Universidad de Valladolid

FACULTAD de FILOSOFÍA Y LETRAS
DEPARTAMENTO de FILOLOGÍA INGLESA
Grado en Estudios Ingleses

TRABAJO DE FIN DE GRADO

Decoding gender: A study of gender in code-switched agreement structures

Patricia Carro García

Tutora: Raquel Fernández Fuertes

2017-2018

Abstract

This dissertation is focused on the perception of grammatical gender by bilingual speakers of English and Spanish from different linguistic profiles. The study was carried out gathering data by means of an experimental task. This task was used to determine what kinds of agreement strategies bilingual speakers prefer in agreement structures (i.e., subject-copula verb-subject complement structures) where their two languages are mixed. The results obtained from the analysis of the data indicate that the grammatical gender properties of English and Spanish influence the way speakers perceive gender agreement in agreement structures, and therefore, that these properties have an impact on the processing of the two languages.

Keywords: Code Switching, grammatical gender, English, Spanish, bilingual speakers, Heritage bilingualism

Resumen

Este trabajo se centra en la percepción del género gramatical por parte de hablantes bilingües de inglés y español de distintos perfiles lingüísticos. El estudio se llevó a cabo recogiendo datos por medio de un test experimental. Este test se usó para determinar qué tipos de estrategias de concordancia prefieren los hablantes bilingües para estructuras que siguen el patrón sujeto-verbo copulativo-atributo, en las que se mezclan sus dos lenguas. Los resultados obtenidos del análisis de datos indican que las propiedades del género gramatical del inglés y el español influyen en cómo los hablantes perciben la concordancia de género en estas estructuras y, por lo tanto, que estas propiedades afectan el procesamiento de las dos lenguas.

Palabras clave: Alternancia de códigos, género gramatical, inglés, español, hablantes bilingües, bilingüismo de herencia

CONTENTS

1. Introduction

Over the past few years, research on bilingualism has increased, and with it our understanding of what being a bilingual entails. One of the issues that has been most widely studied has been Code Switching, a phenomenon by which bilingual speakers move in and out of their different languages (MacSwan, 2000, p. 38), and research on this phenomenon has focused on discovering the rules by which it abides. There are many different theories on the matter, but this study will adopt MacSwan's view, which is the most widely accepted nowadays and which considers that Code Switching is only constrained by the grammars of the languages involved (2000, p. 43).

Many researchers have focused on the Code Switching that involves two languages with different grammatical gender systems. This is the case of studies like Vanden Wyngaerd's (2016) which deals with Code Switching between Dutch and French within the determiner phrase (DP), or Liceras et al.'s (2008) which focuses on code switched Spanish-English DPs. Both these studies used experimental data, more specifically, data gathered by using Acceptability Judgment tasks.

Within this theoretical frame, the present dissertation will focus on the property of grammatical gender within agreement structures (i.e. subject-copula verb-subject complement structures) that contain Code Switching between an English subject and Spanish adjectival subject complement, as illustrated in (1).

(1) The moon *es bonita*.

[The moon is beautiful] (Klassen & Liceras, 2016, p. 78)

The study aims to determine what type of agreement strategy is preferred by different groups of bilingual adults: two groups of heritage speakers, one living in the U.S.A. and another living in Spain, and two groups of sequential bilinguals, a group whose first language is Spanish (L1 Spanish speakers) and another one whose first language is English (L1 English speakers).

The dissertation is divided into 7 main sections. Section 2 offers a theoretical background where Code Switching and its main approaches are described, followed by an overview of how grammatical gender works in both Spanish and English. Following this theory, section 3 presents the four hypotheses that this study will be testing. Section 4 explains the methodology followed, which includes a description of the four groups of

participants tested, as well as of the Acceptability Judgment task used in order to do so. In section 5, the results obtained from the analysis of the data gathered are presented, and this is followed by section 6, where a discussion of these results is offered. Finally, section 7 presents a conclusion to the study carried out in this dissertation. A final section offers a list of the works cited through the course of this study, and an appendix included at the end lists all the abbreviations and acronyms used.

2. Background

2.1. Code Switching: restrictions and use

Code Switching is a phenomenon which appears in bilingual speech and in situations of languages in contact. MacSwan (2000) defines it as "a speech style in which fluent bilinguals move in and out of two (or more) languages" (p. 38). The phenomenon is illustrated in the following sentences:

(1) Yo anduve *in a state of shock* por dos días.
    [I went around in a state of shock for two days]   (Pfaff, 1979, p. 296)

(2) Sometimes I'll start a sentence in Spanish *y termino en español*.
    [Sometimes I'll start a sentence in Spanish and finish in Spanish]   (Poplack, 1980, p. 594)

These examples show two different grammatical points in which Code Switching can appear. In (1), the subject and the verb are in Spanish, a switch to English occurs in the obligatory adjunct *in a state of shock* and this is followed by a switch back to Spanish in the second adjunct. In (2), the switch between English and Spanish occurs from the conjunction that unites the two propositions onwards. The case in (1) is what is called intra-sentential Code Switching, while the case in (2) is referred to as inter-sentential Code Switching.

Code Switching was originally perceived negatively and considered a sign of lack of proficiency of the bilingual speaker. Weinreich (1953) maintained that the ideal bilingual "switches from one language according to appropriate changes in the speech situation [...] but not in an unchanged speech situation and certainly not within a single sentence" (as cited in Muysken, 2000, p. 1). Later studies have debunked this notion. In her 1980 seminal study, "Sometimes I'll start a sentence in Spanish y termino en español: toward a typology of code-switching", Poplack examined Code Switching from different

2

speakers within a Puerto Rican community in the United States. In this study, she concluded that a high level of bilingual competence is required in order to produce Code Switching. Moreover, she also found relevant statistical evidence that intra-sentential switches are more frequent in balanced bilinguals than they are in Spanish-dominant bilinguals (p. 609).

There have been many proposals put forward to explain the phenomenon of Code Switching (e.g. what constrains it, at which grammatical points it can occur). The most prominent approaches at the moment, and the ones that will be discussed, are the Matrix Language Frame model and the Generative approach proposed as part of the Minimalist Program.

2.1.1. The Matrix Language Frame model

The Matrix Language Frame model (henceforth, MLF) was developed by Myers-Scotton in the early nineties. The main principle underlying this theory is the Uniform Structure Principle, as stated below, which applies to unilingual speech as well as to bilingual speech:

> The Uniform Structure Principle:
> A given constituent type in any language has a uniform abstract structure and the requirements of well-formedness for this type must be observed whenever the constituent appears. In bilingual speech, the structures of the Matrix Language are always preferred, but some Embedded structures […] are allowed if Matrix Language clause structure is observed.   (Myers-Scotton, 2006, p. 243)

This model highlights the asymmetric relationship between the languages involved in the process of Code Switching. And the basic idea behind it is the following: A grammatical frame is provided by one language and elements from another language are inserted into that structure. The language that provides the grammatical frame or structure is referred to as the Matrix Language (henceforth, ML) and the language that provides the inserted elements is called the Embedded Language (henceforth, EL). Myers-Scotton and Jake (2009, p. 338) state that, within a corpus, "the ML may vary from clause to clause".

The MLF makes a division between content and system morphemes. Content morphemes are "those that either assign or receive thematic roles" (Myers-Scotton, 2006, p. 244). Therefore, verbs and nouns are prototypical content morphemes, although discourse markers are also included in this category because the meaning of the elements that follow these morphemes will be restricted by them. Content morphemes are often

provided by the EL. System morphemes, on the other hand, are "all affixes (bound morphemes) and some function words that stand alone" (Myers-Scotton, 2006, p. 245). Some adverbs and prepositions are also included in this category. These morphemes come from the ML, as stated by the Morpheme Order Principle and the System Morpheme Principle (Myers-Scotton, 2006, p. 244). The sentence shown in (3) is presented as an example to support the presence of an ML and an EL. The fact that the English words *language Italian* do not follow the English word order is evidence that the morpheme order is provided by Italian, which is the ML.

(3) No porque quiero dispressare a mi *language Italian*.
[Not that I want to undervalue my Italian language]   (Clyne, 2003, p. 87)

The main critique that has been made to the MLF is that some grammatical examples are unexplainable under this model, as in (4), while many ungrammatical examples are allowed, as in (5). MacSwan (2005, p. 10) argues that this is due to the sole reliance on naturalistic data, which provides positive evidence but is unable to provide negative evidence.

(4) Tus *coworkers haven't had a vacation yet, righ*t?
[Your coworkers haven't had a vacation yet, right?]   (MacSwan, 2005, p. 8)
(5) *The students had *visto la película italiana*.
[The students had seen the Italian movie.]   (MacSwan, 2005, p. 9)

In (4), both English and Spanish contribute system morphemes (i.e. Spanish contributes the determiner *tus* and English contributes the verb *haven't had*), which makes the ML unclear. The sentence is grammatical regardless of which of the two is the ML, which would suggest that the ML "does not play a role in determining grammaticality" (MacSwan, 2005, p. 8). Meanwhile, the example in (5) is accepted under the MLF, since it presents a Spanish embedded structure within an English clause, but it is ill-formed.

Additionally, the MLF assumes a hierarchy between languages which is accompanied by a set of rules that do not apply to unilingual speech: that the ML structure will always be preferred, and what types of morphemes are provided by which language. Implying that Code Switching works according to a separate set of rules would mean that there is

a separate grammar by which this phenomenon abides. That is, in order to explain Code Switching we need to resort to rules that are not present if unilingual speech were considered. This premise poses a theoretical weakness.

In the following section, the main ideas offered by the Minimalist Program approach will be presented, which solve some of the problems that arise from the MLF model.

2.1.2. The Minimalist Program

A different view from the one defended by the MLF is the one proposed within the Minimalist Program (henceforth, MP), which argues that there is not a separate grammar by which Code Switching abides. Instead, the MP view defends that "nothing constrains code switching apart from the requirements of the mixed grammars" (MacSwan, 2000, p. 43). That is, in order to explain Code Switching we need to resort to the same rules that are used to analyze unilingual speech. An important difference between this model and the MLF is that the MP considers that both naturalistic and experimental data are essential to form a theory of Code Switching, since only the latter is able to provide negative evidence.

The MP proposes that there are two main components in human language: a computational system, which is the group of rules that organize the lexicon and is universal across languages, and a lexicon, which constitutes the vocabulary of a language and varies from one language to another. Derivations, or structures, are formed by a series of operations. The operation *Select* chooses the items from the lexicon that will participate in the derivation, and this subset of items is referred to as the *Lexical Array*. *Merge* forms syntactic objects from the *Lexical Array* and, later, *Move* is applied to build new structures. *Move* is triggered by feature valuation. Strong features result in overt movement that is realized phonetically, and weak features result in covert movement which is not realized phonetically.

MacSwan (2000, p. 45) sees Code Switching as the "consequence of mixing two lexicons in the course of a derivation." Elements from both lexicons are selected into the *Lexical Array* and they are feature checked in the same way they would in unilingual speech. Therefore, the word order and the structure of the sentence depends on the feature strength of the different lexical components, which varies according to the language to which they belong. MacSwan also develops a theorem to explain ungrammaticality in certain code-switched structures: the PF (Phonological Form) Disjunction Theorem. Under the PF Disjunction Theorem, Code Switching is not possible at the PF level

because it would "generate 'unpronounceable' elements which violate FI (Full Interpretation)" (p. 45). This explains why sentences like the one in (6) are non-viable.

(6) *Juan está *eat*-iendo

    [Juan is eating]   (MacSwan, 2000, p. 46)

Elements like *com+ie+ndo* constitute X^0 elements, they are formed at the PF level. Switches below X^0 are prohibited because they are impossible to realize phonetically.

During the course of this study, the Minimalist approach to Code Switching will be applied. The hypotheses and the analysis of the data will be based on the theory discussed above, and along these grounds the experimental data that are at the bases of this study will be explained. Since grammatical gender is the feature with which this dissertation is concerned, the next section will offer some remarks on grammatical gender, both for English and Spanish separately, and for grammatical gender in the context of English-Spanish Code Switching.

2.2. Some remarks on grammatical gender

Grammatical gender is an abstract feature which serves to the purpose of noun classification (Corbett, 1991), and is, in most cases, "not deducible from the meaning of the noun" (Klassen & Liceras, 2017, p. 79). This means that grammatical gender is disassociated from semantics and morphology (Fernández Fuertes et al, 2016), as seen in (7) versus (8).

	SEMANTICS	SYNTAX	MORPHOLOGY
(7) a. niño [boy]	male	masculine	niñ-o
b. niña [girl]	female	feminine	niñ-a
(8) a. sol [sun]	X	masculine	sol-Ø
b. síntesis [synthesis]	X	feminine	síntes-i-s

                                         (Fernandez Fuertes et al, 2016, p. 243)

The examples in (7) show a correlation between semantics (biological gender), grammatical gender, which is the syntactic dimension of gender (masculine or feminine), and the ending of the word, which is the morphological dimension of gender (-o/-a). However, the examples in (8) do not follow this pattern. Neither of them possesses biological gender, so the correlation between semantics and grammatical gender cannot

be imposed, and the endings for both words are not transparent in terms of what grammatical gender they belong to.

Adjectives and articles are also marked for gender by virtue of gender agreement with the noun they accompany, as in (9).

(9) La$_{fem}$ mesa$_{fem}$ blanca$_{fem}$.
[The white table.]

When talking about gender agreement, we can make a distinction between two types. The first type occurs inside the DP, as in (10), and it is what is called concord, while the second type involves a mediation by the verb, as in (11), and is called agreement (Valenzuela et al., 2012, p. 483). This study is focused on the latter (i.e. on agreement in copula constructions).

(10) La$_{fem}$ ciudad$_{fem}$ ruidosa$_{fem}$.
[The noisy city.]

(11) La$_{fem}$ ciudad$_{fem}$ es ruidosa$_{fem}$.
[The city is noisy.]   (Valenzuela et al., 2012, p. 484)

The examples in (10) and (11) are both in Spanish due to the fact that not all languages have grammatical gender. This will be explained in the section below.

2.2.1. Grammatical gender in Spanish vs. English

In Spanish, all nouns have inherent grammatical gender, which can either be masculine or feminine, and which is a lexical feature, as explained above. The assignment of either masculine or feminine gender to nouns is mainly arbitrary. However, for those nouns whose referents are animate entities, the specific grammatical gender coincides with the biological, or socially constructed, gender. Like so, nouns like the one in (12) have an arbitrarily assigned gender (in this case, feminine), which is inherent to the noun, and need to be memorized together with the noun itself. On the other hand, grammatical gender assignment for words like the ones in (13) depend on the biological gender of the referent. This way of assigning gender is not necessarily the case in other languages that, like Spanish, have grammatical gender. For example, in German "das Mädchen", ("the girl") has neuter gender, despite having an animate referent with a feminine biological

7

gender. That is why, as explained above, grammatical gender needs to be separated from semantics (biological gender) and from morphology (the -o, -a endings in Spanish which find no correlate in other languages that also have grammatical gender).

   (12) Mes**a**

        [table]

   (13) Herman**o** [brother]   /   Herman**a** [sister]

However, despite of the disassociation between grammatical gender and morphology, several studies have found that a correlation exists between these different linguistic areas (e.g. Harris, 1991; Roca, 1989; Fernández Fuertes et al., 2016). In fact, gender marking in Spanish is, in most cases, morphologically regular, with the majority of masculine nouns ending in -o and the majority of feminine nouns ending in -a. The examples in (12) and (13) follow this pattern. However, not all nouns end in the canonical morphological forms for masculine and feminine. Some masculine nouns may end in a vowel other than -o or in a consonant, and the same is true for feminine nouns, as seen in (14) and (15).

   (14) El coch**e**$_{masc}$

        [The car]

   (15) La nue**z**$_{fem}$

        [The nut]

Harris (1991) argues that the masculine feature is unmarked in Spanish, and the marked feature is the feminine. This means that the masculine acts as the default value out of the two options. This choice is not shared by all languages that have gender, for example, in Asháninka the unmarked feature is the feminine, as pointed out in Sánchez and Mayer (2018).

English, on the contrary, does not possess grammatical gender as a formal feature. However, it does have a semantic gender system, where some words with animate referents are masculine or feminine (e.g. boy/girl). Gender is also distinguishable in third person pronouns (i.e. he/she) and it is also tied to biological gender. However, there is no productive system of gender marking in English, and even though some words are marked for gender, the grammatical gender feature as such does not exist in English as opposed to Spanish (Mills, 1986; Corbett, 1991; Namai, 2000).

2.2.2. Gender in Spanish-English Code Switching: agreement strategies

There are two possible outcomes in agreement constructions that contain Code Switching: Spanish subject and English subject complement, as in (16a), or English subject and Spanish subject complement, as in (16b), with the copula verb always being in the same language as the subject complement.

(16) a. El libro *is old*.

       b. The book *es viejo*.

      [The book is old]

Out of these two outcomes, gender agreement can only be forced upon the example in (16b), by making the English subject agree in gender with the Spanish adjective. This is done when bilingual speakers assign the gender feature of the Spanish translation equivalent to the English subject, as in (17):
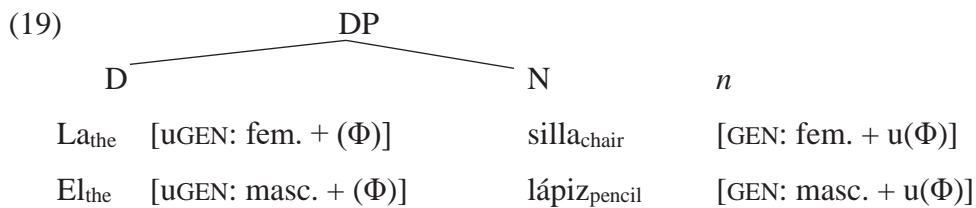
(17) The book [el libro$_{masc}$] es viejo$_{masc}$.

If no such agreement is enforced, then the examples in (18) will be also judged as well formed or possibly produced by bilingual speakers.

(18) a. The house [la casa$_{fem}$] es viejo$_{masc}$.

      [The house is old]

      b. The book [el libro$_{masc}$] is vieja$_{fem}$.

In the examples above, (18a) illustrates an instance of default gender agreement (following Harris 1991, as described in section 2.2.1.), in which the feminine noun agrees with the default masculine form of the adjective; while (18b) is an instance of lack of agreement with *book* inheriting the Spanish masculine gender feature and *vieja* having feminine gender features.

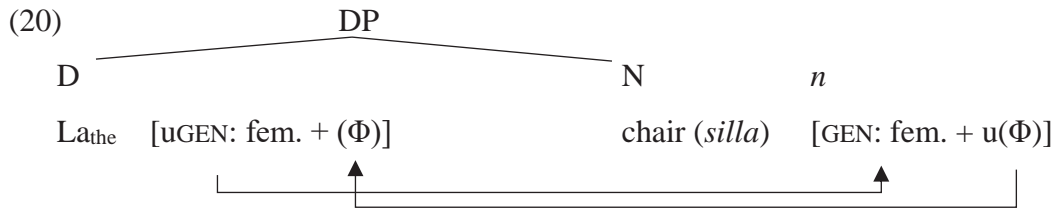Therefore, when dealing with the second outcome in (16), bilinguals can choose between two different agreement strategies: agreement between the Spanish translation equivalent of the English subject and the Spanish adjective, as in (17), and default agreement between the English subject and the Spanish masculine adjective, as in (18a). Below, these two strategies will be explained in more detail.

9

The first strategy, agreement by means of the Spanish translation equivalent, is the one where the speaker produces an adjective that agrees in gender with the translation equivalent of the English noun. This has also come to be known as the 'analogical criterion' (Otheguy and Lapidus, 2005). Liceras et al. (2008) take MacSwan's minimalist approach and propose an account of the analogical criterion which they call the *double-feature valuation mechanism*. This account states that in the process of gender agreement two features are involved and so the features of "GENDER (GEN) and GENDER AGREEMENT (Φ) have to be valued and deleted" (Liceras et al., 2008, p. 835). There is a relationship between GEN and Φ, where GEN is a Noun (N) feature on the Determiner (D) and Φ is a D feature on N. The uninterpretable GEN feature in D has to be valued and deleted by being matched with the interpretable GEN feature in N. Likewise, the uninterpretable Φ feature in N has to be matched with the interpretable Φ feature in D in order to be valued. This relationship is illustrated in (19) for unilingual (non-switched) structures.

(19)
$$DP$$

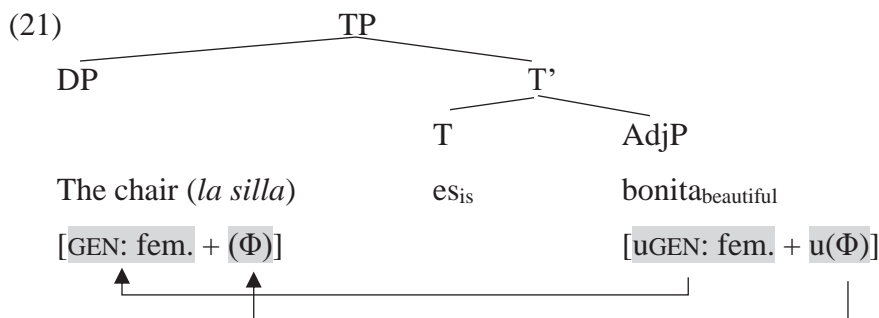| D		N	*n*
La$_{the}$	[uGEN: fem. + (Φ)]	silla$_{chair}$	[GEN: fem. + u(Φ)]
El$_{the}$	[uGEN: masc. + (Φ)]	lápiz$_{pencil}$	[GEN: masc. + u(Φ)]

(Liceras et al., 2008, p. 836)

Because of the lack of grammatical gender in English, this valuation mechanism does not take place in English sentences. However, in switched DPs, the presence of a Spanish D with a Φ feature makes the valuation of GEN necessary. In order to do this, the Spanish N could impose its Φ feature on the English N. As can be seen in (20), the uninterpretable GEN feature in the Spanish D *la* is valued by the interpretable GEN feature in *chair*, which it has inherited from the Spanish translation equivalent *silla*. The uninterpretable Φ feature inherited by *chair* is also valued by the interpretable Φ feature in *la*. The valuation mechanism does not crash because both of them have the feminine value, meaning that the switch follows the analogical criterion.

(20)

```
                           DP
          _____|_____
         D                                   N          n
      La_the   [uGEN: fem. + (Φ)]       chair (silla)   [GEN: fem. + u(Φ)]
```
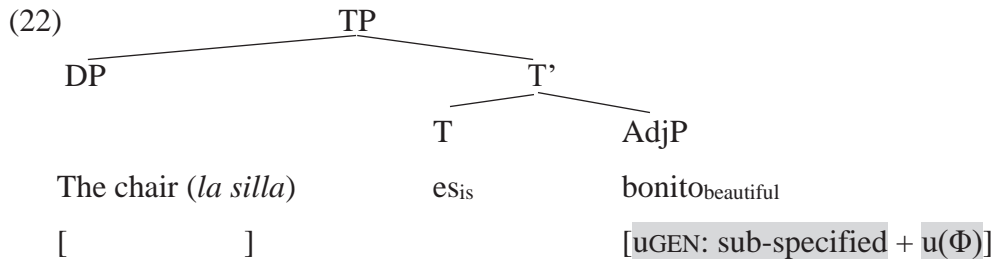
(Liceras et al., 2008, p. 837)

The *double-feature valuation mechanism* can also explain the analogical criterion in switches inside agreement structures, which are the focus of this study. Example (21) illustrates how the Spanish adjective *bonita* carries two uninterpretable features. The subject DP *the chair* has inherited valued GEN and Φ features from it Spanish equivalent, and these features are valued with the uninterpretable ones in *bonita*. In these structures the valuation process is unidirectional, unlike in the process in DP structures, as illustrated by the directionality of the arrows.

(21)

```
                           TP
          _____|_____
       DP                                    T'
                                    _____|_____
                                   T                 AdjP
       The chair (la silla)       es_is           bonita_beautiful
       [GEN: fem. + (Φ)]                         [uGEN: fem. + u(Φ)]
```

The other agreement strategy bilinguals might use in Code Switching is the masculine gender as a default option. As seen before, masculine is the default gender in Spanish, which might lead bilingual speakers and learners to select it, disregarding the gender of the Spanish translation equivalent. The example in (22) shows that in agreement structures where the masculine as a default option is realized, the Spanish adjective does not need to share its features with those of the noun and so the adjective will always carry masculine default features.

(22)

```
                        TP
          ┌─────────────┴──────────┐
        DP                         T'
                              ┌─────┴──────┐
                              T           AdjP

  The chair (la silla)       es_is       bonito_beautiful

  [            ]                         [uGEN: sub-specified + u(Φ)]
```

In this case, there is no clash of features even if the Subject DP contains a [+fem] feature (i.e. when the Spanish translation equivalent of the English subject is feminine) because the GEN feature in the adjective is sub-specified.

3. Hypotheses

Following the theory reviewed in the section above on Code Switching and gender agreement, the hypotheses of this study will be described in detail during the course of this section.

The focus of this study is the type of gender agreement operation that is established in code switched agreement structures where English provides the subject and Spanish provides the adjectival subject complement. Bearing in mind the notions presented in section 2.2 on grammatical gender, the hypotheses will mainly deal with two issues: the analogical criterion and the masculine as default agreement. It is argued that these two issues will have a different impact in the various speaker groups, depending on their dominant language or their L1. Therefore, the four hypotheses below capture the predictions for each of the four groups of participants (which will be described in section 4.1. below). They are ordered in terms of the degree of Spanish dominance of the participant groups.

First, agreement following the analogical criterion, as in (23), will be favored by L1 Spanish speakers (group 1), due to the fact that they will be influenced by the grammatical gender properties of their L1. They will impose the grammatical gender features of the Spanish translation equivalent upon the English subject.

(23) The **city** [Spanish fem.] *es ruidosa* [Spanish fem.].

[The city is noisy.]

Secondly, simultaneous bilinguals who were born and raised in Spain, and whose dominant language is therefore Spanish (group 2), will also prefer the analogical criterion

because they will impose the gender features of Spanish upon the English subject. However, they might do so to a lesser extent than the L1 Spanish speakers, since exposure to English from a young age might reduce the weight these gender features have when judging the acceptability of structures.

Thirdly, simultaneous bilinguals who were born and raised in either the U.K. or the U.S.A. (group 3), and whose dominant language is English, will prefer the masculine as default option, as in (24). The fact that their L1, English, does not possess grammatical gender might influence these speakers, which will cause them to favor the non-marked option in Spanish, and they will not impose the gender features of the Spanish translation equivalent upon the English subject. The reason for this is that the English subject does not carry gender features, as English does not have grammatical gender. However, their other L1, Spanish, might influence their perception, so these speakers will prefer this option to a lesser extent than L1 English late bilinguals.

(24) The **city** [Spanish fem.] *es **ruidoso*** [Spanish masc.].
[The city is noisy.]

In the fourth and final hypothesis, it is argued that L1 English late bilinguals (group 4) will also show a preference for agreement using the masculine as default option, due to the influence of their L1. As stated above, these speakers will show a stronger preference for the masculine as default option than the simultaneous bilinguals born and raised in English speaking environments.

The four hypotheses proposed above for the four groups show a hierarchy, depending on how much weight Spanish features have in the mind of these bilinguals. Therefore, the stronger their Spanish, the higher the preference for the analogical criterion; and the weaker their Spanish, the higher the preference for the masculine default. This hierarchy is illustrated in Table 1.

Table 1. Hierarchy of the participant groups

1	→	2	→	3	→	4
L1 Spanish	→	L1 Spanish-Heritage English	→	L1 English-Heritage Spanish	→	L2 Spanish
Analogical Criterion	→	Analogical Criterion	→	Masculine default	→	Masculine Default

The next section will offer a more thorough description of the four participant groups, as well as an explanation of the task used to collect the data analyzed in this dissertation.

4. Methodology

This section will describe the methodology followed during the development of this dissertation. The information below will be focused on the participants, the experimental task, and the analysis of the experimental data which are at the bases of this work

4.1. Participants: the four experimental groups

The participants are divided into four groups according to their country of origin and to their L1. Each group is formed by 6 adult participants, making a total of 24 participants included in this study. All of them have been tested by the UVALAL (University of Valladolid Language Acquisition Lab). During the 2017-2018 academic year, I was granted a scholarship by the Ministry of Education to collaborate with the UVALAL, which is why I have been able to participate in the data collection of two of the groups: the L1 Spanish group (group 1) and the English heritage group (group 2). Also, when it comes to data codification, I have participated in the data classification of the participants from the four groups that are included in this dissertation.

The first group I will discus is formed by L1 Spanish-L2 English speakers. The 6 participants inside this group are late bilinguals, that is, they have grown up in an L1 Spanish household and have not acquired their second language (English) until later in life, once they learnt it in school. Despite living in a Spanish dominant setting, at the time of testing these participants used English on a regular basis at university. The participants in this study took an English level test and scored at least a C1 level, and their ages range from 20 to 24 years old.

The second group is the English heritage speakers in Spain. This group consists of 6 early bilinguals that have acquired both English and Spanish from birth and inside a natural context. These participants were raised in Spain and have, therefore, grown up in a setting where Spanish is the dominant language in the society. The situation at home is different in that at least one parent is an L1 English speaker, so they use English in the family setting. Consequently, even if both English are Spanish are their two L1s, Spanish

14

is their dominant language and English is their heritage language. The ages in this group range from 19 to 25 years old.

The third group also involves early bilinguals and, in particular, it comprises a group of 6 Spanish heritage speakers in Florida (U.S.A.), whose ages are between 20 and 22 years old. They were all born and raised in the U.S.A., but have acquired Spanish from birth and in a natural context because both their parents are L1 Spanish speakers. These participants use Spanish in their family and social settings, but have grown up in an English dominant environment, making English their dominant language and Spanish their heritage language.

The fourth and final group comprises late bilinguals, and consists of L1 English-L2 Spanish speakers. These 6 participants have been born and raised in an English-speaking country (three of them in the U.K. and three of them in the U.S.A.), and English is therefore their L1. They learnt their second language, Spanish, later in life, in an institutional setting. Their levels of Spanish vary, but they use Spanish frequently as part of their studies or their job and they were able to complete the task successfully. The ages of these participants are between 22 and 27 years old.

4.2. Data elicitation task: acceptability judgments

Participants were tested as part of a coordinated project carried out by the UVALAL from the University of Valladolid and the LARLAB (Language Acquisition Research Laboratory) from the University of Ottawa. They had to complete three different experimental tasks, as well as a background questionnaire, a self-assessment questionnaire and a consent form, all designed by Liceras and Fernández Fuertes. The tasks were two Acceptability Judgment Tasks and one Production Task, designed to study English-Spanish Code Switching as it appears in both concord and agreement structures.

For the present study, which is focused on the perception of gender in agreement structures, only the Acceptability Judgment Task for agreement structures will be considered.

The Acceptability Judgment Task used for this study is focused on agreement structures, and its purpose is to test how bilinguals perceive English Spanish Code Switching in copula agreement constructions where the switch occurs between the subject, on the one hand, and the subject complement, on the other hand. More specifically, the aim of the present study is to focus on the type of grammatical gender

agreement operation between an English subject and a Spanish subject complement, as will be addressed below.

Participants were presented with a short dialogue, consisting of a question and an answer, like the one in (25), and then were asked to judge each answer by choosing one of four emoticons. These emoticons represent a scale from 1 to 4, 1 meaning "sounds very bad" and 4 meaning "sounds very good".

(25) Q: ¿Te gusta el paquete?

[Do you like the package?]

A: El paquete *is beautiful*.

[The package is beautiful.]

Before performing this task, participants had already completed a series of practice items which allowed them to get familiar with how the task works. The task consists of 58 items in total, and out of these the target structures for the present study are 12. The remaining 46 are, therefore, distractors and fillers: 10 sentences were distractors, as in (26), which include structures unrelated to Code Switching, such as compounds, half of which are correct and half are incorrect; and the remaining 36 are fillers which include Code Switching structures where the switch occurs at a different grammatical point from the one on which this study is focused: 24 items are focused on person agreement between subject and verb, as in (27), and 12 are focused on gender agreement between a Spanish subject and an English adjectival subject complement, as in (28). The 12 experimental sentences that are the target of this study are focused on gender agreement between an English subject and a Spanish subject complement, as in (29).

(26) Es un uñacorta.

[It is a nail-cutter.]

(27) The boy bebe agua.

[The boy drinks water.]

(28) La taza *is white*.

[The cup is white.]

(29) The milk *es delicioso*.

[The milk is delicious.]

Table 2 offers an overview of the organization of the different experimental items in this task that will be relevant for the purpose of this study.

Table 2. *Acceptability Judgment Task items*

Experimental items				Total
MM	MF	FF	FM	12
3	3	3	3	

The 12 experimental items are divided into 4 groups according to the grammatical gender combinations of the translation equivalent of the English subject and of the Spanish subject complement. Considering this, there are two possibilities with two different instantiations each: to follow the analogical criterion or not to follow it. In the first case, agreement follows the analogical criterion by having the gender of the Spanish subject complement coincide with that of the translation equivalent of the English subject, as in (30) for masculine (MM) and (31) for feminine (FF). In the second case agreement does not follow the analogical criterion, which results in having a masculine subject and a feminine subject complement (MF), as in (32), or a feminine subject and a masculine subject complement (FM), as in (33).

(30) The **tree** [Spanish masc.] *es alto* [Spanish masc.]. (MM)
   [The tree is tall.]
(31) The **flower** [Spanish fem.] *es hermosa* [Spanish fem.]. (FF)
   [The flower is beautiful.]
(32) The **castle** [Spanish masc.] *es preciosa* [Spanish fem.]. (MF)
   [The castle is gorgeous.]
(33) The **milk** [Spanish fem.] *es delicioso* [Spanish masc.]. (FM)
   [The milk is delicious.]

The 4 groups of participants that have been described in section 4.1. above have been tested using the Acceptability Judgment Task just presented. Taking into account that 24 participants were tested in total, and that each task contained 12 target structures, the total corpus of sentences analyzed comprises 288 judgments. These will be analyzed in the following section.

5. Results

This section will offer a description of the results obtained after analyzing the data gathered from the Acceptability Judgment Task previously described in section 4.2. In the process of analyzing these data, the items marked as 1 or 2 (out of 4) are considered to have been judged as incorrect by the participants, while those marked as 3 or 4 are considered to have been judged as correct.

The results will be organized in four groups, coinciding with the four groups of participants and the four hypotheses to be tested in this study.

5.1. L1 Spanish-L2 English speakers

It was predicted that the L1 Spanish-L2 English group (group 1) would show a preference for the sentences that follow the analogical criterion in the Acceptability Judgment Task (section 3). This type of sentence is shown in example (23) and repeated below in example (34).

(34) The **city** [Spanish fem.] *es ruidosa* [Spanish fem.].

[The city is noisy.]

Figure 1 illustrates the number of items judged as correct and as incorrect out of the 18 total items in each category.
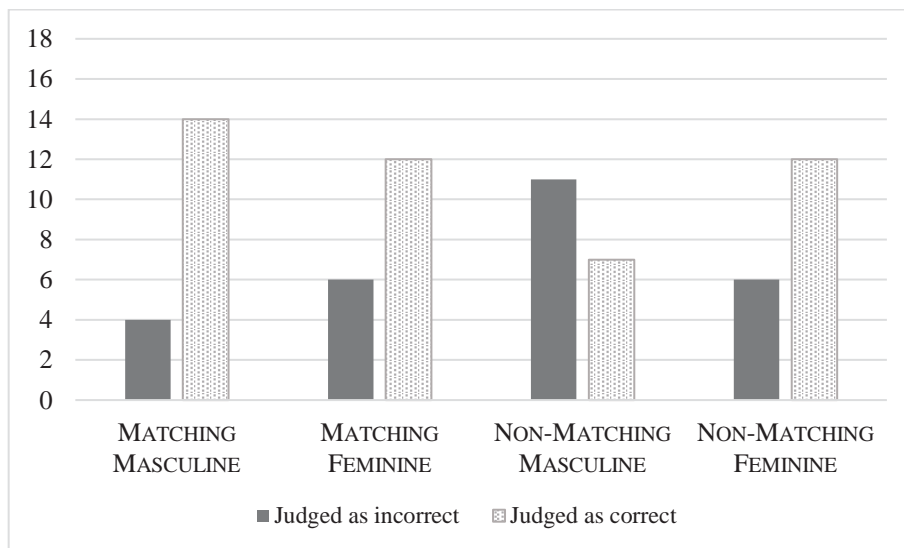

Figure 1. *L1 Spanish-L2 English speakers' judgments*

In both categories of items that abide by the analogical criterion, i.e. Matching Masculine (Masculine-Masculine = MM) and Matching Feminine (Feminine-Feminine = FF), the number of sentences that were judged as correct is greater than the number of answers marked as incorrect. Within the MM category, 14 out of 18 sentences are considered correct, and in the FF category, 12 out of 18.

As for those sentences that did not follow the analogical criterion agreement strategy, i.e. Non-Matching Masculine (Feminine-Masculine = FM) and Non-Matching Feminine (Masculine-Feminine = MF), the results vary. Within the FM category, most items were judged as incorrect by this group of speakers (11 out of 18 items). However, within the MF category most items were judged as correct (12 out of 18).

5.2. English heritage speakers in Spain

The English heritage speakers group (group 2) was predicted to have similar results to the L1 Spanish speakers group, since they were both assumed to have a greater influence from Spanish in their perception of gender agreement. Therefore, the prediction is that this group will favor those sentences in which the analogical criterion is used as an agreement strategy, although to a lesser extent than the L1 Spanish speakers because of the heritage speakers' experience with and exposure to English. The different judgments made by this group of speakers are pictured below in Figure 2.
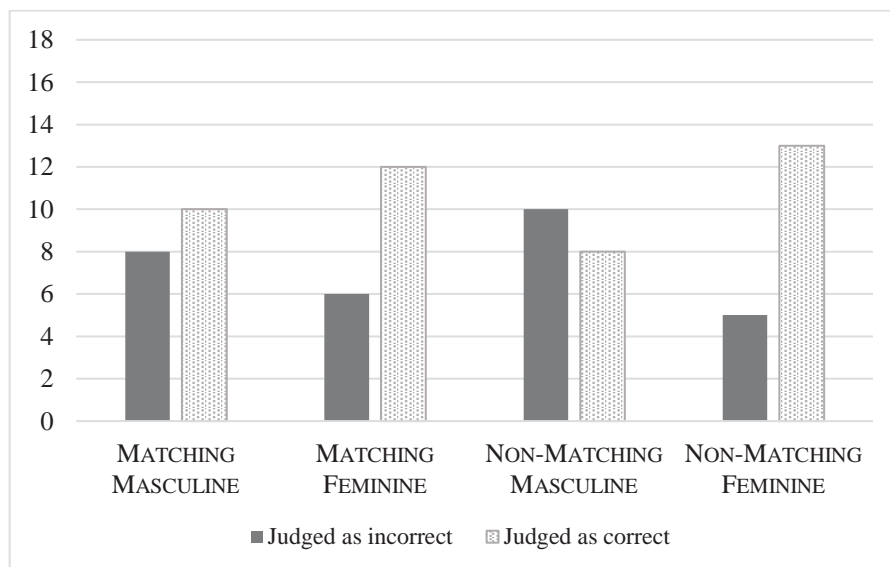
Figure 2. English heritage speakers' judgments

Within the sentences that follow the analogical criterion strategy, most participants marked MM and FF items as correct. However, the difference between sentences marked as correct and sentences marked as incorrect is more noticeable in the FF category (12 out of 18 were marked as correct) than in the MM category (10 out of 18). The FM sentences were generally considered incorrect, with only 8 out of 18 sentences marked as correct. Finally, within the MF category most items were judged as correct (13 out of 18).

5.3. Spanish heritage speakers in Florida

For the Spanish heritage group (group 3), it was predicted that they show a preference for the masculine as default strategy, since their dominant language is English. An example of this option is shown in (24) and repeated below in (35).

(35) The **city** [Spanish fem.] *es **ruidoso*** [Spanish masc].
       [The city is noisy.]

However, because of the influence of Spanish in gender perception, these participants might favor the masculine default to a lesser extent than L1 English speakers. Figure 4 offers an illustration of the different judgments made by these speakers.
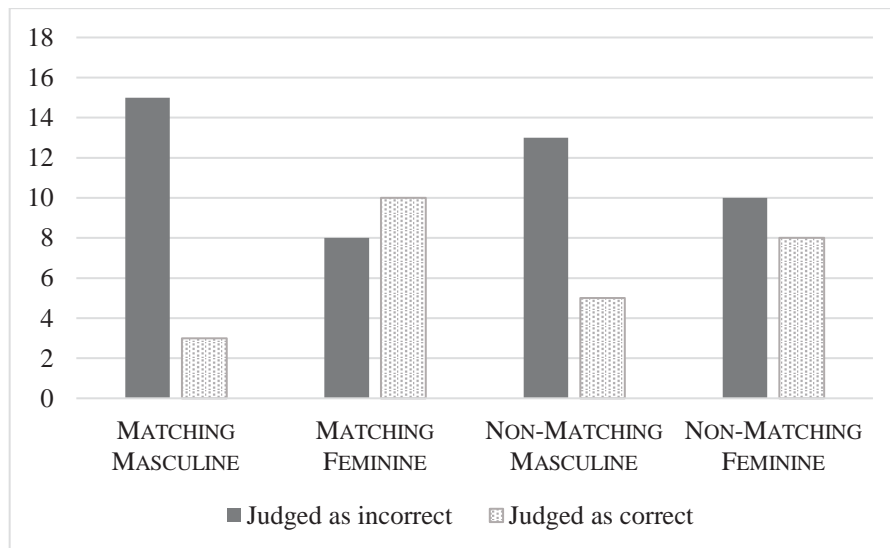

Figure 3. Spanish heritage speakers' judgments

The items that belong to the MM category are largely considered incorrect, with only 3 out of 18 sentences marked as correct. Most FF sentences were marked as correct, but

by a small margin (10 out of 18). The FM sentences (those that follow the masculine as default strategy) are mostly considered incorrect, having only 5 out of 18 items judged as correct. Finally, the MF sentences are also mostly judged as incorrect, although to a lesser extent than the FM sentences (8 out of 18 sentences were considered correct).

5.4. L1 English-L2 Spanish speakers

The prediction made for the L1 English-L2 Spanish group (group 4) was that they would show a preference for those sentences that followed the masculine as a default option as agreement strategy due to the influence of their L1. Illustrated in Figure 3 are the number of items marked as correct and incorrect by this group of speakers.
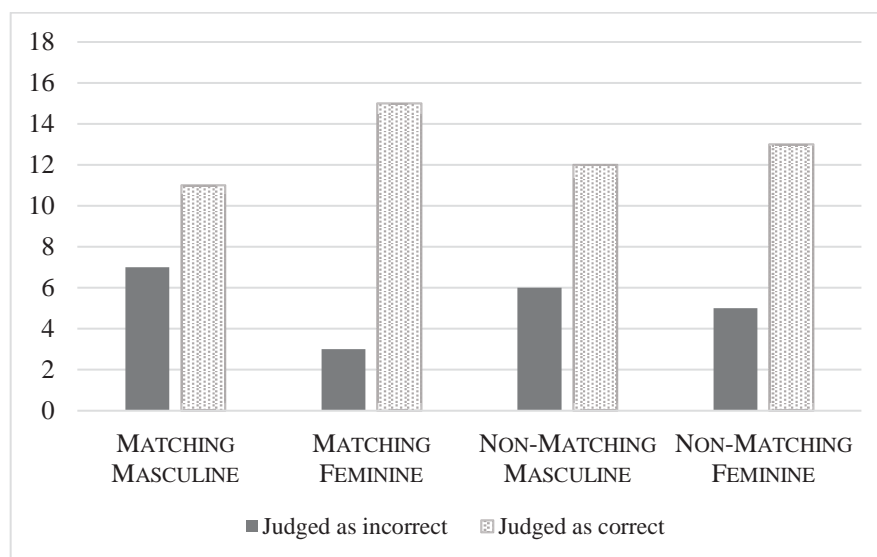
Figure 4. L1 English-L2 Spanish speakers' judgments

Items that belong to both the MM and the FF categories are mostly considered correct by these participants, although the FF sentences are considered so in a much higher extent (15 out of 18 items marked as correct vs. 11 out of 18 in the MM category). Items within the FM category, which are the ones following the masculine as default strategy, are also mostly considered correct (12 out of 18 items). As for the MF category, most items are also marked as correct (13 out of 18).

6. Discussion

Following the analysis of the data and the results obtained from it, the hypotheses presented in section 3 are addressed below in order to determine whether they are confirmed or rejected.

First, it was predicted that the L1 Spanish speakers (group 1) would favor agreement following the analogical criterion, due to the influence of their L1, and would, therefore, show a preference for structures like those in (36) and (37), as opposed to those in (38) and (39). The results illustrated in Figure 1 show that this hypothesis is confirmed, and that structures that follow the analogical criterion (MM in 36 and FF in 37) are generally judged as correct by this group, as opposed to FM structures, like the one in (39), which are generally judged as incorrect. However, the results also show that in general, speakers have also judged as correct MF structures, as in (38), that do not follow the analogical criterion, something that was not predicted by the hypothesis.

(36)   The **tree** [Spanish masc.] *es alto* [Spanish masc.].   (MM)
         [The tree is tall.]

(37)   The **flower** [Spanish fem.] *es hermosa* [Spanish fem.].   (FF)
         [The flower is beautiful.]

(38)   The **castle** [Spanish masc.] *es preciosa* [Spanish fem.].   (MF)
         [The castle is gorgeous.]

(39)   The **milk** [Spanish fem.] *es delicioso* [Spanish masc.].   (FM)
         [The milk is delicious.]

This unexpected high rate of correct judgments for the MF structures could be due to the weight the Spanish grammatical gender feature has for these speakers. The marked gender feature in Spanish is the feminine, while the masculine is the unmarked or default feature. Perhaps the importance of gender marking in the mind of these speakers is what led them to rate these items as correct, despite the fact that the adjective would not agree with the Spanish translation equivalent of the English subject.

Secondly, it was also predicted that English heritage speakers (group 2) would favor structures that follow the analogical criterion agreement strategy, since their dominant language is Spanish and the features from this language will be stronger. However, they were predicted to favor them to a lesser extent than L1 Spanish speakers, because their heritage language will have a higher influence in this group. The results displayed in Figure 2 show that, in fact, English heritage speakers do judge most sentences that follow the analogical criterion (36 and 37) as correct, and indeed they do so to a lesser extent than L1 Spanish speakers, especially in the MM category. Sentences that belong to the FM category, as in (39), are also usually judged as incorrect, but less so than what the L1

22

Spanish speakers did and MF sentences like the one in (38) are widely judged as correct against prediction. Once again, these sentences being judged as correct may be due to the weight of the grammatical property of gender and to the importance of gender marking in Spanish, their heritage language.

The third hypothesis predicted that Spanish heritage speakers (group 3) would show a preference for the default masculine option as agreement strategy, but to a lesser extent than the L1 English group because of the greater influence their heritage language might have on them. The results summarized in Figure 3 completely reject this hypothesis. Both categories of items that use masculine agreement (MM and FM) are considered highly incorrect, while the only category that shows a slightly higher rate of items judged as correct is the FF one. Even if this hypothesis is rejected, this group still does not behave like the L1 English group, since they do show some sensitivity to Spanish grammatical gender. In fact, there is a higher rate of sentences judged as incorrect in those sentences where the adjectival subject complement is masculine (MM and FM). This would again show that the influence Spanish has on speakers causes them to show a preference for feminine agreement due to the importance of gender marking and considering that feminine is, in fact, the marked option for gender.

The fourth and final hypothesis predicted that the L1 English speakers (group 4) would show a preference for those structures that followed the agreement strategy of masculine as a default option, because the lack of gender in their L1 would lead them to use the default gender. The results illustrated in Figure 4 show that, while most structures that follow the masculine as a default option have indeed been judged as correct by this group of speakers, all other structures have also been judged as correct. Therefore, no preference for the masculine as default option is showcased, and this hypothesis is rejected. Considering that this group does not show any preference for any specific structure, it seems that gender is not an issue for them; that is, they have paid no actual attention to gender. This is to be expected, since their L1 (English) lacks grammatical gender and, therefore, they are not sensitive to it.

Figure 5 below offers a summary of the percentages of expected responses for each hypothesis (which coincides with each group of participants). The expected responses for the L1 Spanish group and the English heritage group were that they would mark MM and FF items as correct and FM and MF items as incorrect, while those expected for the L1 English group and the Spanish heritage group were that they would mark MM and FM sentences as correct and FF and MF sentences as incorrect. This figure shows how the

third hypothesis (concerning L1 English speakers) and the fourth hypothesis (concerning Spanish heritage speakers) were rejected, since the percentages of expected responses are quite low. It also shows how, while hypotheses one (concerning L1 Spanish speakers) and two (concerning English heritage speakers) are mostly confirmed, the rate of expected responses for the MF category is still very low.
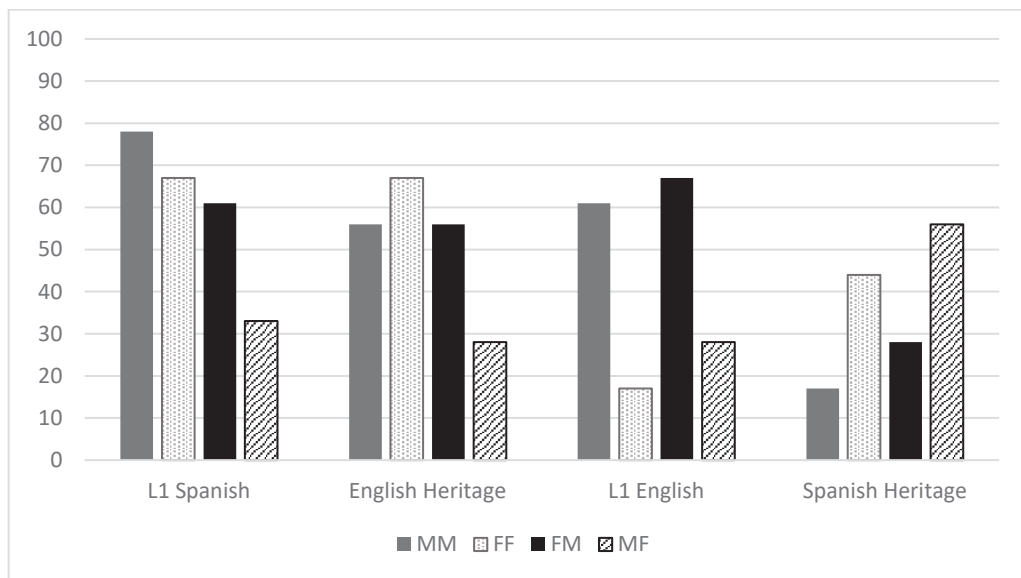
Figure 5. Percentages of expected responses

The results for both the L1 Spanish and the English Heritage groups show that they do in general show a higher preference for matching items. These results are in line with Liceras et al.'s (2008) findings that L1 Spanish speakers prefer matching items over non-matching items. However, the unexpected acceptability for non-matching feminine items does not fall in line with either the previous study nor the hypotheses proposed in this dissertation.

The results for L1 English and Spanish heritage speakers are also unexpected. Even though the answers of the L1 English group for the MM and FM categories coincided with the expected answers, the fact that all of the categories were generally judged as correct rejects the possibility of it being due to a preference for the masculine as a default option. However, the findings might suggest that L1 English speakers do not have any sensitivity to gender, which would explain why they do not seem to prefer one strategy over the other. As for Spanish heritage speakers, even though the hypothesis was rejected, they did seem to have influence from their heritage language as seen in the importance they attribute to gender marking.

The high rate of unexpected answers may also be due to the fact that the size of this study was small (24 participants). Future studies will be needed in order to corroborate the findings from this dissertation, more specifically, those related to the importance that gender marking has in the minds of bilingual speakers from different linguistic profiles.

7. Conclusion

The present study has dealt with a phenomenon that consists in bilingual speakers moving in and out of the languages they speak: Code Switching. In particular, this study has been concerned with Code Switching within a single sentence, i.e. intrasentential Code Switching, and the focus has been on agreement structures where the subject was in English and the verb and the adjectival subject complement were in Spanish. Bearing this in mind, the aim of this study was to examine what type of agreement strategy would different groups of bilingual speakers prefer, considering how one of their languages (English) has no grammatical gender properties. In order to do this, an Acceptability Judgment Task was used in order to test how different gender agreement strategies are perceived by speakers. This study compared the results from four different groups of participants (L1 Spanish speakers, English heritage speakers in Spain, Spanish heritage speakers in the U.S.A. and L1 English speakers), to establish whether or not their L1 or their dominant language played a role in their preference for one strategy over another.

The analysis of the data collected from the Acceptability Judgment Task provided mixed results concerning the hypotheses proposed. First, it was proposed that L1 Spanish speakers would show a preference for structures that followed the analogical criterion due to the influence that the gender properties of Spanish might have in these speakers' minds. This was partially confirmed, since they showed a high preference for structures following the analogical criterion. However, they also showed a high preference for structures where the Spanish translation equivalent of the English subject was masculine but the adjectival subject complement was feminine (MF). It has been suggested that this might be due to the importance of gender marking in Spanish, which might lead them to choose feminine over masculine even if the adjective does not agree with the Spanish translation equivalent of the English subject. The second hypothesis, concerning English heritage speakers in Spain, predicted similar results, although the preference would be so to a lesser extent than that for L1 Spanish speakers due to the influence of their heritage language: English. The results confirm that they show a high acceptance for structures following the analogical criterion (less so than Spanish speakers, as predicted), but, again,

there is an unexpectedly high acceptance for MF structures. Once more, it has been proposed that the influence of their dominant language (Spanish) made them prefer structures marked for gender even if the gender of the Spanish translation equivalent of the English subject does not match that of the Spanish subject complement. Thirdly, it was predicted that Spanish heritage speakers in the U.S.A. would show a slight preference for the masculine as default option, but it was rejected as well because speakers marked most items as incorrect, with the exception of matching feminine structures. This might indicate that they are influenced by their heritage language in their preference for structures marked for gender. Finally, the fourth hypothesis predicted that L1 English speakers would show a preference for the masculine as default strategy, since masculine is the unmarked gender in Spanish. This hypothesis was rejected, because speakers showed a high acceptance of all types of structures. This might be due to the fact that English does not have a grammatical gender system, which makes these speakers not sensitive to gender, in general, and to gender agreement, in particular.

These results show that the grammatical gender properties of both English and Spanish do influence speakers in their perception of gender in agreement structures. That is, the specific grammatical properties of the two languages (in this case, gender) have an impact on how speakers process their two languages in contact. However, further research will be needed with a larger participant pool to determine whether or not this trend is maintained.

References

Clyne, M. (2003). *Dynamics of Language Contact: English and Immigrant Languages*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511606526

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

Fernández Fuertes, R., Álvarez de la Fuente, E., & Mujcinovic, S. (2016). The acquisition of grammatical gender in L1 bilingual Spanish. In A. Alba de la Fuente, E. Valenzuela, & C. Martínez Sanz (Eds.), *Studies in Bilingualism* (Vol. 51, pp. 237–279). Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/sibil.51.10fue

Harris, J. W. (1991). The Exponence of Gender in Spanish. *Linguistic Inquiry*, *22*(1), 27–62.

Klassen, R., & M. Liceras, J. (2017). The representation of gender in the mind of Spanish-English bilinguals: Insights from code-switched Adjectival Predicates. *Borealis – An International Journal of Hispanic Linguistics*, *6*(1), 77. https://doi.org/10.7557/1.6.1.4100

Liceras, J. M., Fernández Fuertes, R., Perales, S., Pérez-Tattam, R., & Spradlin, K. T. (2008). Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional–lexical mixings. *Lingua*, *118*(6), 827–851. https://doi.org/10.1016/j.lingua.2007.05.006

MacSwan, J. (2000). The architecture of the bilingual language faculty: evidence from intrasentential code switching. *Bilingualism: Language and Cognition*, *3*(1), 37–54.

Macswan, J. (2005). Codeswitching and generative grammar: A critique of the MLF model and some remarks on "modified minimalism." *Bilingualism: Language and Cognition*, *8*(1), 1–22. https://doi.org/10.1017/S1366728904002068

MacSwan, J. (2009). Generative approaches to code-switching. In B. E. Bullock & A. J. Toribio (Eds.), *The Cambridge Handbook of Linguistic Code-switching* (pp. 309–335). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511576331.019

Mills, A. E. (1986). *The Acquisition of Gender* (Vol. 20). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-71362-0

Muysken, P. (2000). *Bilingual Speech: A Typology of Code-Mixing*. Cambridge University Press.

Myers-Scotton, C. (2006). *Multiple Voices: An Introduction to Bilingualism*. UK: Blackwell.

Myers-Scotton, C., & Jake, J. (2009). A universal model of code-switching and bilingual language processing and production. In B. E. Bullock & A. J. Toribio (Eds.), *The Cambridge Handbook of Linguistic Code-switching* (pp. 336–357). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511576331.020

Namai, K. (2000). Gender features in English. *Linguistics*, *38*(4), 771–780.

Otheguy, R., & Lapidus, N. (2005). Matización de la teoría de la simplificación en las lenguas en contacto: El concepto de la adaptación en el español de Nueva York. In R. Otheguy & M. Lacorte (Eds.), *Contactos y contextos lingüísticos: El español en Estados Unidos y en contacto con otras lenguas*. Madrid & Frankfurt: Editorial Iberoamericana / Vervuert Verlag.

Otheguy, R., & Lapidus Shin, N. (2003). An adaptive approach to noun gender in New York contact Spanish. In R. Núñez-Cedeño, L. López, & R. Cameron (Eds.),

Current Issues in Linguistic Theory (Vol. 238, p. 209). Amsterdam: John

Benjamins Publishing Company. https://doi.org/10.1075/cilt.238.17oth

Pfaff, C. W. (1979). Constraints on Language Mixing: Intrasentential Code-Switching

and Borrowing in Spanish/English. *Language*, *55*(2), 291–318.

Poplack, S. (1980). "Sometimes I'll start a sentence in Spanish y termino en español":

toward a typology of code-switching. *Linguistics*, *18*(7), 581–618.

Roca, I. M. (1989). The Organization of Grammatical Gender. *Transactions of the

Philological Society*, *87*, 1–32.

Sánchez, L., & Mayer, E. (2018). Typological differences in morphological patterns,

gender features, and thematic structure in the L2 acquisition of Ashaninka Spanish.

Languages. Retrieved from http://www.mdpi.com/journal/languages

University of Valladolid. (2009). *Verifica del Grado en Estudios Ingleses*.

Valenzuela, E., Faure, A., Ramírez-Trujillo, A. P., Barski, E., Pangtay, Y., & Diez, A.

(2012). Gender and Heritage Spanish Bilingual Grammars: A Study of Code-mixed

DeterminerPhrases and Copula Constructions. *Hispania*, *95*(3), 481–494.

Vanden Wyngaerd, E. (2017). The adjective in Dutch-French codeswitching: Word

order and agreement. *International Journal of Bilingualism*, *21*(4), 454–473.

Appendix: List of acronyms and abbreviations used

Φ: Gender agreement

D: Determiner

DP: Determiner Phrase

EL: Embedded Language

Fem.: Feminine

FF: Feminine agreement following the analogical criterion

FI: Full Interpretation

FM: Masculine agreement not following the analogical criterion

GEN: Gender

L1: First Language

L2: Second Language

Masc.: Masculine

MF: Feminine agreement not following the analogical criterion

ML: Matrix Language

MLF: Matrix Language Frame

MM: Masculine agreement following the analogical criterion

MP: Minimalist Program

N: Noun

PF: Phonological Form