard and Substandard Dutch on Gender Assignment in Second Language German. *Language Learning*, *67*(2), 431-460. https://doi.org/10.1111/lang.12230

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413-1432. https://doi.org/10.1007/s11222-016-9696-4

Velnić, M. (2020). Acquisition of a Transparent Gender System: A Comparison of Italian and Croatian. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.571674

White, L., Valenzuela, E., Kozlowska–Macgregor, M., & Leung, Y.-K. I. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics*, *25*(1), 105-133. https://doi.org/10.1017/S0142716404001067

Wu, R., & Schiller, N. O. (2023). Gender Congruency Effects in Spanish: Behavioral Evidence from Noun Phrase Production. *Brain Sciences*, *13*(4), 696. https://doi.org/10.3390/brainsci13040696