Spanish modal infinitival constructions in English–Spanish bilingual grammars

Vanesa Alonso González

A thesis submitted in partial fulfillment of the requirements for the

Doctorate in Philosophy degree in Spanish

Department of Modern Languages and Literatures

Faculty of Arts

University of Ottawa

Acknowledgements

Writing a doctoral dissertation is a solitary endeavour but it is impossible to do it alone. The following words are dedicated to those that helped and encouraged me to achieve one of the most challenging ventures of my life. For them, I kept on going when my path towards submitting this thesis became invisible, obscure, tangled, or even absent.

First, my thesis supervisor, Dr. Liceras. From the very beginning of my PhD, this outstanding scholar opened my eyes to multiple branches of linguistics and, with kindness and generosity, provided me with countless opportunities to grow within the academic world. From a personal dimension, professor Liceras infused me with the enthusiasm and gave me the necessary support to start, to continue, and to finish my thesis research. I thank you for your precious time, your unconditional dedication, your immensurable expertise, and your relentless curiosity.

Dr. Liceras was with me for the last part of this marathon—the hardest part—but before her, there were two extraordinary women that put me on this road: Professor Karant and Dr. Rodrigo. Professor Karant, the best teacher I have ever had, situated me on the starting line of second language teaching and acquisition; Dr. Rodrigo, who show me the possibilities of creative writing in L2 environments, took me along the route of applied linguistics. I thank you for your kindness, your wisdom, and your guidance.

From this internal academic circle, my gratitude expands to the members of the internal committee: Dr. García Soriano, Dr. García-Alcaraz, and Dr. Mañas Navarrete. Dr. García Soriano gave me her inestimable insight into theoretical linguistics; Dr. García-Alcaraz lent me her keen, organized, and challenging view on methodology and statistics; Dr. Mañas Navarrete provided me

with the tools and knowledge to understand the world of R. I thank you for helping me in tackling these challenges.

I turn now to the external members of my thesis committee: Dr. Roca Urgell and Dr. Valenzuela. I thank you for your knowledge, your feedback, and your dedication. Also, my thanks have to reach the members of the *Language Acquisition Research Laboratory* at the University of Ottawa. I had the luxury of sharing this process with you, from its inception as a Comprehensive Exam to its completion as a full-fledged thesis. I thank you for your questions, your answers, and your patience. From the University of Windsor, special thanks must be given to Professor Alena Auchynnikava. I thank you for your inestimable friendship and timely guidance on statistical matters. Also, behind the scenes, there have been three key figures whose work has made the conclusion of this research possible: Marco, in charge of making these pages legible; Shalina Tripodi, for handling the bureaucratic reins; and the members of the Research Ethics Board of uOttawa. I thank you for your professionalism and your timely responses.

This dissertation has been also possible due to the funding received from the Ontario Graduate Scholarship program (Government of Ontario and the University of Ottawa), the International Doctoral Scholarship, the Admission Scholarship, and the Excellence Scholarship (University of Ottawa), along with travel grants from the University of Ottawa, the University of Illinois Urbana–Champaign, and the Canadian Association of Hispanists. I thank these institutions for supporting academic quests on the Humanities.

Now it is my turn to thank those at the heart of the thesis research, my participants: the 103 participants (though we ended up using the experimental data of only 96 of them) that provided me with the experimental data and the Spanish heritage siblings whose code-switched utterances conformed our naturalistic data. I would also like to acknowledge those who helped me to find

participants (sometimes participants themselves)—one of the most challenging tasks of the whole process. Among them are committed family members, dear friends, esteemed colleagues, favourite neighbors, and generous strangers. Along with these individual helpers, I must refer to those academic institutions that have participated in recruiting participants for my research: the University of Valladolid, Universidad Nacional Autónoma de México in Ottawa, Georgia State University, Universidad Autónoma de Madrid, Easy Español, the British Council School in Madrid and Somosaguas, Jesús y María Secondary School, the Cervantes Institute of Dublin, Manchester and Chicago, the Escuela Oficial de Idiomas de Pozuelo de Alarcón, Spectrum Spanish, the International School of Valladolid, AIL Español, and English4All. I thank you all for providing me with the most important part of my thesis—the data that gave answers to my research questions.

I have mentioned those that have helped me to instrumentally build this thesis, but there is a bigger story to tell and it concerns my closest family and my friends: friends from forever, whom I left in Spain but who have kept me company and given me support in these distant lands through the years; and friends from recent times, who have been with me at different stages of my pilgrimage through these North American lands. I thank you for your incalculable friendship, for your unwavering support, and for your tremendous patience, in good and bad times.

As to my dear family, those furthest only in distance, I could not have reached this point without my mother, Francisca, my father, Antonio, my sister, Raquel, my brother-in-law, Carlos, my nephew, Arturo, and my grandmother, Petri—you are the nature and nurture that have made the person writing this thesis. To those closest also in proximity: my beloved husband and my precious children—you three are my purpose and my strength, my source of joy, and my ultimate reward. I thank you for everything. God bless you.

Table of Contents

Abstract

The aim of this thesis is twofold: first, to determine whether a group of Spanish-native adult speakers and six groups of English–Spanish bilinguals are able to differentiate modal periphrases (MP), as in *debo/puedo* comer (I must/can eat), from nominal infinitival clauses (NIC), as in *prefiero/deseo* comer (I prefer/wish to eat). These two structures are superficially similar (inflected verb + infinitive) and have modal meaning, but differ in their internal structure (Gómez Torrego, 1999; RAE-ASALE, 2011; Vivanco, 2019). The second aim is to determine whether there is a pattern of inclusion/exclusion of the English infinitival marker *to* in intra-sentential English–Spanish code-switching in AUX/MOD–V and V–V sequences which depict MPs, NICs, and the periphrastic future (PF).

To achieve the first goal, we designed a written Grammaticality Judgment Task (GJT) where participants were asked to judge sentences containing three syntactic processes that differentiate MPs from NICs. This GJT was administered to 96 participants: (i) 10 Heritage Spanish speakers, (ii) 10 Heritage English speakers, (iii) 14 L1Spa–L2Eng Immigrant speakers, (iv) 10 L1Eng–L2Spa Immigrant speakers, (v) 22 L1Spa–L2Eng Students, and (vi) 10 L1Eng–L2Spa Students. A Control group (CG) of 20 native Spanish speakers, born, raised, and residing in Spain, was also recruited.

The results of the first task showed positive overall performance (84% accuracy). There were no significant differences between the groups, but *Grammaticality* and the three syntactic processes played a significant role in determining *Accuracy*. Our participants' performance was better with grammatical than with ungrammatical items, except for the Heritage English group, who were more accurate than the rest when judging ungrammatical MPs. As for the three syntactic processes, 'That' sentences yielded a significantly higher level of accurate answers, followed by

'What' and 'Demonstrative' sentences. Despite the significant role played by *Grammaticality* and the syntactic processes, the high rate of accurate responses and the lack of significant differences between the overall performance of an ample spectrum of English–Spanish bilinguals could indicate the acquisition of MPs and NICs, characterized by their lack of saliency, their variety of grammatical constraints, and their semantic complexity.

To investigate the second goal, we designed a written Two-alternative Forced-choice task (2AFC), where participants were asked to choose one sentence within pairs of Spanish–English code-switched infinitival constructions (MP, NIC, and PF) with and without the English infinitival marker *to* (Spanish *a* in PF) between the inflected verb and the infinitive. This task was administered to the 76 participants in the bilingual groups but not to the native Spanish Control group. We also compared the experimental data with the corresponding patterns observed in a longitudinal corpus of two Spanish heritage siblings (4 and 8 years old).

The results of the second task showed a significant pattern of exclusion of the infinitival *to* in MPs and of inclusion of *to/a* in the PF. With respect to NICs, the probability of including *to* was above 0.5, at a significant distance from the other two. There were no significant differences between the groups of bilinguals, and they were more likely to exclude *to* when the inflected verb was in Spanish. These results were partially in accordance with the patterns detected in the naturalistic data: the exclusion of *to* in MPs, the inclusion of *a* in the PF, and the exclusion of *to* in NICs when the inflected verb was in Spanish. There were no code-switched infinitival constructions with an English inflected verb in the corpus. These results contribute new data to discerning potential patterns that govern AUX/MOD–V and V–V code-switched sequences and provide a new perspective on the debate concerning switches between functional and lexical categories.

List of Tables

List of Figures

List of Abbreviations

2AFC	Two-alternative Forced-choice
2L1	Two languages at birth
AIC	Akaike Information Criterion
ANOVA	Analysis of Variance
AO	Age of onset
AUX	Auxiliary
BFLA	Bilingual First Language Acquisition
BLP	Bilingual Language Profile
BTH	Block Transfer Hypothesis
CG	Control Group
CI	Crosslinguistic influence
COMP	Complementizer
CS	Code-switching
DP	Determiner Phrase
EL	Embedded language
Eng	English
FHC	Functional Head Constraint
GC	Government Constraint
GJT	Grammaticality Judgment Task
HS	Heritage speakers
INE	Instituto Nacional de Estadística (National Institute of Statistics)

ImS	Immigrant speakers
L1	First Language
L2	Second Language
L2S	L2 students
LD	Language Dominance
LOR	Length of Residence
MA	Minimalist Approach
MID	Modality-indicating device
ML	Matrix language
MLF	Matrix language frame
MOD	Modal
MP	Modal Periphrasis
NIC	Nominal Infinitival Clauses
NICE	Negation, Inversion, Code, Emphatic affirmation
NP	Noun Phrase
PF	Periphrastic Future
PFR	Principle of Functional Constraint
RT	Reaction time
Spa	Spanish
TAG	Tree Adjoining Grammar
TAM	Temporal, aspectual, modal
V	Verb
VP	Verb Phrase

1. Introduction

Modal verbs are infrequently addressed in Spanish descriptive and pedagogical grammars for second and foreign language teaching because there is no distinct class of items which express modality. While there are many ways to convey modality in Spanish (Rabadán, 2006), the use of the subjunctive mood is the most studied by L1 and L2 researchers and teachers of Spanish as a second and foreign language alike. Another way, however, of expressing modality in Spanish is the use of modal infinitival constructions. We will use the term modal infinitival constructions to refer to those constructions that contain an inflected verb with modal meaning plus an infinitive. Among these infinitival constructions there are two subgroups: Modal Periphrases (MP) with inflected verbs like *poder* and *deber*—the auxiliary verbs that convey most of the meanings expressed by English modal verbs (Carlisle, 1980; Rabadán, 2006)—, and nominal infinitival clauses (NIC), including those headed by verbs such as *preferir* or *desear*.

Even though the two subgroups of modal infinitival constructions show semantic and structural similarities, they fundamentally differ in their internal structure (Gómez Torrego, 1999; Hadlich, 1971; RAE-ASALE, 2011; Vivanco, 2019): periphrases are the unions of two or more verbs that constitute only one predicate head while an NIC is a sequence of two predicates with more than one predicate head (Gómez Torrego, 1999; Topor, 2005). These differences are visible through a set of syntactic tests used to identify periphrases and differentiate them from NICs, such as the impossibility of pronominalizing the infinitive or of adding an embedded clause to the inflected verb (García Fernández, 2006; Gómez Torrego, 1999; Topor, 2005).

The variety of these grammatical constraints (the syntactic processes already mentioned) and their lack of saliency (apparent absence of a superficial structural contrast since both

constructions in Spanish are formed by an inflected verb plus an infinitive) combined with the semantic complexity conveyed by modality as well as the absence of explicit instruction in second and foreign language contexts constitute a set of factors that could contribute to interference in their acquisition by L2 learners. To address this issue, this research will, first, try to investigate whether Spanish native speakers and three different types of English–Spanish bilingual speakers (heritage speakers, L2 students and immigrant speakers) are able to differentiate MPs from the superficially similar NICs.

Second, this thesis will explore modal infinitival constructions in light of one of the main constructs of bilingualism: code-switching between functional and lexical categories, in particular at the juncture of modal auxiliary verbs and lexical infinitives. Since the late 70s and early 80s, when the idea of rule-governed code-switching started to consolidate in the field of bilingualism, there have been different proposals to determine whether and how lexical and functional categories intervene in intra-sentential code-switching. We will test these theories using both experimental and naturalistic data in search of patterns that could clarify the potential code-switching constraints which could affect AUX/MOD–V sequences and V–V sequences, as in (1–6):

(1) The student can (to) *sacar buenas notas*. [MP]

(2) *La ganadora puede* (to) celebrate the victory. [MP]

(3) The model wishes (to) *proponer un cambio*. [NIC]

(4) *El comprador desea* (to) lower the price. [NIC]

(5) We′re going (to) *ir a la fiesta*. [PF]

(6) *Vamos* (a) visit my parents. [PF]

These three infinitival constructions—MPs, NICs, and the periphrastic future (PF)—have Spanish or English directionality (the language of the inflected verb is English or Spanish) and are

formed with or without the English infinitival marker *to* (and the Spanish preposition *a* in the Spanish PF).

Both approaches to the analysis of modal infinitival constructions are presented in two parts. The first part discusses the differences between NICs and MPs, and the second part investigates code-switching patterns within these constructions. Taking this into consideration, the structure of this thesis is as follows: Chapter 1 will address the linguistic dimension of the theoretical framework by introducing the concept of modality as well as the different ways of expressing modality in Spanish and English. This distinction will reveal the semantic similarities between the Spanish verbs *deber* (must/should) and *poder* (can) and the English modal verbs, as well as showing that both languages represent two opposite but complementary ways of expressing modality: Spanish relies on its mood system (subjunctive, imperative and indicative) while English prioritizes modal verbs. Taking this differentiation into account, the chapter will focus on two main means of expressing modality in Spanish, namely, MPs and NICs.

Chapter 2 will first introduce an overview of bilingualism that will constitute the psycholinguistic part of the theoretical framework of this thesis and will then discuss data from a longitudinal study of code-switched infinitival constructions produced by two Spanish heritage children who have been observed in their natural environment by their parents for two years. The first part of Chapter 2 on bilingualism will explore its definition, the types of bilinguals in general and those specific English–Spanish bilinguals that will participate in our experimental study. It will also provide a description of our participants' sociolinguistic context through the exploration of the status of both languages in the participants' places of residence.

After describing the particular characteristics and linguistic context of our participants, we will introduce three of the main constructs of bilingualism: language dominance, crosslinguistic

influence and code-switching. After discussing the broad concept of code-switching, Chapter 2 will address the specific code-switching constraints between lexical and functional categories proposed within the Generative syntax tradition, more specifically, code-switching at the juncture of an auxiliary modal verb and an infinitive, which characterizes modal infinitival constructions. The way in which the participants deal with the specific constraints will be investigated in the naturalistic data of our longitudinal corpus and the experimental data collected via the experimental task.

Chapter 3 will provide an overview of this twofold study, including the research questions and hypotheses, the description of the participants who completed our two experimental tasks, the results of the Bilingual Language Profile (BLP: Birdsong, Gertken & Amengual, 2012), the description of the experimental tasks, the data collection procedure and a data analysis overview. Chapter 4 will present the results of the two experimental tasks along with the detailing of the modelling process carried out in the statistical program R to configure the models that best fit our data set. Chapter 5 will conclude with the discussion of the results and a summary of the main findings together with contributions to the field. This final chapter will also include the limitations of the study and suggestions for future research.

To the best of our knowledge, this thesis constitutes the first empirical attempt to investigate the status of NICs and MPs in the mind of six groups of Spanish–English bilingual speakers from these two different but complementary perspectives using both experimental and naturalistic data.

2. Modal infinitival constructions

The main purpose of this chapter is to introduce our Spanish modal infinitival constructions, namely those made up of an inflected verb with modal meaning followed by an infinitive. This chapter consists of three sections. In section 2.1 we present an introduction to the concept of modality as well as several proposals concerning how to classify its meanings; in section 2.2 we will explore different ways to express modality in Spanish and English, contrasting both modality systems with examples of some of the most common modal utterances in both languages; in section 2.3 we will focus on modal infinitival constructions in Spanish, starting with the contrast between Spanish and English when defining auxiliary and modal verbs as opposed to full verbs (in particular, *deber* and *poder* in contrast to *desear* and *preferir*) together with the discrepancies around distinguishing periphrases from nominal infinitival clauses in Spanish. This chapter will finish with a brief reference to the epistemic meaning of the verb *poder* (can/could/may/might), which poses a counterexample to the syntactic differences between modal periphrases and nominal infinitival clauses.

2.1. Modality and meaning

The concept of modality has been approached from different perspectives. From a philosophical point of view, its study dates from the 4th century BCE with Aristotle (Otaola Olano, 1988; Ridruejo Alonso, 1999). Since then, there have been three main disciplines that have dealt with modality: logic, linguistics and semiotics (Otaola Olano, 1988). From a linguistic perspective, the Swiss linguist Charles Bally (1950)—who defined modality as the "soul of the sentence"—argued that within every statement there are two elements: dictum and modus. Dictum refers to the "representation" of what is said whereas modus refers to the "attitude of the speaker". Equating

modality with the speaker´s attitude is an idea that is echoed in both English grammars (Carter & McCarthy, 2006) and Spanish grammars (RAE-ASALE, 2011).

Although modality is related to the speaker's attitude, the spectrum of meanings that modality encompasses goes beyond the speaker´s attitude, which explains why there have been many attempts to categorize its multiple and, sometimes, overlapping meanings. For example, Bybee et al. (1991) established three types of modality: agent-oriented (desire, obligation, ability, root possibility, permission), epistemic (possibility, probability), and speaker-oriented (imperative, hortative, optative). On the other hand, Hacquard (2011, as cited in Cournane, 2020), considers that there are two mayor divisions within modality: *forces*, that discern between what is possible (7a,b) and what is necessary (7c,d); and *flavours*, that distinguish between root (1a,c)—a subset of meanings[1] related to what is external to the speaker, the main event, and its participants—and epistemic (1b,d), related to the speaker´s knowledge:

Table 1

Modality: spectrum of meanings (Hacquard, 2011)

Sentence	Type of modality
7a. Felipe could do the exam. 'Felipe podría hacer el examen'	Root (deontic) possibility
7b. Felipe could be a good friend. 'Felipe podría ser un buen amigo'	Epistemic possibility
7c. I have to find my keys. 'Tengo que encontrar mis gafas'	Root (deontic) necessity

[1] These meanings include: bouletic (relative to desires), deontic (relative to social norms), dispositional (relative to biological needs), teleological (relative to goals), alethic (relative to analytic statements), or of ability (relative to competence). (Carrasco Gutiérrez, 2019; Cournane, 2020; Hacquart, 2011; Kratzer, 1981).

7d. The character with the red socks has to be the culprit. 'El personaje con los calcetines rojos tiene que ser el culpable'	Epistemic necessity

Note. Examples adapted from Cournane, (2020, p. 3)

Within this frame of meanings, Palmer (2001) proposes a distinction between Propositional and Event modality with a set of different meanings attributed to each one. Propositional modality refers to the speaker´s attitude towards a proposition, which can be either epistemic or evidential. Event modality refers to potential events and can be either deontic or dynamic[2]. Not all modality falls into this classification: the author also mentions presupposed propositions, negatives/interrogatives, and wishes and fears (which share both deontic and epistemic characteristics), among others. The classification proposed by Ridruejo Alonso (1999), Pihler (2014), or Criado de Diego & San Mateo Valdehíta (2020) is more concise; they focus on the distinction between epistemic modality (relative to the knowledge of the speaker; for instance, expressing doubt) and deontic modality (relative to rules; for example, giving an order). This binary distinction is commonly used when addressing this concept: Hoye (2013) considers epistemic and deontic as "the two main types of modality" (p. 285) and Portner (2009) as "an important traditional way of classifying varieties of modality" (p. 2).

In particular, Portner (2009) establishes an initial categorization of modality that is not based on meaning but in what portion of the discourse modality manifests itself. As a result, we could have sentential modality (at the sentence level), sub-sentential modality (within the clause),

[2] Some of these terms also come from the philosopher Von Wright (1951), who establishes four modal categories: (i) alethic, including those notions related to what is necessary, possible, contingent, and impossible; (ii) epistemic, which includes what is conceived of as certain, doubtful, or false; (iii) deontic, which implies what is obligatory, allowed, or forbidden; (iv) and existential, related to concepts such as universal, existent, and non-existent.

and discourse modality. In his analysis, he further centers on sentential modality when introducing another classification of modality: epistemic, priority, and dynamic modality. Epistemic, along the same lines as previously mentioned, would be those linguistic forms (modal verbs, in particular) related to the speaker's knowledge; priority would encompass deontic, bouletic, and teleological meanings; and dynamic would include volitional (ability, opportunity, disposition) and quantificational (existential and universal) meanings. Another approach is the one given by Love & Curry (2021). In their revision of different categorizations of modality, they conclude that epistemic, deontic, and dynamic modalities "are accounted for in some way across most studies" (540).

What makes the study of modality even more complex is that languages have linguistic forms (i.e., modal verbs) that can express root and epistemic meanings (see Table 1). These variable-flavour modal verbs "are ubiquitous across Indo-European languages" (Cournane, 2020, p. 4), However, according to this author, this variability does not extend to the force of the modal in English, expressed separately by a particular set of verbs that either convey possibility, like *can* or *might*, or necessity, like *must* or *have to*. Although there are languages with variable-force modals —Cournane (2020) highlights Gitksan among other Western North American languages. In particular, Spanish could be considered among the latter group because, as the (Pan-Hispanic) Dictionary of the RAE-ASALE (2005) points out, *deber* can convey necessity, as in "debo cumplir con mi misión" (I must fulfill my mission[3]) in the novel "Satanás", by Mendoza Zambrano (2002) or possibility, as in "Marianita, su hija, debe tener unos veinte años", (Marianita, her daughter, should be around twenty years-old[4]) in the novel "La fiesta del chivo", by Vargas Llosa (2000).

[3] Our translation.
[4] Our translation.

This is just a glimpse of what the concept of modality may imply, since a deep dive into "the soul of the sentence" is far from the scope of this thesis. For further reference, however, we will be using the terms mentioned in Hacquard´s (2011) classification of modality relative to *flavour* (root and epistemic) and *force* (possibility and necessity) because they cover a broad spectrum of terms. Most particularly, the meanings that will be considered in the experimental tasks are the root meanings of *poder* and *deber*—bouletic (relative to desires), deontic (relative to rules), and ability (relative to competence)—with the intention of narrowing the broad spectrum of meanings into a concrete selection (*flavour* ⊇ root ⊇ bouletic, deontic, ability[5]) and to avoid a particular ambiguity that exists around an epistemic use of the verb *poder* that will be briefly explored in section 2.3.2.1.

2.2. Ways of expressing modality

There are many ways to convey modality in Spanish (Rabadán, 2006) and in English (Cournane, 2020). Palmer (2001), however, introduces a larger framework stating that "there are two ways in which languages deal grammatically with an overall category of modality. These are to be distinguished in terms of (i) modal system and (ii) mood" (p. 4). In a modal system, modality is mainly expressed through a set of modal verbs, whereas in a mood system is primarily based on the binary contrast of assertion (declarative) and non-assertion (non-declarative) expressed in the indicative and the subjunctive moods, respectively, generally through verb inflection.

[5] These three root meanings (bouletic, deontic, and ability) are three of the six language functions or uses of language specified on the curriculum plan of the Cervantes Institute (2006): (i) to give and ask for information, (ii) to express opinions, attitudes, and knowledge, (iii) to express interests, wishes, and feelings, (iv) to influence the interlocutor, (v) to interact socially, (vi) to structure the discourse. The third function is equivalent to a bouletic meaning, the fourth function is related to a deontic meaning, and the second function can be used to express the ability to do something, which directly connects with ability.

The presence and co-existence of both modal and mood systems varies across the languages of the world. The author distinguishes between those languages, like German, that express modality through both systems, from other languages, like English or Spanish, where one of the systems is more prevalent than the other: while English favours a modal system, Spanish favours mood. However, as we will see in the following sections, an abundance of meanings can be conveyed through modality in both languages, both by verbal and non-verbal cues[6].

2.2.1. Modality in English

English has a modal system that puts the subjunctive mood in a secondary position, barely present in daily use (Kastronic & Poplack, 2021; Palmer, 2001; Hoye (2013) states that the use of modal expressions, (among them, the modal verbs), has acted to the detriment of the subjunctive, stating that the subjunctive presently "occurs in subordinate clauses in association with a very small class of verbs like demand, insist, require; and in main clauses it is restricted to archaic or formulaic utterances such as *'so be it!*'" (p. 39). As a result, English expresses modality through a set of core modal verbs (*can, could, may, might, would, shall, should, must* and *ought to*[7]), semi-modals (*used to, need*, and *dare*), and emergent modals (Leech, 2003) such as *be able to*, *be going to*, and *have got to*.

Apart from modal verbs, modality in English appears in less grammaticalized "modal constructions" (Krug, 2000) such as adjectives (i.e., possible), adverbs (i.e., maybe), other auxiliaries or verbs (Carnie, 2013; Portner, 2009), and other modal utterances (Hoye, 2013). Hoye highlights the combination of modal verbs and their "satellite" adverbs (i.e., can´t possibly) that behave as "units of modal expression" (p. 284). Palmer (2001) also concedes the future tense has

[6] See Cestero Mancera et al. (2020) Diccionario audiovisual de gestos españoles (uah.es) where some modality meanings (i.e., influence the interlocutor — deontic meaning) can be expressed through gestures and expressions.
[7] *Ought to* is also considered a semi-modal by other authors (Love & Curry, 2021)

an epistemic meaning, as in "someone knocked at the door, it will be Christopher", along with negation in terms of *force* meanings (possibility and necessity) when it is used within questions, as in "aren't you going to study for the exam?"

Recent corpora studies on the expression of modality in English (Leech, 2013; Love & Curry, 2021), point to the decreasing use of modal verbs in favour of semi-modals and other less grammaticalized modal constructions. Most particularly, in their study on informal spoken English between the 1990s and the 2010s, Love & Curry (2021) specify different trends between each group. Within core modal verbs, *could*, *might,* and *would* have increased in frequency, *must*, *shall,* and *will* have decreased in use, and *can*, *may* and *should* remain the same. Within the semi-modals, *ought to* and *need* (as an auxiliary) are less frequent than *dare* and *used to*. In reference to "lexical modality-indicating devices" (MIDs[8]), the use of *able to*, *need*, *possible*, *probable*, and *sure* has increased in comparison to other expressions such as *have to* and *want to*.

2.2.2. Modality in Spanish

Palmer (2001) states that Romance languages such as French, Italian, or Spanish, have a morphologically marked mood system (indicative, subjunctive, and imperative[9]), where mood is considered a grammatical category along the same lines as tense and aspect, "but they (Romance languages) also have a set of modal verbs (…) the modal verbs have not been fully grammaticalized" (p. 104). However, as occurs with English and its alternative ways of expressing modality, the mood system is only one way of conveying modality in Spanish.

[8] Love & Curry (2021) adopt the term MID (modality-indicating device) from Mubarak (2015). Their study on the frequency of MIDs, among other modal expressions, includes the following: *able, need, possible, probable, sure, have to, want, allow, certain,* and *going to*.

[9] In terms of modality, the imperative mood will be centered on expressing the deontic modality.

Otaola Olano (1988) considers that there are both explicit and implicit forms of expressing modality. Explicit forms occur "when the modal verb is different from the dictum verb, as in "Creo que llegaré pronto" (I believe that I will arrive soon) (100); whereas implicit forms manifest themselves through syntax in different forms: a modal verb (*deber*: must, should), a modal adverb (*quizá*: maybe), grammatical mood (subjunctive or imperative), and certain adjectives of judgment or appraisal (necessary, possible).

Along the same lines, Ridruejo Alonso (1999) lists several options, among them intonation, modal verbs, syntax, and adverbs[10]. The RAE-ASALE's grammar (2011) adds verb inflection, as in "Ven, Juan" / "Juan viene" (Come, Juan/ Juan comes) (p. 228), and Rabadán (2006) adds tense, aspect, and modal periphrases. Among all these possibilities, this study will focus its attention on the latter because Spanish modal periphrases communicate most of the lexical meanings that English modal verbs can express[11], especially through two verbs: *poder* and *deber* (Carlisle, 1980; Rabadán, 2006). In a contrastive analysis between English and Spanish, Rabadán (2006) concludes that periphrases headed by *poder* and *deber* transmit most of the modal meanings (possibility, obligation, and necessity) conveyed by the English modal verbs. Carlisle (1980), in another contrastive study, states that, along with tense and mood, *poder* and *deber* cover most of the semantic range expressed by English modals. Palmer (2001) also considers *poder* and *deber*, along with their equivalents in French (*pouvoir* and *devoir*) and Italian (*potere* and *dovere*), as "cognates" of the English modals in terms of meaning.

[10] For a diachronic study on the development of epistemic and evidential adverbs followed by the complementizer que, as in "Seguro que…" (Surely that…) see Kocher (2017)

[11] The verbs that can be inflected verbs of modal periphrases in Spanish are *poder* (can, may, might, shall), *deber* (should, must, ought to), *tener que, haber de* and *haber que* (have to), (Gómez Torrego, 1999).

There are other infinitival constructions whose inflected verbs can also express modal meanings. This is the case of nominal infinitival clauses (also called completives) with lexical verbs such as *querer*[12], *preferir,* or *desear* (to want, to prefer, or to wish). Semantically, these verbs can convey modal *root* meanings—as in "desea ganar la carrera" (she wishes to win the race) or "prefiero comer pescado en vez de carne" (I prefer to eat meat instead of fish), where both *desea* and *prefiero* express desires, that is to say, bouletic meanings—but they differ syntactically from modal periphrases despite their apparent structural similarity (an inflected verb followed by an infinitive). These differences will be discussed in the next section.

2.3. Modal infinitival constructions

As has been mentioned before, there are two types of modal infinitival constructions: modal periphrases (MP) and nominal infinitival clauses (NIC). In English, as a language characterized by a modal system with a series of modal verbs considered functional categories, there is no problem in differentiating MPs from NICs[13]. In Spanish, the superficial structure of both constructions—an inflected verb followed by an infinitive—has led to more discussion. In this section we will examine the contrast between the two languages in terms of the categorization of what constitutes an auxiliary and a modal verb, along with the differences between MPs and NICs in Spanish, which will finally lead us to the particularities of the epistemic use of the verb *poder*.

[12] The verb *querer* has been considered as a semi-modal with certain syntactic properties that differentiate it from other restructuring verbs (See Gómez Torrego, 1999; Grano, 2015; Picallo, 1990; A. o.)

[13] The attention is skewed towards distinguishing finite from non-finite clauses (infinitival clauses) (Carnie, 2013; Radford, 1988)

2.3.1. Modal and auxiliary verbs

In English, there appears to be a clear path towards comprising a list of English auxiliary and modal verbs that distinguishes them from each other and from full or lexical verbs. Within the main parts of speech, there seems to be agreement around auxiliary and modal verbs as belonging to functional categories (Carnie, 2013); more particularly, modals belong to the auxiliary structure in English syntax (Hakutani, 1978). When determining their specific characteristics, Jespersen (1940) made an initial distinction between modal and auxiliary verbs by affirming that a modal verb is followed by a "base infinitive", but it was Chomsky (1957, p. 39) who described auxiliary and modal verbs in more detail:

(28) (iii) *Aux → C (M) (have + en) (be + ing) (be + en)*[14]

(iv) *M → will, can, may, shall, must*

In particular, Chomsky argued that what differentiates modals from full verbs is the immutable order of (M + have + be), the impossibility of placing two modals together, and the absence of *do* in negation, questions, and tag questions. Almost a decade later, Twaddell (1963) and Palmer (1965) identified four basic features that differentiate lexical verbs from auxiliaries. Apart from negation, interrogation, and inversion, already mentioned by Chomsky, they added "echo-repetition or code, and stress for insistence or emphatic affirmation" (Palmer, 1965 as cited by Carlisle, 1980, p. 43; Twaddell, 1963). In fact, these features are what Huddleston (1976a, 1976b) named as the NICE (Negation, Inversion, Code, Emphatic affirmation) properties of English auxiliary verbs, which can be extended to modal verbs as opposed to full verbs:

[14] C = number, person, and tense; M = modal; en = past participle (perfective aspect); ing = gerund (continuous aspect). Chomsky will later remove (be+en)— the use of the verb *to be* in the passive voice—from this definition of AUX.

- **Negation:** auxiliary/modal verbs can be negated and occur with *not*, as in "I can´t sleep"
- **Inversion:** auxiliary/modal undergo inversion (the verb moves to a pre-subject position) in interrogative constructions as in "could I enter?"
- **Code:** auxiliary/modal verbs allow the following verb phrase to be deleted, as in the ellipsis "she could go and so could I (go)" or the tag question "we can go swimming, can't we?"
- **Emphatic affirmation:** auxiliary/modal verbs can be prosodically emphasized: "you must go to bed now."

However, as Palmer (2001) and Carnie (2013) state, modal verbs have their own formal features[15]: (i) they do not co-occur[16]; (ii) they do not follow other auxiliaries; (iii) they have no final *-s* for their third person singular; (iv) they have no non-finite forms—so they cannot be preceded by the infinitive marker *to*; (v) they have no imperatives; (vi) they are suppletive negative forms; (vii) they do not follow *not*; (viii) there are formal differences between modal verbs depending on their root or epistemic sense. In this respect, Palmer (2001) states the following differences:

- Deontic MUST has negative *mustn't* and a suppletive *needn't*, but epistemic MUST has no morphologically related negative.
- MAY NOT negates modality when deontic (no permission), but the proposition when epistemic ("It may be that it is not so")
- MAY and MUST followed by *have* are always epistemic, never deontic.
- MAY is replaceable by CAN only in the deontic use, though CAN'T may be epistemic.

[15] Some of these features related to modal verbs were already mentioned by Chomsky (1957)
[16] Carnie (2013) points to an exception to this rule with the combination of *might* and *could* done by some speakers from Southeastern United States, i.e., %He might could leave (p. 267)

(Palmer, 2001, p. 103)

When we move this debate to Spanish, we find a different scenario. Considering modal verbs as functional or lexical categories is an ongoing debate (Bosque, 1989, 2015; Bravo et al., 2015; Fábregas, 2005; Haspelmath, 2015; Iordăchioaia et al., 2013; Panagiotidis, 2015; Vivanco, 2019; A. o.), and this is especially apparent when we deal with auxiliaries and modals in Spanish. As a result, there is neither a clear list of modal verbs nor a clear definition of what constitutes an auxiliary; therefore, there is no agreement on whether to consider *poder* and *deber* as functional or as lexical categories.

In Spanish, the use of the term *modal* is often linked to the meaning that the verb conveys (root or epistemic modality); for example, Kovacci (1990) considers as modal verbs not only *deber*, *poder,* and *soler* (must/should, can/could/might/may, and used to), but also *mandar*, *querer,* or *decidir* (to order/to want/to decide), because they "describe the speaker´s attitude" (p. 129). Gómez Torrego (1999) regards *querer* and *poder* as modal verbs, while making a clear distinction between them: the latter is an auxiliary, the former, a full verb.

In relation to what constitutes an auxiliary in Spanish, there are authors that only consider *haber*, *ser,* and *estar* (to have and to be) as auxiliaries (Castillo, 2003; Hernanz Carbó, 1980; Klein, 1968; Olarrea, 2010; Schroten, 1974) while regarding *poder* and *deber* as fully lexical verbs. For these authors, when these verbs accompany an infinitive, they comprise a compound or multiclausal sentence with two independent verbs, "a compound sentence integrated by a main sentence and a completive infinitival sentence" (Hernanz Carbó, 1980, p. 423), a scenario that would allow its components (the inflected verb and the infinitive verb) to be separated and to establish possible subordination or coordination relationships between them. In addition to this argument, Spanish modal verbs, such as *poder* or *deber*, do not share some of the previously

mentioned properties that characterize English modal auxiliary verbs (Castillo, 2003; Hernanz, 1980; Schroten, 1974); for example, Spanish modal verbs can co-occur, as in (8) or they can appear in various places in the sentence structure that do not necessarily follow the order that characterizes AUX [(M) (have + en) (be + ing) (be + en)], as in (9):

(8) María tiene que poder cantar
*María has to can sing
'Maria has to be able to sing'

(Hernanz, 1980, p. 418)

(9) Juan habría debido llegar antes
*Juan would-have must/should arrive sooner
'Juan should have arrived sooner'

(Castillo, 2001, p. 119)

On the other hand, there are authors that also apply the term auxiliary to those inflected verbs that form periphrases (Gómez Torrego, 1999; Hadlich, 1971; RAE-ASALE, 2011; Vivanco, 2019). Gómez Torrego (1999) describes periphrases as "the union of two or more verbs that constitutes only one predicate head" (p. 3325). This syntactic union is formed by what he considers an inflected auxiliary verb—"a mere syntactic tool" that provides verbal morphemes linked to the verbal inflection (i.e, mood, tense, aspect, person, and number)—and a main verb, the head of the structure, that can be an infinitive, a gerund, or a past participle. The discussion in this thesis will be limited to modal infinitival periphrases and their contrast with nominal infinitival clauses with modal meaning[17].

[17] It should be noted that there will be an inclusion of aspectual periphrases (i.e., periphrastic future, *ir a*/going to) as experimental items in the second task of this research study. This inclusion will not contribute to the debate on the distinction between periphrases and nominal infinitival clauses but will pertain to the second part of the thesis research on bilingualism and code-switching.

As for modal auxiliary verbs, Gómez Torrego talks about "auxiliary uses[18]" of certain verbs and the fact that there is only one verb that always acts as an auxiliary: *soler* (to tend to/used to). In fact, *deber* without the infinitive has a completely different meaning, i.e., "debo mucho a mi hermana" (I owe a lot to my sister); however, according to the author "when verbs such as must, can, have to, or use to (*deber, poder, tener que,* or *soler*), among others, connect with an infinitive, they are always auxiliaries" (p. 3344). Contrary to *deber*, the verb *poder* does not lose its meaning when it appears without an infinitive, i.e., "no puedo con el examen" (I can't handle the exam).

An additional classification is the one provided by Picallo (1990). In her study on Catalan modal verbs, she makes a distinction between "true modal verbs" (*poder*, may/can; *deure/deber* must; *haver de/tener que*, to have to), semi-modals (*voler/querer*, to want; *gosar/atreverse*, to dare), and aspectuals (*començar a*, to start to; *acabar de*, to finish; among others). Conversely, Camacho (2018), who associates modality to the meaning of the verb (i.e., *poder* denotes capability or ability), distinguishes between lexical verbs and auxiliary verbs. Within this latter group he includes *haber, ir, estar, ser, poder, querer,* and *seguir*, (to have / to go / to be / can / to want / to continue) and he also specifies that most auxiliary verbs can behave as lexical verbs as well. Along these lines, in her analysis and classification of Spanish auxiliary verbs, Fontanella de Weinberg (1970) determines that when *poder* and *deber* express an epistemic meaning, they can be considered auxiliary verbs, but when they express root meanings, they are full verbs.

Parallel to this debate, there have been numerous accounts in previous literature of "restructuring constructions[19]": sequences of modal, aspectual, or motion verbs followed by an

[18] This idea of "auxiliary verb uses" has its echo in English with be, have and do (Carnie, 2013) that behave as main verbs when, for example, *be* is used as a copular verb (i.e., *my sister is beautiful*), *have* indicates possession (i.e., *my neighbour has two dogs*) and *do* implies accomplishment or performance (i.e., *do me a favour*).

[19] The process of restructuring implies the transformation "of a biclausal structure into a simple clause with a complex 'verb'" (Rizzi, 2000, p. 148), a precondition to clitic-climbing. Something that was also argued for by Kayne (1989)

infinitive, first introduced by the Italian linguist Luigi Rizzi[20] (1976, 1978, 1982) in relation to Romance Languages. This term led to the existence of restructuring verbs including modal verbs such as *deber*, must/should; *poder,* can/may/could/might; and other verbs, such as *querer*, want; *soler*, used to; *empezar,* begin; *terminar,* finish; *continuar*, continue; or *ir.* go; *estar*, be; and *haber*, have; can also be part of these verb sequences as the only auxiliaries (Rizzi, 1982). Thus, these restructuring constructions would include infinitive periphrastic constructions and nominal infinitival clauses.

Finally, on auxiliary chains there is the Split Hypothesis (Bravo et al., 2015; García Fernández et al., 2017). These authors make a distinction between lexical/opaque and functional/transparent auxiliary verbs[21]. The former are those that absorb the temporal and aspectual information, whereas the latter are those that contribute this information in such a way that functional auxiliaries, contrary to lexical auxiliaries, can modify other verbs but not be modified themselves in a verb sequence. For these authors, *ser* + participle, *estar* + gerund, *ir a* + infinitive, *haber* + participle, *acabar de* + infinitive are functional auxiliary verbs, whereas *poder* + infinitive, *tener que* + infinitive, *empezar a* + infinitive, etc., are lexical auxiliary verbs.

We will take the position that periphrastic verbs with modal meaning such as *poder* and *deber* are auxiliary verbs (a functional category) when they are part of modal infinitival constructions (inflected verb with a modal meaning + infinitive), as opposed to full or lexical verbs

who considers these infinitival sequences not complex VPs but multiclausal constructions, suggesting that clitic-climbing is due to 'a property of the infinitive INFL node in the null subject languages that allows clitics to move to the main clause' (Kayne, 1989, as cited by Picallo, 1990).

[20] Parallel to Rizzi (1976), Aissen & Perlmutter (1976) introduced a similar hypothesis named Clause Union (two clauses into a single clause) to account for the phenomenon of Clitic Climbing, distinguishing two types of verbs, trigger and non-trigger verbs, depending on whether they allow clitic climbing or not within an infinitive construction.

[21] For further references to auxiliaries' different grades see Anderson (2006), Cardinaletti & Giusti (2006), and Emonds (2006), A. o.

in NICs such as *preferir* and *desear*, consequently establishing a different syntactic relation with the infinitive. These differences will be described in the following section.

2.3.2. Modal periphrases vs nominal infinitival clauses

The superficial similarities in terms of meaning and structure that exist between MP and NIC have led them to be considered indistinguishable from each other, which would place *poder* and *deber* as well as *desear* and *preferir* as non-auxiliaries or full lexical verbs accompanied by an embedded infinitive (Hernanz Carbó, 1980; Klein, 1968, 1982; Olarrea, 2010; Schroten, 1974). One of the main arguments against regarding *poder* and *deber* as auxiliaries (functional categories) is the one exemplified by Castillo (2003), who considers that "Spanish modals are not auxiliaries but, rather, main verbs that subcategorize for a subordinate clause: specifically, epistemic modals are raising verbs (as in 10) (…) whereas non-epistemic modals are control predicates (as in 11)" (p. 122):

(10) Juan$_i$ puede [t$_i$ llegar pronto]

'Juan may arrive soon'

(Castillo, 2003, p. 121)

(11) Tú$_i$ puedes [PRO$_i$ llegar pronto]

'You can arrive soon'

(Castillo, 2003, p. 122)

Raising verbs, such as *parecer* (to seem), have common syntactic elements with *poder* and *deber*, as both raising and periphrastic verbs do not assign thematic roles. The difference between them lies in the fact that, with raising verbs, the infinitive that comes after the inflected verb can

be substituted by a pronoun (12–13), which, according to Gómez Torrego (1999), distances raising verbs from being an auxiliary in a periphrastic construction:

(12) Parece haber poca gente/que hay poca gente.

'There seem to be few people/It seems that there are few people'

(Gómez Torrego, 1999, p. 3342)

(13) Lo/Eso parece.

'It seems so'

(Gómez Torrego, 1999, p. 3342)

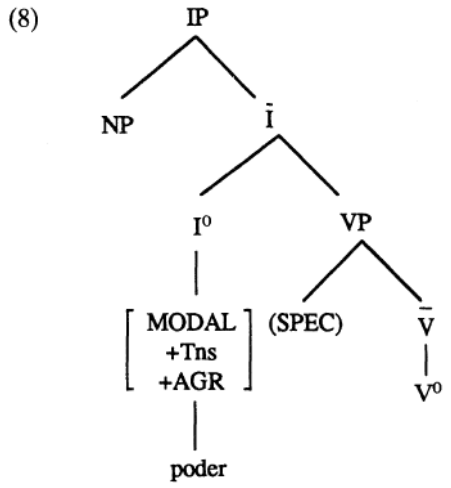
Castillo (2003) admits that *poder* and *deber* do not accept clitic pronominalization and explains this fact by affirming that they are not "case-assigners". However, with the exception of unaccusative verbs, all verbs assign case. The specific circumstance of *poder* and *deber*, as we have mentioned before, is that they do not assign thematic roles, something that would explain why *poder* and *deber* are not control verbs either.

Furthermore, Picallo (1990) provided a series of arguments against considering epistemic modal verbs as raising predicates and root modals as control predicates by considering those verb sequences with *poder* and *deber*[22] as "base-generated monosentential complexes where the infinitive verb is the one that establishes the argument structure" (285). In her analysis of Catalan modals and semi-modals—similar to Spanish modals in many respects—she made a particular distinction between these verbs in terms of their meaning: if the modal verb has an epistemic meaning, it is a constituent of INFL (Figure 1); in the case of the verb expressing a root meaning, it is considered a VP adjunct (Figure 2).

[22] In Catalan, *deure* (*deber*, must) has only epistemic meaning.

Figure 1

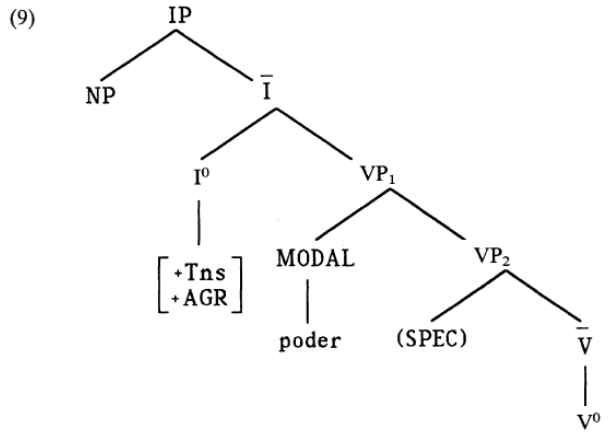
Tree structure corresponding to an epistemic interpretation of the verb 'poder'.



Note. Tree structure taken from Picallo, M. C. (1990, p. 288).

Figure 2

Tree structure corresponding to a root interpretation of the verb 'poder'.



Note. Tree structure taken from Picallo, M. C. (1990, p. 289).

To determine the main differences between MP and modal NIC, we will focus on what characterizes them. Verbal periphrases have two main properties: a single argument and the TAM meaning (temporal, aspectual, or modal) (Vivanco, 2019); the main difference with modal NIC is the first one. In addition, there are a series of syntactic tests used to identify periphrases and distinguish them from NIC (García Fernández, 2006; Gómez Torrego, 1999; Topor, 2005):

- The infinitive of an MP does not have nominal status, only verbal status, and, as a consequence, the infinitive cannot be replaced by a pronominalization; namely, it cannot be converted into a nominal element such as a demonstrative (14–15) or an interrogative (16–17):

(14) Prefiero comer > prefiero eso

'I prefer to eat > I prefer that'

(15) Debo comer > *debo eso

'I must eat > *I must that'

(16) Juan prefiere comer > ¿Qué prefiere Juan?

'Juan prefers to eat > What does Juan prefer?'

(17) Juan puede comer > *¿Qué puede Juan?

'Juan can eat > *What can Juan?'

- The infinitive is the only verbal form that determines the argument structure, this implies that it is the one that assigns thematic roles and, consequently, there is only one subject in a periphrasis. It also implies that nominal infinitival clauses (but not verbal periphrases) can have embedded clauses, "that sentences", as in (18) versus (19).

(18) Prefiero que comáis bien

'I prefer that you eat well'

(19) *Puedo que comáis bien

‘*I can that you eat well’

- Modal periphrases cannot create cleft sentences.

(20) Deseo comer > Lo que deseo es comer

‘What I wish is to eat’

(21) Puedo comer > *Lo que puedo es comer[23]

‘*What I can is to eat’

- In the periphrasis, the reflexive passive affects the entire periphrastic head: if the NP is plural, the auxiliary is as well, as in (23) versus (25).

(22) Se desea contratar buenos profesores

‘It is wished for to hire good teachers’

(23) *Se desean contratar buenos profesores

‘It is wished for to hire good teachers’

(24) *Se debe contratar buenos profesores

‘Good teachers must be hired’

(25) Se deben contratar buenos profesores

‘Good teachers must be hired’

- The inflected verb in a periphrasis cannot be passivized, as in (26) versus (27).

(26) Ganar la libertad es deseado por todos

‘Winning freedom is desired by everyone’

(27) *Ganar la libertad es debido por todos

[23] In the nominal substitution, interrogative substitution, and cleft sentences, the syntactic transformation is only be possible in English and Spanish if the auxiliary verb is followed by the light verb hacer (to do) -i.e., ¿Qué debo hacer? (What should I do?)-, in which case, *deber* would act as an auxiliary of *hacer*.

'*Winning freedom is must by everyone'

This list, which does not exhaust the syntactic tests that can be applied to distinguish periphrases from nominal infinitival clauses, will be the basis for the content of the first two experimental tasks of this study. These tasks aim to determine whether monolingual Spanish and English–Spanish bilingual speakers can differentiate between these two infinitival constructions. In order not to introduce too many variables in the experimental tasks, only three syntactic tests out of those previously described will be included: the pronominalization of the infinitive into a demonstrative, the pronominalization of the infinitive into an interrogative pronoun, and *that* sentences.

2.3.2.1. Epistemic *poder*

A particular mention should be given to a specific use and form of the verb *poder*. We have just stated that, within modal periphrases, a construction such as "*Puedo que comáis bien" (*I can that you eat well) is ungrammatical, yet *poder* can be followed by an embedded sentence under three conditions: third person singular, present simple, and an epistemic/possibility interpretation where the speaker manifests his or her point of view (i.e., making an inference), or if it is an expression of possibility (force modal meaning). If one of the three conditions is missing, for instance, when the subject (32) or the tense (33) changes, or the verb has a root meaning (i.e, ability), as in (29), the sentence becomes ungrammatical or does not make sense:

(28) Puede que vengamos

'It may/could be that we come'

(29) ?Puede que vengamos

'*He can that we come'

(30) Es posible que vengamos

'It´s possible that we come'

(31) Quizá/tal vez vengamos

'Maybe we will come'

(32) *Podemos que vengamos

'*We could that we come'

(33) *Pudo que vengamos

'*He could that we come'

A possible explanation for these differences is to consider this particular use of the verb *poder*, (present tense, third person singular, and epistemic or possibility meaning) as a raising construction (Rizzi, 1982); the expletive *it* (used in the translation of 28) "is only allowed in non-thematic positions, which are the hallmark of raising" (Carnie, 2013, p. 438). Superficially, the sentence (28) looks similar in English to those sentences including atmospheric verbs[24] with a null expletive subject ("it rains" / "it could be").

Another explanation follows the English translation of "puede ser que vengamos" (it may/could be that we come), where the verb *to be* is included in the sentence, giving, as a result, a complete periphrasis without losing its original meaning. In order to further explore this issue, we have done a succinct corpus search using the Spanish Real Academy corpus CREA (Corpus de Referencia del Español Actual). The search parameters were as follows: *puede que* in oral texts from Spain since 1975. The rationale behind this selection is that oral tests are closer to the vernacular variety of the language that written texts (Díaz-Campos, 2014). The specific objective was to determine whether the sentences allowed the inclusion of the verb *to be* (i.e., puede ser que

[24] Atmospheric verbs are considered lexical impersonal verbs by the Spanish Royal Academy (2011, p. 223).

/ it may be that) without sacrificing its epistemic/possibility meaning. The result of this search was 64 cases in 47 documents; from these 64 cases, 3 were not considered because there was a comma between puede and que, i.e., “porque cada uno se busca el dinero como puede, que es lo que, quieras o no…” (“because everybody looks for money as they can, which is, whether you like it or not…”). All 61 cases allowed the inclusion of the verb *to be* without losing or changing the epistemic/possibility meaning of the verb *poder*. As can be seen in the following example taken from the CREA (34), *poder* is used to infer (epistemic meaning):

(34) Algunos tienen trescientos años de antigüedad y puede (ser) que incluso más.
‘Some have three hundred years of seniority, and maybe even more’.

With this, it could be inferred that *puede que* is a periphrasis without the explicit inclusion of the verb *to be*, a verb that may have been lost in the same way that “it may be (that)” (31) evolved into *maybe* in late Middle English (XV B. C.) (Oxford Learner’s Dictionaries). It could also be the case that there has been a process of grammaticalization of *puede que* that can only be used in this morphologically unaltered form (present, third person, singular) with the meaning of possibility; in fact, in the dictionary of the RAE-ASALE (2014), *puede que* is considered an ‘adverbial expression’ with the specific meaning of *acaso*, *quizá* (perhaps, maybe).

Regardless of these possible accounts of this specific modal use of the verb *poder*, the experimental tasks in this study will only contain sentences where the verb *poder* has root meanings to avoid any ambiguity. This will also affect the verb *deber*, (although this verb does not allow embedded sentences regardless of its meaning). Consequently, the experimental items in all tasks will only convey deontic root meanings (deontic and ability).

3. On bilingualism

"Bilingualism is a result of languages (people, cultures, and/or regions) in contact" (Mahootian, 2019, 2). Migration, trade, war, conquest, colonization, geopolitical frontiers, and, more recently, mass media, the Internet, and, potentially, artificial intelligence, among other determinants, have driven and will drive the human being to encounter, acquire, and learn other languages. As a result, about half of the world's population is bilingual (Grosjean, 2021; Mahootian, 2019), and, according to Eberhard et al. (2021), two or more languages are known and used by 60%–70% of the world population daily. In Chomsky´s words (2000): "in most of human history, and in most parts of the word today, children grow up speaking a variety of languages, (…) that´s just a natural state of human beings" (p. 59).

Bilingualism and multilingualism have been observed and studied for at least five centuries (Wei, 2012), but it was in the 1970s when it became topic of scientific research for three specific disciplines: psychology, sociology, and linguistics, deriving into three main approaches for conceptualizing language—psycholinguistics, sociolinguistics, and formal linguistics. This thesis is mainly framed within formal linguistics, as it focuses on the structural constraints of code-switching in modal infinitival constructions.

We will specifically focus on two issues: the bilingual speaker and the outcomes of language contact. With respect to the bilingual speaker, we will first discuss different types of bilinguals and will then dedicate special attention to the specific characteristics of the bilinguals that will participate in this study, describing not only their linguistic background but also their sociolinguistic environment and taking into consideration the particularities of their countries of origin and residence. The second part of the chapter will be devoted to the main factors surrounding language contact, which will lead to a review of the main theories that have postulated code-

switching constraints that relate to lexical and functional categories in general, with a special focus on switches involving modal infinitival constructions, be they modal periphrases or nominal infinitival clauses. The chapter will conclude with the inclusion and a quantitative analysis of modal infinitival constructions in spontaneous data produced by two Spanish–English bilingual children.

3.1. Bilingualism: one speaker, two languages, and an I-language

When defining bilingualism, the first challenge is discerning between all those factors that occur within the pairing of one person with more than one language. Edwards (2014) points out how early definitions in the 1950s were too restrictive. These definitions were followed by broader conceptualizations of the term that still do not cover the complex reality of bilingualism, which, in unison with its multiple nuances, has led, according to Edwards (2014), to partial definitions.

Most definitions of bilingualism and multilingualism seem to have, at their core, the presence of a number of languages, their use, and the ability of the user (Aronin, 2022). As such, different authors define bilingualism by referring either to language use (Grosjean, 2021; Silva-Corvalán, 2014) or to language knowledge, proficiency, and competence of two or more languages in everyday life (Flynn et al., 2005; Montrul, 2008). This results in *balanced* versus *dominant* bilingualism (Lambert, 1955). This results in more nuanced types of bilingualism in response to the diverse parameters present in the language history of the speaker, such as the age of onset (AO) or the order of acquisition.

Other elements to consider when defining bilingualism and classifying bilinguals, as noted by Edwards (2014), are the differences between *receptive* (or passive) and *productive* (or active) bilinguals, along with the distinction between *primary* and *secondary* bilingualism (which consider the form of acquisition and the context in which the languages are spoken). However, if there is

anything in common among all these attempts to classify bilinguals and multilinguals it is that the boundaries between some of those different classifications fall into different subclassifications that are far from categorical, which, from a methodological point of view, makes the study of bilingualism and multilingualism even more challenging (see Goto Butler (2014) for further discussion on the definition of bilingualism and multilingualism).

From a psycholinguistic perspective, Wei (2014), who defines a bilingual or a multilingual person as "someone who can function in two or more languages in conversational interaction" (p. 33), emphasizes the neurological implications of being bilingual or multilingual in comparison to being monolingual. In fact, one of the central debates within psycholinguistics has been the organization of the lexicon and the grammar in the mind of the bilingual. This has led to the well-known *Differentiation Hypothesis* (De Hower, 1990; Genesee, 1989; Goodz, 1989; Meisel, 1994; Padilla & Liebmann, 1975), the *One System Hypothesis* (Swain, 1972; Volterra & Taeschner, 1978), and a third model that aligns with the *Differentiation Hypothesis* but adds the idea that the systems of the bilingual are interdependent due to the existence of crosslinguistic interaction or transfer (Serratrice, 2014). Beyond this debate and for the purpose of this thesis, we will consider two specific parameters to define the English–Spanish bilingual participants of this research: age and order of acquisition. These will be further explored in the following sections.

3.1.1. Types of bilinguals

The classification of bilinguals according to age and order of acquisition of both languages is intrinsically related to the type and level that the effects of the lexicon and the grammar of one language may or may not have on the other language. In terms of age of acquisition, there is a distinction between *early* acquisition and *late* acquisition of a second language. However, there is disagreement as to when to trace the age line between the two of them. For some authors, early

bilinguals would be those that acquire the second language between birth and the age of four or five (Johnson & Newport, 1991; Lenneberg, 1967; Liceras & Senn, 2009; Penfield, 1953; Penfield & Lamar, 1959). For other authors, early simply refers to childhood, whereas late refers to acquiring a second language after puberty or during adulthood (Montrul, 2008).

When it comes to order of acquisition, there is a major distinction between *simultaneous* and *sequential* bilingualism. Strictly speaking, the term *simultaneous*, also referred to by some authors as Bilingual First Language Acquisition (BFLA) (Genesee, 2000; Meisel, 2001; Silva-Corvalán, 2014), is applicable to those bilinguals that start acquiring both languages at birth (2L1). De Hower (1990), in line with Padilla and Lindholm (1984), rejects the term *simultaneous* and only uses the term BFLA versus *sequential* bilinguals, whereas other authors, such as Montrul (2008), not only accept the term *simultaneous* but apply it when the acquisition of the second language starts at some point between birth and the age of three[25], "before the linguistic foundations of the languages are in place" (p. 17). In other words, when the acquisition of the second language takes place after the third year, it is considered *sequential* bilingualism.

Within *sequential* bilingualism, there is also a distinction between those speakers who acquire the second language before receiving formal schooling, between four and six years of age (early second language acquisition or early sequential acquisition), and after this period (late child L2 acquisition and late adult sequential bilingualism) (Montrul, 2008). Along with these definitions, we also find the concept of *deep* bilinguals, which includes speakers characterized by a starting point of bilingualism between birth and two years of age, and a high level of competence of both languages in adulthood, when both languages are fully developed (López, 2020).

[25] When acquisition of the second language occurs within this period of time (between one and three years of age) it is considered successive bilingualism (Silva-Corvalán, 2014).

3.1.2. Bilinguals in the current study

Having established an introductory framework on the complex spectrum of bilingual speakers, this study will focus on three specific groups of adult English–Spanish bilinguals whose differences and similarities are at the core of many discussions not only in the field of language acquisition, but also in the field of teaching and learning a second or a foreign language: heritage speakers (HS), L2 students (L2S), and immigrant speakers (ImS). In particular, (i) Spanish HS living in an English-speaking environment and English HS living in a Spanish-speaking environment, (ii) Spanish L2S with L1 English and English L2S with L1 Spanish living in countries where the majority language is their L1, (iii) and Spanish ImS living in an English-speaking environment and English ImS living in a Spanish-speaking environment. It should be noted that the Spanish and the English level of the speakers is advanced[26].

With respect to HS, also referred to as second-generation immigrants (Silva-Corvalán, 1994), there are certain elements that condition their language environment: as children of a first generation of immigrants, they are exposed to a minority (immigrant) language from childhood (Liceras & Senn, 2009; Montrul, 2014; Polinsky & Kagan, 2007). This home language is usually in competition with the majority language, in terms of input and use, by its ubiquity, recognition and official status. There are others who define HS in a broader sense as those speakers who have a connection (familial, ethnic, or emotional) to a particular language, regardless of having been exposed to a given language as a child (Fisherman, 2014; McCarty, 2002). For the purpose of this dissertation, HS will be considered those who have been exposed to the minority language since birth.

[26] In this study we are not considering the early stages of Spanish L2 acquisition, especially in relation to L2S, due to the relatively complex structures that will be under investigation.

As noted by Montrul (2008), when HS and L1 acquisition are compared, the amount of input HS receive varies and the quality of this input is usually restricted and less varied in comparison with the HS' L1 acquisition. HS can experience fossilization, attrition[27], and incomplete acquisition. HS outcome is variable and typically incomplete, and, as a result, there can be disparities at every level of language, with inflectional morphology being the most affected, especially nominal morphology (gender, number, and case marking) over verbal morphology (Subject–Verb agreement and tense paradigms) (Montrul, 2014).

In terms of similarities, the setting for both HS and L1 acquisition is naturalistic, the mode of acquisition is aural, and the age of onset (AO) of the minority language is birth. In regards to the AO of the majority language—which in most cases will become the dominant language—we find again the previous classification of order of acquisition: (i) *simultaneous* bilinguals, when the exposure to both the minority and the majority language occurs before the age of 5; (ii) *sequential* bilinguals or child L2 students, when the exposure to the majority language occurs in preschool, 4–5 years; and (iii) *late* child L2 students, when the exposure to the majority language happens between the ages of 7 and 11 (Silva Corvalán, 1994). There are, however, authors that only consider early bilinguals—be they sequential or simultaneous—as HS (Polinsky, 2018).

L2S are frequently compared to HS. Similar to HS, L2S vary in the quantity and quality of their input, with personal and affective factors also playing a role. L2S also experience fossilization, lack of acquisition, attrition, and incomplete acquisition. Regarding the differences, the setting for L2S can be both instructed and/or naturalistic, and the mode of exposure can be

[27] Attrition is defined as a "non-pathological neurocognitive destabilization or adjustment in language processing and a reconfiguration or repurposing of the L1 and L2 in adapting to new linguistic-experiential situations" (Birdsong, 2024). According to this author, attrition can occur in L1 or L2 and it can be reversed when domain-based dominance (contexts and frequencies of use) returns to initial values.

both aural and written. For L2S, the majority language is their L1, and it maintains its status as the dominant language (unless they emigrate), whereas, for HS, the majority language, although acquired after the minority language, tends to become their dominant language. Finally, the L1 of the L2S is fully developed when they have initial contact with the L2, something that can also occur with HS whose AO of the L2 is at 4–5 years of age (child L2 leaners) or between 7 and 11 years of age (late child L2 leaners), though the main difference is still the fact that the L1 of the L2S is the majority language, whereas the majority language for the HS is their L2.

When we talk about immigrant speakers, we are typically referring to those who emigrate as adults (Liceras & Senn, 2009). In terms of the AO of the L2, ImS face two potential scenarios: their contact with the new majority language occurs at the time they arrive to the new country, where they can receive formal instruction or not; or they may have had previous contact as L2S in their countries of origin. Most references to immigrant speakers, especially in the context of North America, deem the first scenario as the most common. In both cases, they are sequential bilinguals with different AOs, where the most common setting for L2 acquisition is naturalistic as well as informal. ImS have many characteristics in common with L2S, with an essential difference shared with HS: the majority language of the new country is their L2, which, depending on the age at immigration and the length of residence (LOR), among other factors, may have an impact on both the ultimate degree of attainment of the L2 and their L1 grammars (Liceras, 2017; Stevens, 2006). For example, among the effects on their L1 grammar highlighted in previous literature, attrition is one of the most prominent, especially when the speaker has been living in the new country for more than 10 years (Köpke and Schmid, 2004; Schmid and Köpke, 2007: Selinger, 1996).

A key element in the process of acquisition for ImS—which can be extended to the other two type of bilinguals in this study (HS and L2S)—is the relationship of the individual with his or

her linguistic environment (the presence, accessibility, and status of the languages of the speaker). Although specifically attributed to HS, Montrul (2014) reckons that the "sociopolitical environment where the two languages are used—in addition to specific family circumstances and patterns of language used during the period of language development—play a crucial role in the ultimate attainment of the heritage language" (p. 184). This is something that may also be applied to other bilingual speakers, especially ImS. To understand the linguistic environment of the participants in this study, we are going to introduce some information in relation to the different status of English and Spanish as foreign or immigrant languages in their respective countries: English in Spain, Spanish in the United States and Canada[28].

3.1.2.1. The (socio)linguistic context of Spanish–English bilinguals

While Spanish in Canada is a minority language with some growth among other immigrant languages, it is far below the social, economic, political, and educational presence of the official languages of the country (English and French). Spanish in the United States, although considered a minority non-official language (except in New Mexico and the Commonwealth of Puerto Rico), has a stronger political, cultural, mediatic, and educational imprint in the country, to the point of forecasting the United States as having the second largest Spanish-speaking community in the world in 2060, only below Mexico (Bueno Hudson, 2021). At the same time, English in Spain, despite also being a minority language, is considered a prestige language with language policies that reinforce this impression. We will see how this change in the status of a foreign language in a country could counterbalance the usual pattern that generally characterizes immigrants, especially in the United States, where the second generation of immigrants, still bilingual but dominant in

[28] Apart from Spain, the USA, and Canada, there are two other countries of residence in this study (Bolivia and the United Kingdom) that will not be explored in the next section because of the small number of participants residing in these places (1 in Bolivia, 3 in the UK); as we will see in Chapter 4, these participants represent the 4% of the total of participants.

the majority language, leads to a third generation of monolinguals in the majority language (Edwards, 2014).

According to Statistics Canada (2022), Spanish is the fourth most spoken immigrant language (this excludes the Canadian official languages, English and French, and the Indigenous Languages[29]). It is spoken at home by 8.1% of the population, preceded by Cantonese (10%), Punjabi (13.4%), and Mandarin (13.7%). Within large urban centers, Mandarin is the main non-official language spoken at home in Toronto and Vancouver, whereas Spanish and Arabic (both at 16%) are the first ones in Montreal. From an educational perspective, Spanish is the most studied foreign language in Canada, although it varies among the provinces, with Alberta and Quebec being the provinces where Spanish has a stronger presence (Fernández Alonso et al., 2021).

In the United States, where English is the only official language, Spanish is the largest minority language in the country, spoken by 41.8 million of people (12.5% of the population) and by 61.6% of speakers within the group of "Languages Other Than English" (LOTE) in American homes (Dietrich & Hernandez, 2022), far surpassing the second most spoken minority language, Chinese[30] (5.2% within LOTE). However, as Lacomba (2023) points out, during the last decades there has been a decline in the number of Latin American immigrants in the United States. This, paired with an increase in the number of Latin Americans born in the US, may affect the use of Spanish in the future given the decline in the use of Spanish that occurs in the third generation of immigrants, where "75% of Latin Americans tend to speak English more fluently, and around 24% of Latin Americans of this generation report being Spanish–English bilingual[31]" (p. 40).

[29] More than 70 distinct Indigenous languages are currently spoken by the First Nations, Métis and Inuit peoples in Canada (Statistics Canada, 2023).
[30] The US Census Bureau does not distinguish between Mandarin, Cantonese, or any other Chinese languages.
[31] Our translation.

Parallel to these demographic movements, another factor to consider in relation to the status of Spanish in the United States is its position as a language of study. At the university level, Spanish is the language with the highest numbers of students since 1995, though there has been a decrease in enrollment since 2013, especially in Community Colleges (Looney & Lusin, 2019); in K–12, Spanish is also the foreign language with the greatest number of students (American Councils for International Education, 2017).

Spanish is the official language in Spain, though there are other languages that have the status of official languages in their respective Autonomous Communities (i.e., Catalan in Catalonia, the Valencian Community, and the Balearic Islands; Euskera or Basque in the Basque Country; and Galician in Galicia). In Spain, 56.3% of the population between 18 and 64 knows a foreign language, and English is the most spoken of these (40.3% of the population), followed by French (14%), German (2.8%), Italian (2.6%), and Portuguese (1.7%), (Spanish National Institute of Statistics, INE, 2016). According to the last Eurobarometer from the European Commission (2012) on attitudes towards bilingualism, 82% of Spaniards consider English as one of the two most useful languages for personal development, and 92% regard English as the most useful language for children to learn for their future. This positive perception of English by the Spanish population is also reinforced at an educational level, since English is the most studied foreign language in the country from kindergarten to Grade 12. In the 2021–2022 academic year, the percentage of students with English as a foreign language was 85.1%,in Senior Kindergarten, 98.9% from Grade 1 to Grade 6, 98.3% from Grade 7 to Grade 10, and 94.8% in Grades 11–12 (Spanish Ministry of Education and Vocational Training, 2020). These data suggest that English is considered a language of prestige in Spain.

Having described the main characteristics of our participants and their language environments, we now proceed to explain which of the main constructs related to bilingualism will be examined in our study.

3.2. Bilingualism: languages in contact

At the core of the theoretical and empirical research on bilingual language acquisition there are three main constructs that have caught the attention of contemporary scholars: language dominance (LD), crosslinguistic influence (CI), and code-switching (CS). LD has mainly been defined in relation to proficiency (Gertken et al., 2014; Grosjean, 1982; Unsworth et al., 2018; Wei, 2000; A.o.), though it is not the only approach: Wapole (2000) defines LD in terms of velocity of development, Nicoladis and Secco (1998) consider vocabulary size, and Fernández Fuertes and Liceras (2018) refer to the "grammaticalization of features in the various languages" (p. 160). LD has also been measured in terms of proficiency, use, and a combination of these two factors (Lloyd-Smith et al., 2021). It is worth noting that balanced bilingualism is rare (Silva-Corvalán, 2014), and it is more often the case that one language is more dominant than the other, and, depending on the particular circumstances of the speaker (quality and quantity of input, affective factors, age, mode, and order of acquisition, etc.), this equilibrium between both languages can change in favour of one or the other along the individual's lifespan.

According to Serratrice (2014), there is agreement that a bilingual speaker has two independent but not isolated systems, in comparison to the lone system of a monolingual speaker. This leads us to discuss the second construct: crosslinguistic influence. In this study, we face different scenarios with our three types of bilingual speakers. For L2 adult students and ImS, for whom, according to the *Full Transfer/Full Access Hypothesis* (Schwartz & Sprouse, 1996), the L1 is fully developed when initial contact with the L2 occurs, the L1 can facilitate development and,

at the same time, be the cause of errors made in the target language. Sometimes those errors are due to the contrastive features that characterize both languages[32]. This transfer from, or interlinguistic influence between, L1 and L2, however, acquires a different dimension when dealing with HS. In the case of HS, especially when the acquisition of both languages is simultaneous from birth or, at least, before the age of three, the output shows a great deal of variability depending on the quantity and the quality of the input in both languages through the years (something that could also be applied to L2 adult students and ImS) which may lead either to attrition or lack of acquisition in either of the languages (in the case of L2 adult students and ImS, lack of acquisition would essentially affect their L2).

Interlinguistic influence can occur at the phonetic, phonological, semantic and morphosyntactic levels, with the latter being the focus of this study. Silva-Corvalán (2014) states that the effects caused by crosslinguistic influence at the morphosyntactic level can be either permanent or temporary. This influence, according to Silva-Corvalán, can manifest itself as a delay or acceleration of acquisition of one or both languages, the production of non-target L1 constructions, the use of parallel constructions in misplaced pragmatic contexts, or as a transfer of "forms and meanings" from one language to the other. Another classification of cross-linguistic effects at the morphosyntactic level is the one provided by Serratrice (2014), who mentions three potential outcomes: omission, overproduction, and unattested L1 constructions. Also at the heart of this crosslinguistic influence lies the combination of features in the functional and the lexical categories of the language, with the dominant language being the one that "provides the functional category whenever that category is highly grammaticized" (Fernández Fuertes & Liceras, 2018,

[32] Among the main contrastive features that distinguish the languages of this study, it should be noted that Spanish is a [+null-subject] language, as opposed to English which is a [–null subject] language; also, Spanish is more morphologically oriented whereas English is more syntactically oriented (Baker, 1995).

pp. 159), which means that, in the case of English–Spanish bilinguals, a code-switched utterance such as "La$_{(fem)}$ table" will predominate over "The mesa", because the Spanish Determiner is the one that carries a Gender feature, not the English one.

The third component of this trio of constructs on bilingualism is CS, the alternation of two or more languages in an utterance. This alternation can be between sentences (inter-sentential), as in (35), within a sentence (intra-sentential), as in (36), a single word in a sentence (insertional switching or borrowing), as in (37), and in tag sentences (tag-switching), as in (38):

(35) *Llegamos a los Estados Unidos en los 60.* New York was our home.
'We arrived in the United States in the 60s. New York was our home.'

(36) Code-switching among bilinguals *ha sido la fuente de numerosas investigaciones*.
'Code-switching among bilinguals has been the source of numerous studies.'

(37) *El estudiante leyó el libro en el* reference room.
'The student read the book in the reference room.'

(38) It is raining a lot these days, ¿*verdad*?
'It is raining a lot these days, isn´t it?'

English–*Spanish* (Toribio, 2001, p. 205)

Code-switching, which has been shown to be a sign of "sophisticated linguistic knowledge" (Montrul, 2008, p. 105), can vary depending on the language dominance of the speaker to the point that the type and degree of competence in CS is associated to the degree of competence in both languages; for example, balanced bilinguals are keen to code-switch intra-sententially, whereas unbalanced bilinguals tend to favour inter-sentential CS (Belazi et al., 1994; Poplack, 1980; Toribio, 2001). Another way to exemplify the role of language dominance on bilingual CS is through the *Ivy Hypothesis* (Bernardini & Schlyter, 2004), according to which "elements of the

lower portions of the (syntactic) tree are lexically instantiated in the Weaker Language, whereas elements of the higher proportions are realized in the Stronger Language" (p. 50), as in "ett table" where the Determiner *ett* is contributed by Swedish (Stronger Language in Bernardini & Schlyter´s example) and the Noun "table" by French (Weaker Language in Bernardini & Schlyter´s example).

Having said this, our study will only focus on intra-sentential CS in spite of the fact that our participants vary in terms of their dominant language—depending on the type of speaker (HS, L2S, ImS), their place of birth, and their current residence (Spain, Canada, or the United States). The participants are English dominant, Spanish dominant, or balanced bilinguals, something that has been determined by administering the Bilingual Language Profile (BLP: Birdsong, Gertken, & Amengual, 2012). This focus on intra-sentential CS is due to the specific characteristics of the observed constructions (modal periphrases and nominal infinitival clauses), where the change of language (Spanish into English or English into Spanish) will occur between the inflected verb (functional, as in 39–40 or lexical, as in 41–42) and the infinitive, with the potential inclusion or exclusion of the infinitival marker *to* between the verbs:

(39) The partner must (to) *aumentar el negocio.*

'The partner must increase the business'

(40) *La profesora debe* (to) give clear examples.

'The teacher must give clear examples'

(41) The worker prefers (to) *salir del edificio*.

'The worker prefers to leave the building'

(42) *El compañero prefiere* (to) call his friends.

'The workmate prefers to call his friends'

Having introduced the object of study of the second part of this research—CS within modal infinitival constructions—in the following section we will explore some of the main grammatical approaches that are said to account for the specific constraints that affect this type of CS, namely intra-sentential CS between lexical and functional categories.

3.2.1. Code-switching constraints: lexical and functional categories

The study of CS has an official starting point in the late 1960s and the beginning of the 1970s (Myers-Scotton, 1993; A.o.). However, as noted by Benson (2001), there were earlier studies in this field in the first part of the twentieth century based on "language diaries of bilingual children and anthropological-linguistic investigations of bilingual communities" (p. 27), particularly in the United States. Most of those initial studies considered CS as a linguistic phenomenon characterized by randomness and lack of rules and it was in the late 1970s and the 1980s when it was observed that CS was rule-governed. Since then, one of the main aims of linguistic research has been to determine those rules.

As stated by Wei (2014), there are two opposing approaches to the study of the structural constraints of language mixing: one that only considers the interaction of two grammatical systems at an abstract level, and another that also incorporates the context in which CS occurs. This thesis will center on the former approach, mainly due to the type of data that will be analyzed. Although we will explore some naturalistic data within this section, the greater amount of data pertaining to this research has been elicited using experimental tasks.

Since the eighties, especially after Poplack´s (1980, 1981) ground-breaking articles about CS constraints, there have been different proposals to determine the grammatical restrictions on intra-sentential CS. Among these theories, we are going to introduce, succinctly, those that have reflected on their data switches concerning AUX/MOD–V and V–V (i.e., modal infinitival

constructions). As we have seen, these constructions can imply the switch between a functional and a lexical category, as in "Puedo eat/I can comer", or between a finite lexical verb and its complement, as in "Deseo (to) eat/I wish (to) comer". An early account of these specific restrictions in English–Spanish bilinguals can be found in Timm (1975) who observed that switches between a finite lexical verb and an infinitive, as in (43 to 45), or between a verb and its modal auxiliary, as in (46–47), were either rejected or not present in the analyzed data (a short story and a taped conversation):

(43) *(They) want *a venir*

(44) **Quieren* come

(45) **Quieren* to come

(46) *(I) must *esperar*

(47) **Debo* wait

English–*Spanish* (Timm, 1975, p. 478)

In relation to the switching between finite verbs and their infinitive complements (–45), Timm noted, however, that the sequence [Inflected verb (English) + particle (English) + infinitive (Spanish)], as in '(They) want to *venir'*, was regarded as "possible, though not tasteful" (p. 478) by speakers. Considering the second proposed constraint, the impossibility of switching between auxiliaries and verbs, Poplack (1981) presented a counterexample in her formulation of CS constraints[33]: "So you take the ham… as they're *ablandando, ya que está un poquito hirviendo, tu le echas el güeso del jamón*" (So you take the ham… as they're softening, as they're boiling a

[33] *Free Morpheme Constraint*: a switch may occur at any point in the discourse at which it is possible to make a surface constituent cut and still retain a free morpheme (1981, p. 175).
Equivalence Constraint: Codes will tend to be switched at points where the surface structures of the languages map onto each other (1981, p. 175).

little, you throw in the ham bone) (1981, p. 174). Similar counterexamples can be found in another pioneering study made by Pfaff (1979).

From the very beginning, as observed by some authors (Lopez, 2020; MacSwan, 2014a, 2014b), finding exceptions and counterexamples has been the pattern in subsequent endeavours to provide categorical assertions on CS constraints. However, in our next review of the main CS theories concerning modal infinitival constructions, we will neither include the counterexamples already mentioned by the previous literature nor will we enumerate the many theories around CS constraints formulated within the Generative syntax tradition[34]. We will limit ourselves to those theories and proposals that have included modal infinitival constructions in their data in order to further contrast their assumptions with the results gathered from our naturalistic and experimental data. We will see how the discussion of these theories focus on the acceptance/rejection of switches between functional (AUX/MOD) and lexical categories (i.e., modal periphrases), rather than on switches between lexical categories (i.e., nominal infinitival clauses), a discussion that, as we have mentioned, started with work such as Timm (1975) but which has not reached a definitive conclusion 50 years later.

We will start with the only proposal that does not question the possibility of switches between functional and lexical categories: the *Government Constraint* (GC) proposal, formulated by Di Sciullo, Muysken, and Singh (1986), according to which CS is allowed between two categories when there is no government relation between them. Contrary to other theories, the GC would permit CS between an auxiliary verb or a modal verb and a non-finite verb (i.e., modal periphrases) because "modals and tense auxiliaries do not govern the V, which can have a different

[34] For a further exploration on the theories on CS see Boztepe (2003), Chan (2003), López (2020), MacSwan (2009, 2014a, 2014b), Myers-Scotton (2002), Panhwar & Buriro (2020), A. o.

language index" (p. 15). In order to prove this assumption, the authors provided the following French–Italian code-switched utterances from Di Sciullo et al.'s (1976) spontaneous corpus data. In (48) there is a sequence with the auxiliary *avere* (have) plus the past participle of the verb *donner* (give). In (49) the verb sequence is formed by the modal verb *potere* (can) plus the infinitive *aller* (go):

(48) No, parce que *hanno* donné des cours.

'No, because they gave (have given) lectures.'

(49) Oui, alors j´ai dit que *si poteva* aller comme ça.

'Yes, so I said that we could go like that.'

French–*Italian*, (Di Sciullo et al., 1986, p. 15)

With opposing conclusions in relation to modal periphrases, we have the Functional Head Constraint (FHC) proposed by Belazi, Rubin, and Toribio (1994), which these authors defined as follows: "The language feature of the complement f-selected by a functional head, like all other relevant features, must match the corresponding feature of that functional head" (p. 228). This principle specifically disallows switching between modal and auxiliary verbs (a functional head) and a VP, as in (50–51) and in (54–55), while allowing switches between a lexical head and its complement, as in (52–53) and (56–57):

(50) *Je serai *sae :fir-t fi-1-ʕašra*.

I will.be went-I at-the-ten

'I will have gone by ten o'clock.'

(51) **N-ku:n parti* à dix heures.

I-will.be gone at ten hours

'I will have gone by ten o'clock.'

(52) Je serai parti *fi-l- ʕašra*.

I will.be gone at-the-ten

'I will have gone by ten o'clock.

(53) *N-ku:n sae.fir-t* à dix heures.

I-will.be went-I at ten hours

'I will have gone by ten o'clock.'

French–*Tunisian Arabic* (Belazi et al., 1994, p. 229–230)

(54) *The police officers have *visto un ladrón*.

'The police officers have seen a thief.'

(55) **Los policías han* seen a thief.

'The police officers have seen a thief.'

(56) The police officers have seen *un ladrón*.

'The police officers have seen a thief.'

(57) *Los policías han visto* a thief.

'The police officers have seen a thief.'

English–*Spanish* (Belazi et al., 1994, p. 229–230)

We should note that the examples included in the article as support for the inadequacy of switches between a modal auxiliary and a VP only refer to auxiliary verbs (to have/*haber* and être/to be/*ser*) not to modal verbs[35]. It was Toribio (2001) who specifically included examples of

[35] AUX/VP switches have been explored in subsequent studies with the result that depending on the auxiliary (*estar*/to be or *haber*/to have) the grade of acceptance varies. When conducting similar eye-tracking studies, Dussias (2003) and Guzzardo Tamargo (2012) arrived at the same conclusion: bilinguals treated AUX/VP switches with *haber* differently from those AUX/VP switches with estar. In another study involving HS and L2S, Giancaspro (2015) reached a similar outcome: both types of participants accepted estar/be + VP switches while rejecting haber/have + VP switches. In addition, the Block Transfer Hypothesis (BTH: López, Alexiadou & Veenstra, 2017) also accounts for these differences alleging that, in the case of estar and contrary to the verb *haber*, the auxiliary verb and the

switches between modal verbs and inflected verbs, as in (58–59), to exemplify what the FHC does not allow:

(58) *The president will *dirigirse al público*

'The president will address the public'

(59) **El candidato puede* prepare his remarks during the flight

'The candidate can prepare his remarks during the flight.'

English–*Spanish* (Toribio, 2001, p. 209)

Another theory around functional heads is the one proposed by González-Vilbazo (2005), the Principle of Functional Restriction (PFR), according to which CS between two functional heads is not allowed. In relation to those switches between AUX/VP, the PFR would not permit CS between an auxiliary and a complement. However, González-Vilbazo's Spanish/German participants allowed different switches depending on the directionality of the switch, since they rejected sentences with a German auxiliary and a Spanish participle, as in (60), but accepted switches between a Spanish auxiliary and a German participle, as in (61):

(60) *Du hast es ihm *contado*.

Have you told him?

(61) *¿Lo has* getroffen?

Have you met him?

German–*Spanish* (González-Vilbazo, 2005, p.87)

progressive morpheme are in different phases, switches between them are permitted. The BTH states that "the material that is transferred to the interfaces is sent in one fell swoop'.

The following table displays the different types of switches accepted and rejected by González-Vilbazo's (2005) German–Spanish participants:

Table 2

Summary of German–Spanish AUX–V switches (González-Vilbazo, 2005)

		Participle	
		German	Spanish
Auxiliar	German	+	-
		(hat erzählt)	(*hat *contado*)
		'has narrated'	'has narrated'
	Spanish	+	+
		(*ha* erzählt)	(*ha contado*)
		'has narrated'	'has narrated'

Note. Table translated from González-Vilbazo (2005, p. 88)

Another study from Bartlett and González-Vilbazo (2013) on the syntactic structure of the Taiwanese determiner phrase (DP), where Taiwanese–Spanish bilinguals rejected or accepted switches between classifiers and nouns depending on the directionality of the switch, lead them to conclude that phrase structure is not universal in the sense that a switch between a head in L1 and its complement in L2 would not be possible when the category of the complement does not exist in L1. Taking into account the directionality of the change is something that has been seen in other CS theories and proposals. This is the case of the *Matrix language frame* model (MLF) (Joshi, 1985; Azuma, 1991; Myers-Scotton, 1993, 2002) that distinguishes between the *matrix* language (ML) and the *embedded* language (EL), with the ML being the one that determines the surface

structure positions for lexical and functional categories in such a way that function morphemes (system morphemes) in code-switched ML + EL fragments can only be drawn from the ML.

Another theory that contemplates the directionality of the change is the *Grammatical Features Spell-Out Hypothesis*, (Liceras et al., 2003; Liceras et al., 2005; Liceras et al., 2006; Liceras et al., 2008), but, in this case, the prevalence of one directionality over the other is determined by the features of the functional categories involved in the switch; more specifically, the bilingual speaker "will make CS choices that will favour the functional categories containing highly 'grammaticalized' features" (2005, p. 829); for example, "*Los*$_{(masc,\ plural)}$ cars" would be preferable to "The *coches*" because the Spanish Determiner is the one with Gender and Number features.

In a specific reference to modal constructions and their features, López (2020) noted that "if the two languages involved in CS have auxiliaries with the same feature content, CS between the auxiliary and the complement should be allowed" (201), as in (62), an example of CS between a modal verb and an infinitive verb belonging to two Creole languages, Sranantongo and Nengee, from Migge (2015), where the modal verb (can) is in Nengee and the infinitive (tell) is in Sranantongo in the main sentence but the contrary occurs in the subordinate; the modal verb (must) is in Nengee and the infinitive (end) in Sranantongo:

(62) I kan *taagi* en taki a tori dati *mu* skotu.
you MOD tell him say DET story DEM must end
'You can tell him that issue must end'

Sranantongo–*Nengee* (Migge, 2015, p.189)

This idea of not focusing on specific constraints but on a broader theoretical context is adopted by another group of authors—representatives of the so-called *Null Theory* (López, 2020)—whose conclusions on infinitival constructions lead to the idea that the rejection/acceptance of these switches depends on the directionality of the switch and the grammars of the languages involved. An example of this is the *Tree Adjoining Grammar* (TAG) model (Mahootian, 1993; Mahootian & Santorini, 1996), based on general principles of phrase structure. This model proposes that "heads (functional or lexical) determine the syntactic properties of their complements in CS and monolingual contexts alike" (1996, p. 470). Referencing infinitival constructions, the authors propose that CS would be allowed within those constructions only if 'subcategorizational restrictions', imposed by the inflected head, are respected. To illustrate this proposal, they mentioned the work on Arabic–French CS by Bentahila and Davies (1983). These authors noticed that in sentences in which an Arabic auxiliary introduces a French complement verb the switch is only accepted when the French verb is accompanied by Arabic inflection, as in (63). This is the case because certain Arabic verbs, such as /xeSS/ (to have to), are 'subcategorized for a finite verb complement''; in other words, the added Arabic inflection for tense and person that transforms the non-finite French verb into a finite verb is what makes the switch acceptable. In fact, the examples in (64–65), with no added inflection or with the reversed pattern (main verb in French and the complement in Arabic), were rejected:

(63) ʕad xSSu j *redouble*

'again he needs imperfect — repeat (again he needs to repeat)'

(64) *ʕad xSSu *redoubler*

'again he needs to repeat'

(65) **je dois* nSeLi

'I must I pray (I must pray)'

Arabic–*French* (Bentahila & Davies, 1983, p. 315–322)

Another proposal within the *Null Theory* is Chan's (2008) work on word order and lexical/functional heads. This author claims that functional heads, contrary to lexical heads, determine the word order of their code-switched elements. Chan's data contains instances of switches between functional and lexical categories, in particular, modal verbs and infinitives, as in (66–68), and between the semi modal *to want*[36] and an infinitive, as in (69):

(66) I have to ttakē my hand

'I have to wash my hand'

English–*Korean* (Choi, 1991, p. 889)

(67) anta *car*-ei *drive* paNNanum
That car-ACC drive do + must
'We must drive that car'

Tamil–*English* (Sankoff et al., 1990, p. 80)

(68) Can I *nigeru*?

'Can I escape?'

English–*Japanese* (Nishimura & Yoon, 1998, p.127)

(69) ωsi *Panjabi learn* kerni a

we Panjabi learn do

'We want to learn Panjabi'

Panjabi–*English* (Romaine, 1995, p. 140)

[36] See Footnote 12, Chapter 2.

One of the most well-known attempts to explore the role of the specific grammars of the languages involved in the switch is MacSwan's (1999, 2000, 2009, 2014a, 2014b) *Minimalist Approach* (MA), a proposal based on Chomsky's Minimalist Program, which states that CS is only constrained by the interaction of a bilingual's two grammars. In his initial outline of the MA and in relation to modal infinitival constructions, MacSwan makes a direct reference to Rizzi's restructuring verbs (1982), focusing on the particular case of the verb *querer* (to want), using the following examples:

(70) He wants to *hacer La cena*

'He wants to make dinner.'

(71) *He wants *hacer La cena*

'He wants to make dinner.'

(72) *He wants *a hacer La cena*

'He wants to make dinner.'

(73) **Quiere* make dinner

'He wants to make dinner.'

(74) **Quiere* to make dinner

'He wants to make dinner.'

English–*Spanish* (MacSwan, 1999, pp. 212–213)

According to MacSwan (1999), who also cites naturalistic English–Spanish data involving 'restructuring verbs' from other authors (Lipski, 1978; McClure, 1981), only the switch in (70) (inflected verb in English, particle in English, infinitive verb in Spanish) would be allowed, something that he explains by saying that "these code-switched constructions converge if and only if restructuring is not forced by movement" (212). As we mentioned in Chapter 2 when discussing

modal infinitival constructions, Rizzi regards *poder* and *deber* as restructuring verbs and, as a result, we could infer that if the sentence includes *deber* or *poder*, the directionality of the acceptable switch would be the one in (64), Eng-MOD + Spa-VP, as in the following adapted examples (75–76)

(75) He must *hacer la cena*

'He must make dinner'

(76) He can *hacer la cena*

'He can make dinner'

The previous theories present us with three different scenarios concerning modal infinitival constructions, which will be tested in this thesis: (i) The first scenario, typified by the FHC or the PFR, would forbid switches between functional and lexical categories (i.e., modal periphrases), but would allow switches between lexical categories (i.e., nominal infinitival clauses); (ii) the second scenario, represented by the GC framework, would allow both type of switches; and (iii) the third scenario, featured in the *Null Theory*, more specifically MacSwan's MA, would take into consideration the specific grammars of the languages involved and the directionality of the switch, only allowing switches where the inflected verb is in English and the infinitive verb is in Spanish, regardless of the type of infinitival construction. In the following section we will test these theories against a small sample of naturalistic data.

3.2.2. Modal infinitival constructions in spontaneous code-switched data

Around 18 months old, when the one-word stage is left behind[37], child syntax is apparent in their multiword utterances and, consequently, in their intra-sentential mixed utterances in bilingual first language acquisition. This early stage of language mixing has been addressed differently in prior literature as questioning whether child and adult CS can be compared on the same basis in relation to grammatical constraints. One perspective regards this early mixing as unrestricted, in the sense that the child has not acquired the constraints present in adult CS (Köppe & Meisel, 1995; Meisel, 1989, 1994), syntactic knowledge that, according to these authors, will be acquired from 3;0 onwards. There is another perspective according to which those rules governing these early switches are the same for adult CS, consequently validating early child data as a means to determining what it is allowed (or not) in CS (Cantone, 2007; Cantone & Müller, 2005; Müller & Cantone, 2009).

In relation to complex constructions, modal infinitival constructions in particular, it is precisely from 24 to 36 months old when monolingual children start to acquire them. In their study on early acquisition of verb grammar in French and Austrian German of periphrastic constructions (i.e., modal verb plus infinitive–modal constructions), Bassano et al. (2004) concluded that these structures emerged from around 2;0 onwards in a range of time of about three months. Apart from modal constructions, these authors observed the emergence of two other complex structures, the compound past and analytic future, and their data indicated that modal constructions expressing "very basic states of mind, such as desires and abilities" emerged earlier than the analytic future in both languages. In English, this is the case for present tense auxiliaries along with complement-

[37] During the holophrastic stage of language development where "one word represents an entire thought" (Farrall, 2012) children start producing lexical categories; function words, such as auxiliary verbs, are much less frequent, even absent (Nelson, 1973).

taking verbs (e.g., I can + finite verb), with the earliest and most frequent being *wanna* or *hafta* (Diessel, 2004). She summarizes the development of infinitival clauses in the following way:

(i)	Modality verbs	> manipulative verbs	> cognition-utterance verbs
(ii)	NP–V–VP	> NP–V–NP–VP	
(iii)	Bare infinitives	> *to*-infinitives	> *wh*-infinitives
(iv)	Nonfinite complement	> finite complement	

(Diessel, 2004, p. 72)

In relation to Spanish[38], the fist verbal forms that appear in child speech are the imperative and the infinitive under the structure [a (prep) + protolexeme], as in "amí" (a dormir / to sleep), with the imperative (deontic)–(bouletic) function being one of the first ones expressed by the child (López Ornat et al., 1994). In their longitudinal case study of a monolingual Spanish child, Maria, from 19 months to 4 years of age, the authors saw how this initial structure evolved into [a + infinitive] to settle into the periphrastic future [ir/to go + a/to + infinitive] at around two years of age. At this age, another periphrasis will also appear, the present continuous [estar/to be + gerund], along with the first modal infinitival periphrases (77–81) and nominal infinitival constructions (82–83):

(77) Tenía que trabajá [=trabajar] mamá. (2;00)

'I had to work mom'

(78) No poro [=puedo] salí(r). (2;01)

'I can't go out'

[38] In her study of the bilingual acquisition of her two grandsons between birth and the age of six, Silva-Corvalán (2014) reported some Spanish monolingual complex constructions (i.e., infinitival constructions) before age 2;0: Necesito cortar pelo, papi. / '[I] need to cut [my] hair, daddy.' (1;10.21)

(79) No, no puedo sacalo [=sacarlo]. (2;01)

'I can't take it out'.

(80) Que teno [=tengo] que guadalas [=guardarlas] las cosas. (2;03)

'I have to put away the things'

(81) Hay que meté [=meter] el e pato. (2;03)

'You have put the duck'

(82) Toma, cógelo, que quere [=quiere] ir co [=con] su mamá. (2;02)

'Here, take it, because he wants to go with his mom'

(83) No quero [=quiero] seca(r)me, no quero [=quiero] que me saques [seques]. (2;02)

'I don't want to dry myself, I don't want you to dry me'

When we move to bilingual language acquisition, there are examples in the previous literature of code-switched modal infinitival constructions, either within modal periphrases or nominal infinitival clauses, at different ages and for different pairs of languages:

(84) Kannst du *move a bit*

'Can you move a bit?'

German–*English*. 2;4 to 2;9 (Gawlitzek-Maiwald & Tracy, 1996, p. 913)

(85) *Mami Giulia will* lavare *die hände*.

'Mummy Giulia wants to wash her hands'

Italian–*German*. 2;8 (Taeschner, 1983, p. 95)

(86) devi *finden* un *seil*

'[you] must find a rope'

Italian–*German*. 2;11 (Cantone, 2007, p. 207)

(87) *Tú tienes que* turn it on.

'You have to turn it on'

English–*Spanish*. 3;4 (Silva-Corvalán, 2014, p. 100)[39]

(88) non lo devi *verraten*

'not it CLIT must [you] betray[40]'

Italian–*German*. 3;6 (Cantone, 2007, p. 205)

These are insufficient examples to allow us to propose a consistent pattern. In fact, they only indicate that, as early as 2;4, children switch at the INFL phrase regardless of the directionality of the switch. In our search for this potential pattern, we will turn to the naturalistic data pertaining to this thesis research: a longitudinal study of the language production of code-switched infinitival constructions of two Spanish–English bilinguals who have been observed in their natural environments by their parents for two years in two different periods of their linguistic development due to their age differences. As such, the data originated in natural and uncontrolled discourse contexts, i.e., adult–child and child–child spontaneous interactions mainly at their home and driving to and from different venues. There are two siblings, the older, a male (Child A), was observed from 6;5 to 8;6; the younger, a female (Child B), was observed from 2;7 to 4;8. Both were born and raised in a home where both parents are native speakers of Spanish and first generation immigrants from Spain to North America; first to the United States, where the older son was born, then to Canada, where the younger daughter was born. The parents have a middle socioeconomic status, and both have completed graduate studies at the university.

[39] Although her study was not particularly focused on modal infinitival constructions but rather on grammatical subjects and copulative verbs, the order of constituents (subjects), and the development of verb morphology (tense–mood–aspect), the author reported switches at the boundary between a modal verb and an infinitive, a type of switch that was attested in both children.

[40] Our translation, in Cantone's text *verraten* is translated as *tell*.

The linguistic environment of these siblings is what characterizes, in this case, a Spanish heritage speaker; with Spanish being the minority language in the United States and Canada, and English the majority language—with the addition of French in the case of Canada. The parents (Spanish–English bilinguals) have addressed the children almost exclusively in Spanish from birth, although the children are exposed to English, and occasionally French, as well as Spanish, during nightly reading time. The children are also exposed to audiovisual content in four languages (English, French, Spanish, and Mandarin). Once a year they visit their parents' home country, Spain, and stay there for an average of three weeks per year in an immersive Spanish environment. The children also talk, at least twice a week, with members of their Spanish family, mainly their grandparents.

At home, although the children are exposed to other languages through literature and audiovisual content, their parents actively encourage the use of Spanish between the children; in addition, their mother has started to explicitly teach them Spanish. In addition, taking into consideration Lanza's (2007) parental discourse styles and strategies to foster bilingualism[41], when the children are code-switching while talking to their parents, they tend to either continue the conversation ("moving on") or rephrase the child's utterance completely in Spanish while questioning it ("expressed guess") or affirming it ("adult repetition"). During the period of observation, the siblings lived in the province of Ontario, Canada. Their main social environments were their home, a daycare center (for the younger child), their Elementary school, different venues for their extracurricular activities, and their grandparents' home in Spain. According to the BLP

[41] Lanza's (2007) five parental discourse strategies refer to the moment when the child uses non-parental language in the whole utterance, not specifically when the child code-switches intra-sententially. While referring to the specific strategies within this family, we have adjusted Lanza's continuum to intra-sentential CS. The parents report no instances in which the children talk to them only in English; this only occurs occasionally when the children talk amongst themselves; at those times, their parents encourage them to speak in Spanish.

(Bilingual Language Profile), both are balanced bilinguals. Their specific linguistic background is as follows:

a. **Child A** (BLP Dominance score: -1.996) was born in the state of Georgia, United States, and was first exposed to English at the daycare when he was six months old. By the age of three, his family moved to Canada where he was exposed to a third language, French, at the age of 4;6. Since 7;4, the child has been exposed to a fourth language, Mandarin, as an extracurricular activity, twice a week in a formal environment. For the two years during which the observation took place, the child attended a French immersion school, where 75% of his school time was in French (25% in English), but the after-school program and all his extracurricular activities (except Mandarin) were in English.

b. **Child B** (BLP Dominance score: -5.084) was born in Ontario, Canada, and was first exposed to English at daycare, when she was nine months old. For the two years under observation, the child was exposed to a third language, French, when she started attending her French immersion school at 3;9, where 100% of her school time was in French, but, as with her brother, the afterschool program and all her extracurricular activities were in English.

It should be noted that the compilation of speech samples over this period of time has not been exhaustive; the corpus does not include every code-switched modal infinitival construction uttered by the children in front of their parents. The number of utterances varies in quantity depending on the month (see Appendix I for the whole list of utterances). Consequently, this difference does not imply an increase or a decrease in the production of code-switched modal infinitival constructions by the children either. As explained in Table 3, during the two years for which the data were gathered, a total of 167 code-switched items involving modal infinitival constructions were recorded from both children. Between the ages of 6;5 to 8;6, there were a total

of 83 items recorded from Child A: 63 correspond to periphrastic constructions (76%) and 20 to nominal infinitival clauses (24%). As for Child B, between the ages of 2;7 and 4;8, her total was 84 items: 57 periphrastic constructions (68%) and 27 nominal infinitival clauses (31%). Among the periphrastic constructions, we found modal (i.e, inflected verbs: '*haber que/tener que/have to*' and 'poder/can') and aspectual (i.e., periphrastic future with '*ir a*/going to' and '*empezar a*/begin to'). Within the nominal infinitival clauses, we found mainly the use of the verb '*querer*/want to', with less frequent appearances of the verbs '*necesitar*/need to' and '*intentar*/try to' (See Table 4 for examples of code-switched utterances).

Table 3

Longitudinal corpus — Total of code-switched items

	Periphrastic constructions					Nominal infinitival constructions		
	Modal			Aspectual		Modal		Other
	Haber que	Poder	Tener que	Ir a	Empezar a	Necesitar	Querer	Intentar
Child A	10	31	1	20	1	8	10	2
Child B	0	46	0	11	0	2	25	0
Total	10	77	1	31	1	10	35	2

Table 4

Longitudinal corpus — Examples of code-switched items

	Periphrastic constructions					Nominal infinitival constructions		
	Modal			Aspectual		Modal		Other
	Haber que	Poder	Tener que	Ir a	Empezar a	Necesitar	Querer	Intentar
Child A	'Hay que *bend it*'	'¿Puedes *press go*?'	'Tengo que *play it*'	'Voy a *move on*'	'Puedes empezar a *catch me*'	'Necesito *sharpen* mi *pencil*'	'Yo quiero *check*'	'Intenta *score*'
Child B	NA	'Podíamos *watch movies*'	NA	'Voy a *change directions*'	NA	'Necesitas *get off the floor*'	'¿Quieres *watch me*?'	NA

In all items, periphrastic and nominal infinitival constructions alike, the directionality of the switch was the same as the one depicted in Table 3: inflected verb in Spanish and infinitive in English. There are no utterances with an inflected verb in English and the infinitive in Spanish. With respect to the presence or absence of the infinitival marker *to* for nominal infinitival constructions, where there is a discrepancy between English and Spanish (i.e., '*necesitar* /need to', '*querer*/want to' and '*intentar*/try to'), there is only one utterance, expressed by Child B at 3;5 years old, that does include the infinitival marker, "Yo no quiero to sleep" (I don't want to sleep). In all utterances but this one the infinitival marker *to* is absent. This pattern is present in both children regardless of their age (English inflected verb + Spanish infinitive). In relation to those constructions where there is a particle between the inflected verb and the infinitive both in Spanish and English (i.e.,

'*haber que/tener que/have to*', '*ir a*/going to', '*empezar* a/start to)', that particle was always present and uttered in Spanish.

As we have just described, there is a pattern of CS in these children in their specific sociolinguistic context (previously described), where Spanish is the language spoken and encouraged to be spoken with their parents (where the data were gathered). However, the data also showed some variability on what occurs in the predicate after the English infinitive. We noticed six types of sentences, in terms of their ways of mixing, whose numbers can be found in Table 5:

Type 1: Borrowing or 'insertional CS' with only one English word (i.e., the infinitive verb) introduced in the sentence:

(i) '*Lo quiero* sharpen' (I want to sharpen it) Child A (7;2)

(ii) '¿*Puedo* play *con las cartas contigo*?' (Can I play cards with you?) Child B (3;7)

Type 2: English infinitive + English complement where the complement, generally a direct object, can be a pronoun (Type 2a) or a DP (Type 2b). It should be noted that, within type 2a, there were four utterances where the clitic pronoun was expressed in Spanish and attached to the English verb. Also, as an exception to the whole corpus, there was also a code-switched utterance of Spanish and French within the Type 2b:

(iii) Type 2a — '¿*Puedo* let it go?' (Can I let it go?) Child A (6;8)

(iv) Type 2a (Spa) — '*Hay que* flat-*lo*' (It must be flattened) Child A (8;6)

(v) Type 2b — '*La pelota no puede* touch the floor' (The ball can't touch the floor) Child A (7;3)

(vi) Type 2b (Fr) — *El lobo quiere* manger des cochons (The wolf wants to eat the pigs) Child B (4;5)

Type 3: Mixed utterances where the English infinitive is followed by a Spanish Determiner and an English Noun:

(vii) '*Yo quiero* climb *ese* tree' (I want to climb that tree) Child A (8;5)

(viii) '*Yo voy a* make *un* rocket ship' (I'm going to make a rocket ship) Child B (3;7)

Type 4: English phrasal verb:

(ix) '*Necesitas* hurry up' (You need to hurry up) Child A (7;10)

(x) ¿*Puedes* back up? (Can you back up?) Child B (4;6)

Type 5: Intra-sentential CS where there are no Spanish words in the second half of the sentence:

(xi) '*Puede* spin all the way around' (It can spin all the way around) Child A (8;0)

(xii) '*Mamá, ¿puedo* sit next to you?' (Mom, can I sit next to you?) Child B (3;7)

Type 6: Other mixed utterances:

(xiii) '*Voy a* be nice' (I'm going to be nice) Child A (8;5)

(xiv) '¿*Puedes* stop messing *mi pelo*?' (Can you stop messing up my hair?) Child B (4;8)

Type 7: A combination of previous types (usually a phrasal verb along with an English pronoun):

(xv) '*Voy a* turn it on' (I'm going to turn it on) Child A (8;6)

(xvi) '*Voy a* clean up my mess *muy rápido*' (I'm going to clean up my mess very quickly) Child B (3;5)

Table 5

Longitudinal corpus — Types of sentences

	Child A		Child B	
	Periphrastic constructions	Nominal infinitival clauses	Periphrastic constructions	Nominal infinitival clauses
Type 1	24	11	16	13
Type 2	23	3	18	8
Type 3	5	4	12	3
Type 4	4	1	4	NA
Type 5	4	NA	2	1
Type 6	2	NA	1	NA
Type 7	1	1	4	2

As we can see in the previous table, the most frequent types of sentences were those that involved an individual English lexical item in the Spanish sentence (Type 1) and those that include English complement clauses along with the English infinitive (Type 2). As we will see in the following section, with the description of the study, the type of experimental items that have been used in the code-switched task are of Type 2b, where the complement is a DP, and we have not introduced the presence of clitics in our study, which could be considered for future research.

If we analyze these preliminary data within the framework of the previously described CS theories, we found that only the GC framework, which would allow switches within infinitival constructions, is in accordance with the available patterns. Contrary to the FHC and the PFR which argue for the impossibility of switches between lexical and functional categories, the data, in line with previous child related literature, show that switches between a modal auxiliary and an

infinitive verb are possible. When we compare this pattern with the one in MacSwan's MA theory, different conclusions are reached. According to MacSwan, only the switches in (70), characterized by an English inflected verb followed by a Spanish infinitive, are allowed, disavowing switches such as (73), precisely the pattern shown by these two children[42]:

(70) He wants **to** *hacer la cena*

'He wants to make dinner.'

(73) **Quiere* make dinner

'He wants to make dinner.'

The generalizations that we can make on the basis of this small data set are yet to be tested. The differences and similarities of the patterns shown by these children when comparing them with previous linguistic CS theories are simply a first step towards shedding more light on the various conditions that determine what occurs when two languages are in contact and expressed in a mixed utterance. These results could be attributed to the specific sociolinguistic environment of these children, along with their specific linguistic background, which would imply that not only should the specific characteristics of the grammars to be considered when determining what is allowed in CS but also the particularities of the bilingual speaker. In the following chapter we introduce the experimental study which will provide us with more data to address the issues raised by the different theoretical approaches.

[42] (70) and (73) are repeated examples from page 50.

4. An overview of the study

The purpose of this study is twofold: on the one hand, to determine whether Spanish–English bilinguals with different linguistic backgrounds can distinguish modal periphrases from nominal infinitival clauses in Spanish; on the other hand, to try to shed light on the potential patterns of switches involving Spanish–English modal infinitival constructions (i.e., inflected verb plus infinitive verb): (i) between lexical and functional categories (i.e., *poder*/*deber*/can/must + infinitive verb); (ii) between lexical categories with or without the infinitival marker *to* (i.e., *preferir/desear*/prefer/wish + (to) + infinitive verb); and (iii) within the periphrastic future with the preposition *a* in Spanish (‘*ir a’* + infinitive) and the infinitival marker *to* in English (‘going to’ + infinitive verb). This dual purpose will be achieved through the administration of two tasks to the study participants and the subsequent data analysis.

The first task, with Spanish items, will focus on the distinction of modal periphrases from nominal infinitival clauses. The second task will focus on CS patterns and will include Spanish–English code-switched items (i.e., both types of infinitival constructions). With this double aim, in this chapter we first present the research questions with their related hypotheses. Next, we provide a detailed description of the participants focusing on their language background. This description will be followed by the depiction of the experimental tasks and the data collection procedure.

4.1. Research questions and hypotheses

4.1.1. Distinguishing modal infinitival constructions

Taking into consideration the first aim of the study, the first task, a Grammaticality Judgment Task (GJT), seeks to answer the following research questions:

Research question 1: Are the Control group[43] and six experimental groups of Spanish–English bilinguals—L2 students (L2S), Immigrant Speakers (ImS) and Heritage Speakers (HS)—able to differentiate between modal periphrases and nominal infinitival clauses? Do Spanish–English bilinguals perform differently from the Control group and/or between the groups in terms of accuracy?

Hypothesis

Given that there are no previous empirical studies on this matter, the starting point would be the results of the study carried out by Alonso González (2021) with Spanish speakers, born and raised in Spain and whose first and dominant language was Spanish[44]. Those results, in alignment with previous literature (Gómez Torrego, 1999; Hadlich, 1971; RAE-ASELE, 2011), lead us hypothesize that the Control group will be able to differentiate both infinitival constructions, which means that participants in this group are expected to consider a sentence such as (89) ungrammatical while accepting a sentence such as (90):

(89) *El equipo puede que el director apruebe el proyecto

* 'The team can that the director approves the project.'

(90) El turista desea que la habitación tenga más camas.

[43] We will consider the Control group as those Spanish dominant speakers whose first language is Spanish and who were born within a Spanish-speaking family and living in a Spanish-speaking country (Spain).

[44] The aim of this study was to determine whether two groups of Spanish-native speakers, children and adults, were able to differentiate modal periphrases, as in *debo/puedo comer* (I must/can eat), from nominal infinitival clauses, as in *quiero/deseo comer* (I want/wish to eat). Two Forced-Choice Tasks (FCT) were designed for this purpose: a Grammaticality Judgment Task (FCT-1) and a Fill-in-the-Blank Task (FCT-2). The between-subject variable *group* (children and adults), and the within-subject variables *task* (FCT-1 and FCT-2) and *type of sentence* (there were four different verb combinations that can result in grammatical or ungrammatical sentences depending on whether the inflected verb provided is periphrastic or non-periphrastic) were used to determine their possible effect in the differentiation of these constructions. Results showed a positive overall performance, with 75% of accurate answers. However, the statistical analysis revealed a significant difference between groups and tasks: FCT-2 elicited more right answers than FCT-1, especially among adults, and the adults' performance was significantly better than the children's. In relation to the type of sentence, the results showed that this variable did not play a significant role in distinguishing infinitival constructions.

The tourist desires that the room has more beds.

'The tourist would like the room to have more beds.'

With respect to our bilingual participants, we consider that the variety of grammatical constraints that affect these constructions and the apparent absence of visible distinguishable features that account for their lack of saliency together with the semantic complexity conveyed by modality could interfere with the acquisition of these constructions by those groups whose dominant language is English; namely, the Heritage Spanish, the L1Eng–L2Spa Immigrants, and the L1Eng–L2Spa Students. These groups, contrary to the Control Group, the Heritage English, the L1Spa–L2Eng Immigrants, and the L1Spa–L2Eng Students have (presumably) not received explicit instruction aimed at distinguishing both constructions[45]. If quality and quantity of input could condition the degree of differentiation between MPs and NICs, we hypothesize different outcomes within the bilingual groups when compared to the Control Group considering the differences between them (i.e., Language Dominance, late vs early acquisition, and quantity and quality of input).

In addition to the above-mentioned variables, the type of experimental task (Grammaticality Judgement Task) along with the specific characteristics of its items (untimed written (un)grammatical sentences), may yield the use of explicit knowledge (conscious and potentially verbalizable) when judging the grammaticality of the sentence (Ellis, 2005; Gutiérrez, 2013; Vafaee et al., 2017; Zyzik & Sánchez, 2019). This could n egatively affect the performance

[45] In L1, the identification of periphrases and nominal infinitival clauses is explicitly taught at the secondary school level in Spain. In L2 Spanish, the Cervantes Institute curriculum includes both constructions separately. At the B1 level, it provides information on how to identify a periphrasis but no specific contrast between the two constructions is provided. Instructional books of Spanish as a Second Language not only lack information about the explicit contrast between these constructions but also leave unexplored the identification of periphrases except for Gómez Torrego's *Gramática didáctica del español* (2011).

of those groups of bilinguals whose dominant language is English and who have acquired Spanish in a naturalistic setting; namely, the Heritage Spanish and the L1Eng–L2Spa Immigrant speakers. These groups are expected to have greater implicit (intuitive and not verbalizable) knowledge than explicit knowledge of the Spanish grammatical structures tested in this task (Zyzik & Sánchez, 2019).

Research question 2: Do the grammaticality of the sentence and the syntactic processes play a role in the differentiation of these two infinitival constructions?

Hypothesis

Following Alonso González's (2021) main conclusions, we hypothesize that bilingual speakers will be able to differentiate the two infinitival constructions regardless of the syntactic processes (infinitival pronominalization, either with an interrogative or a demonstrative pronoun, and "that" sentence).[46] With respect to the grammaticality of the sentence, previous literature has shown that grammatical items produce higher levels of accuracy than ungrammatical items for both monolingual and bilingual speakers (Bialystok, E. 1979, 1986; Ellis, 1991; Godfroid et al., 2015; Gutiérrez, 2013; Hedgcock, 1993; Kim & Nam, 2017; Loewen, 2009; Murphy, 1997; Shiu et al., 2018; Vafaee et al., 2017; Zyzik & Sánchez, 2019); thus, we expect more accurate results on grammatical items compared to ungrammatical items.

4.1.2. Code-switching patterns

With regards to the second aim of the research, the second task (Two-alternative Forced-choice Task) seeks to answer the following research questions:

[46] Alonso González's (2021) study included four types of sentences: infinitive pronominalization with an interrogative pronoun, 'that' sentences, cleft sentences, and reflexive passive in plural. From the previous we selected two of the syntactic processes (infinitive pronominalization with an interrogative pronoun and 'that' sentences) because they clearly differentiated both type of infinitival constructions.

Research question 3: Is there a pattern in the way Spanish–English bilinguals with different linguistic backgrounds accept or reject the non-finite marker *to* in switches between a modal periphrasis (modal verb + infinitive), as in "The girl can (to) *escalar la montaña*" or "*La niña puede* (to) keep her dolly"?

Hypothesis

Given that neither English nor Spanish have a non-finite marker *to* (or an equivalent in Spanish) between a modal verb and an infinitive, we expect that participants will omit *to* in those switches occurring within modal periphrases, regardless of the directionality of the change and their different linguistic backgrounds (group), which means that participants will opt for sequences such as (91–92), rather than (93–94):

(91) The partner must *aumentar el negocio*

'The partner must increase the business.'

(92) *La profesora debe* give clear examples

'The teacher must give clear examples.'

(93) The partner must **to** *aumentar el negocio*

*The partner must to increase the business

'The partner must increase the business.'

(94) *La profesora debe* **to** give clear examples

*The teacher must to give clear examples.

'The teacher must give clear examples.'

These preferences would be in accordance with the pattern found in the corpus analyzed in Chapter 3 (See section 3.2.2).

Research question 4: Is there a pattern in the way Spanish–English bilinguals with different linguistic backgrounds accept or reject *to/a* in switches involving the periphrastic future (*Ir a* / Going to + infinitive), as in "I'm going (to) *leer un libro*" or "*Voy (a)* balance my budget"?

Hypothesis

With regards to the periphrastic future, as opposed to the modal periphrases, there is a particle *to/a* between the inflected verb and the infinitive in English and Spanish. Therefore, we expect that participants will include this particle regardless of the directionality of the switch and the group of speakers, a result also noticed in the children's corpus (Section 3.2.2). This would translate in participants opting for (95–96) rather than (97–98):

(95) *Voy a* drink juice

'I'm going to drink juice.'

(96) I'm going to *viajar a Estados Unidos*

'I'm going to travel to the United States.'

(97) *Vamos* visit my parents

*We're going visit my parents.

'We're going to visit my parents.'

(98) We're going *comprar unas flores*

*We're going buy some flowers.

'We're going to buy some flowers.'

Research question 5: Is there a pattern in the way Spanish–English bilinguals with different linguistic backgrounds accept or reject the non-finite marker *to* in switches between a nominal

infinitival clause (lexical modal verb + infinitive), as in "the pupil wishes (to) *mejorar su nota*" or "*el comprador desea* (to) lower the price"?

Hypothesis

We do not expect consistency with the acceptance/rejection of *to* with nominal infinitival constructions, as there is no equivalent in Spanish to the English non-finite marker *to*, which is materialized as the -r suffix in Spanish (Klein, 1985). Taking this into consideration, we can hypothesize different patterns depending on the directionality of the switch. When the inflected verb is in English there would be two scenarios: (i) the inclusion of the infinitival marker *to*, as in (99), which is the only acceptable switch according to MacSwan (1999); or (ii) the exclusion of the infinitival marker *to*, as in (100), given that the Functional Head Constraint rejects "code-switching between functional heads that articulate INFL and their complements" (Toribio, 2001, p. 209) as well as between COMP (complementizer) and its IP complement. When the inflected verb is in Spanish, we hypothesize that our participants will follow the pattern which occurs in our corpus data; namely, the preferred switch would not display the non-finite marker *to*, as in (101):

(99) The worker prefers **to** *salir del edificio*

'The worker prefers to leave the building.'

(100) The worker prefers *salir del edificio*

The worker prefers to leave the building.

'The worker prefers to leave the building.'

(101) *El científico prefiere* clean the laboratory

The scientist prefers to clean the laboratory.

'The scientist prefers to clean the laboratory.'

Research question 6: Do L2 student (L2S), immigrant speakers (ImS), and heritage speakers (HS) perform differently from speakers in our longitudinal case study?

Hypothesis

As we described in the previous chapter (Section 3.2.2.), there is debate over considering switches made by young children as admissible when determining what is allowed or not in adult CS; however, the age of the children in our corpus surpasses this early stage of CS[47], so we do not expect a difference in production with regards to modal periphrases and the periphrastic future; namely, we expect all groups will reject the infinitival marker *to* in modal periphrases and will accept it with the periphrastic future. With respect to the nominal infinitival constructions, our corpus only contains data where the inflected verb is in Spanish and the infinitive in English, which implies that we expect a preference for not including the English infinitival marker *to*, as in (101). When the directionality is in English, we expect our participants to choose those sentences with the English infinitival marker *to*, in accordance with MacSwan (1999). Table 6 shows a summary of the hypotheses formulated.

Table 6

Task 2—Summary of hypotheses

Type of sentence	**Directionality**	
	English	**Spanish**
MP	Without *to* The partner must *aumentar el negocio*	Without *to* *La profesora debe* give clear examples
NIC	With *to* The worker prefers **to** *salir del edificio*	Without *to* *El científico prefiere* clean the laboratory

[47] The earliest recorded switch in the corpus was noted at 2;7.

PF	With *to*	With *a*
	I'm going **to** *viajar a Estados Unidos*	*Voy* ***a*** drink juice

Note 1. MP = Modal Periphrasis; NIC = Nominal infinitival clauses; PF = Periphrastic future; Eng = English; Spa = Spanish

Note 2. The conclusions gathered from the longitudinal corpus coincide with the hypotheses when the directionality is in Spanish; there were no data with directionality in English.

4.2. Participants

A total of 96 adult participants were recruited for this study. These participants belonged to three different pairs of bilingual speakers: immigrants, students, and heritage speakers. The first pair comprised 14 Immigrant speakers with L1 Spanish and L2 English (L1Spa–L2Eng Immigrants) and 10 Immigrant speakers with L1 English and L2 Spanish (L1Eng–L2Spa Immigrants). The L1Spa–L2Eng Immigrants were recruited in English-speaking countries, mostly Canada and United States, although there were some residents of the United Kingdom and one participant residing in Spain[48]. Their average LOR in an English-speaking country was 15 years (*SD* = 4.8). The L1Eng–L2Spa Immigrants were recruited in Spanish-speaking countries, mainly Spain, with the exception of one participant residing in Bolivia. These participants' average LOR in a Spanish-speaking country was 15 years (*SD* = 6). In both groups, the LOR of most participants was above 10 years.

The second type of bilinguals that participated in the study were L2 students in the context of foreign language learning, which meant their L1 was the dominant language (as opposed to a context of L2 immersion, where their L1 is the minority language)[49]. In particular, 22 students

[48] Although the residency of this participant was Spain (not an English-speaking country for an Immigrant participant) at the time of testing, this participant was a newcomer to Spain (previously residing in the United States) without residency roots in the country yet.

[49] An example of foreign language learning would be leaning English as an L2 in Spain; an example of L2 immersion would be learning English as an L2 in the UK.

with L1 Spanish and L2 English (L1Spa–L2Eng Students) and 10 students with L1 English and L2 Spanish (L1Eng–L2Spa Students) participated in the study. Both groups were recruited in academic institutions (language schools and universities) in Spain (for the L1Spa–L2Eng Students), the United States, and Canada (L1Eng–L2Spa Students), which ensured that our participants had a minimum level of B2 in their L2 (English and Spanish, respectively). There was one exception in terms of residence within the L1Eng–L2Spa Student group, this participant resided in Spain when the tasks were administered but had mostly resided in a country where English was the official language[50].

The third group of bilinguals consisted of heritage speakers, more specifically, 10 Spanish heritage speakers (Heritage Spanish) residing in English-speaking countries, such as Canada and the United States, and 10 English heritage speakers (Heritage English) residing in Spain. In both groups, there was one exception in terms of residence, a Heritage Spanish speaker residing in Spain and a Heritage English speaker residing in UK, but their linguistic background helped us to confirm that they grew up and acquired their heritage language in the linguistic environment that characterizes a heritage speaker; namely, these speakers were 2L1, born and raised in a country where the dominant language is not their heritage language: the United States for the Spanish Heritage speaker and Spain for the English Heritage speaker. In fact, to ensure that the participants in this group were heritage speakers, we not only took into consideration their place of residence but also the first language(s) they were exposed to and their LOR in the country where the dominant language differed from their heritage language (English in Spain, and Spanish in Canada and the United States).

[50] These two exceptions, the participants within the L1Spa–L2Eng Immigrants group and the L1Eng–L2Spa Student group, respectively, were tested as potential outliers with no substantial variation in the results.

Finally, and only for the first task, we also recruited a Control Group (CG) composed of 20 native Spanish speakers born, raised, and residing in Spain. The absence of the CG in the second task was due to the bilingual nature of the items used for this experimental task, far from the scope of the Spanish dominant speakers of the CG. Although 96 people were recruited for the experimental tasks, seven more people participated in the study but were discarded for the following reasons: (i) two English Heritage Speakers were eliminated because the AO for Spanish placed them as late L2 students (older than 6 year old)[51], as a consequence, both participants were English dominant, instead of Spanish dominant or balanced bilinguals, as expected for a Heritage English speaker; (ii) one participant was excluded from the group of L1Spa–L2Eng Students because the data obtained from the Bilingual Language Profile (BLP: Birdsong, Gertken & Amengual, 2012) was contradictory; (iii) one participant was excluded from the group of L1Eng–L2Spa Immigrants because the first language stated on the BLP was neither English nor Spanish; (iv) and three participants were excluded from the group of L1Eng–L2Spa Students for different reasons, the first one because the first language stated by the participant on the BLP was neither English nor Spanish, the second one because the participant did not complete the tasks, and the third one because they specifically asked to be removed from the study after performing the tasks.

The specific biographical and linguistic characteristics of the 96 participants are detailed in the following section where we introduce the results provided by the BLP along with some extra questions that were added to this questionnaire in order to get a more complete picture of our participants' sociolinguistic background. Considering the different bilingual profiles that we needed to take into consideration during the recruiting process, this combination of both (BLP +

[51] See section 3.1.1. on 'Types of bilinguals'

extra questions) allowed us to place participants into their corresponding group with a higher level of accuracy.

4.2.1. Bilingual Language Profile

The Bilingual Language Profile (Birdsong, Gertken & Amengual, 2012) is an instrument for assessing Language Dominance (LD) through several variables such as language history, proficiency, use, and attitudes towards a given pair of languages. The four modules, equally measured, consist of 19 multiple-choice responses, all scalar. Developed by the Center for Open Educational Resources and Language Learning (COERLL) from University of Texas, this questionnaire is available for several pair of languages. For this study, the Spanish–English and the English–Spanish versions were selected. The form was adapted to and administered through Survey Monkey. To determine the language dominance index, the score varies between -218 (Spanish) and +218 (English), with scores between -30 and +30 corresponding to balanced bilingualism (Amengual & Simonet, 2020).

More than half of our participants were Spanish dominant (58%), meaning that the number of participants for the various groups was not even, as the recruitment process yielded more participants for the CG and the L1Spa–L2Eng Students group, both characterized by being Spanish dominant. Despite the weight of Spanish dominant speakers, the LD mean of all participants was -38.265 (*SD* = 91.06), close to balanced bilingualism. In the second task, without the CG, the LD mean of the 76 participants was -4.93, balanced bilingualism, (*SD* = 69.713). Table 2 shows a summary of the data related to the participants' LD.

Table 7

Language Dominance — Experimental Groups

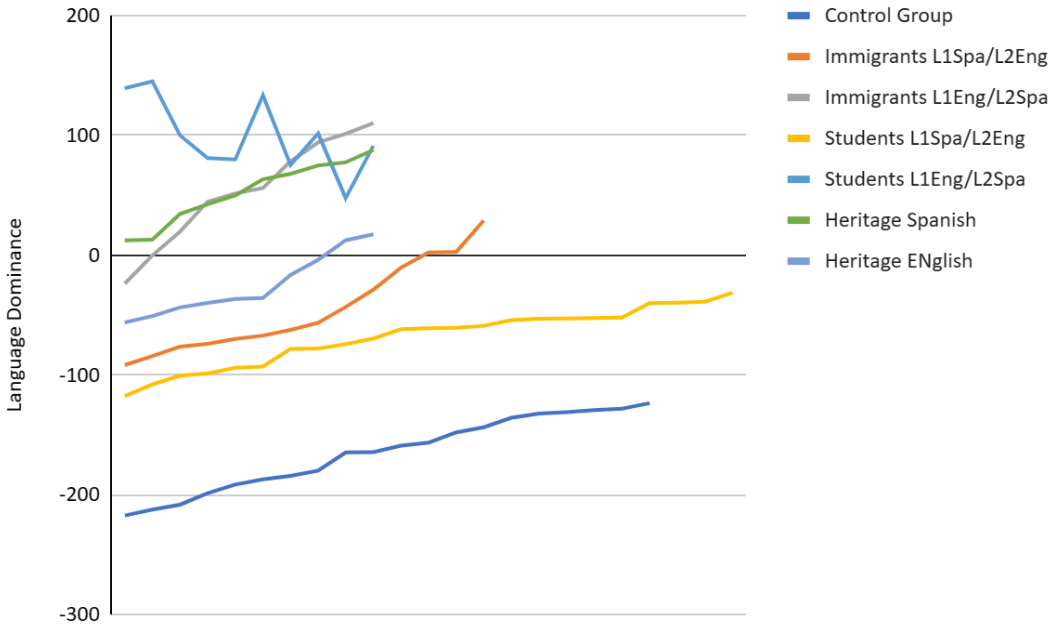
Group	Participants	LD: Mean	LD: SD	LD: #/%
Control	20	-164.9	30.8	20 / 100% (SpD)
L1 Spa–L2 Eng Immigrants	14	-45.1	37.7	5 / 36% (B) 9 / 64% (SpD)
L1 Eng–L2 Spa Immigrants	10	53.6	44.6	3 / 30% (B) 7 / 70% (ED)
L1 Spa–L2 Eng Students	22	-69.9	23.2	22 / 100% (SpD)
L1 Eng–L2 Spa Students	10	99.4	31.5	10 / 100% (ED)
Heritage Spanish	10	52.3	26.5	2 / 20% (B) 8 / 80% (ED)
Heritage English	10	-25.4	26.2	5 / 50% (B) 5 / 50% (SpD)
Task 1	96	-38.265	91.06	15 / 16% (B) 25 / 26% (ED) 56 / 58% (SpD)
Task 2	76	-4.93	69.71	15 / 20% (B) 25 / 33% (ED) 36 / 47% (SpD)

Note. LD = Language Dominance; SpD = Spanish Dominant; ED = English Dominant; B = Balanced; # = number of participants; % = percentage of participants

In addition to these data, Figure 3 depicts the distribution of LD values of each experimental group. Except for the CG values (*M* = -164.9 LD, *SD* = 30.8, range of LD -217.486 / -123.684), most LD values are within the -100/+100 range. Without the CG, the group of participants with higher Spanish dominance is the L1Spa–L2Eng Students group (*M* = -69.9, *SD* = 23.2, range of LD -117.786 / -31.41), followed by the L1Spa–L2Eng Immigrant group (*M* = -45.1, *SD* = 37.7, range of LD -91.724 / 28.88), and the Heritage English group, with proportionally more balanced bilinguals than any other group, (*M* = -25.4, *SD* = 26.2, range of LD -56.218 / 17.342). In the English dominant side, both Heritage Spanish (*M* = 52.3, *SD* = 26.5, range of LD 12.348 / 87.822) and L1Eng–L2Spa Immigrants (*M* = 53.6, *SD* = 44.6, range of LD -23.796 / 110.244) present similar numbers, and the L1Eng–L2Spa Students (*M* = 99.4, *SD* = 31.5, range of LD 47.496 / 145.026) are the group with the highest English dominance.

Figure 3

Language Dominance scores — Experimental Groups



Note. Distribution of LD values of each experimental group

After this descriptive overview of the LD data that we will include as a variable in our data analysis, another section to consider within the BLP is the Biographical section. This section gives the age, gender, current country of residence and level of studies of our participants. The mean age of the participants was 41.9 (*SD* = 16.5, age range 18–80), the younger bilinguals belonged to the L1Spa–L2Eng Students (*M* = 27.5 years old, *SD* = 15, age range 18–52) and the Heritage Spanish (*M* = 32.2 years old, *SD* = 10.7, age range 20–49) groups, followed by the Heritage English group (*M* = 40.8 years old, *SD* = 15.5, age range 19–70) and the L1Eng–L2Spa Students (*M* = 43.5 years old, *SD* = 17.9, age range 27–80). With similar means we find both groups of Immigrants, L1Spa–L2Eng (*M* = 50.4 years old, *SD* = 15, age range 33–78) and L1Eng–L2Spa (*M* = 50.6 years old, *SD* = 13.6, age range 30–78), and the Control Group (*M* = 50.5 years old, *SD* = 13.6, age range 32–76)[52].

In terms of gender, most of our participants were women (61.5%), followed by men (36.5%) and a small number of participants (2) who stated other genders in the BLP. With regards to the current place of residence, most participants reside in Spain (66%) given that the two most numerous groups recruited for this research (Control and L2S with L1 Spanish) live in Spain; these are followed by 16% residing in United States and 14% residing in Canada. There were a small percentage of participants currently residing in the UK (3%) and Bolivia (1%). In terms of education, three participants (3.1%) finished Secondary school, four participants (4.2%) have High School education, 24 participants (25%) were in their first years of university/college, 29 participants (30.2%) had a university/college degree, 26 participants (27.1%) had a Master's degree, and 10 participants (10.4%) had a PhD. A summary of the previous information can be seen in Table 8.

[52] Given the variability in the age of the participants, we checked the potential impact of this variability on the dependent variable for both tasks with the result that it was not significant as a predictor.

Table 8

Bilingual Language Profile — Experimental Groups

Group	Age	Gender	Current place of residence	Highest level of formal education
Control	Mean: 50.5 SD: 13.6	M: 7 (35%) F: 13 (65%)	Spain: 20 (100%)	Secondary School: 0 (0%) High School: 2 (10%) 1st years of university/college: 4 (20%) University/college: 8 (40%) Master: 4 (20%) PhD: 2 (10%)
L1 Spa–L2 Eng Immigrants	Mean: 50.4 SD: 15	M: 9 (64%) F: 5 (36%)	Canada : 6 (43%) Spain : 1 (7%) UK : 2 (14%) USA : 5 (36%)	Secondary School: 0 (0%) High School: 0 (0%) 1st years of university/college: 0 (0%) University/college: 6 (43%) Master: 5 (36%) PhD: 3 (21%)
L1 Eng–L2 Spa Immigrants	Mean: 50.6 SD: 13.6	M: 5 (50%) F: 5 (50%)	Bolivia : 1 (10%) Spain : 9 (90%)	Secondary School: 0 (0%) High School: 1 (10%) 1st years of university/college: 0 (0%) University/college: 6 (60%) Master: 3 (30%) PhD: 0 (0%)
L1 Spa–L2 Eng Students	Mean: 27.5 SD: 9.6	M: 6 (27%) F: 16 (73%)	Spain: 22 (100%)	Secondary School: 1 (4.5%) High School: 0 (0%) 1st years of university/college: 14 (64%) University/college: 4 (18%) Master: 2 (9%) PhD: 1 (4.5%)
L1 Eng–L2 Spa Students	Mean: 43.5 SD: 17.9	M: 1 (10%) F: 7 (70%) O: 2 (20%)	Canada: 2 (20%) Spain: 1 (10%) USA: 7 (70%)	Secondary School: 0 (0%) High School: 0 (0%) 1st years of university/college: 2 (20%) University/college: 0 (0%)

				Master: 5 (50%) PhD: 3 (30%)
Heritage Spanish	Mean: 32.2 SD: 10.7	M: 3 (30%) F: 7 (70%)	Canada: 6 (60%) USA: 3 (70%) Spain: 1 (1%)	Secondary School: 2 (20%) High School: 0 (0%) 1st years of university/college: 3 (30%) University/college: 2 (20%) Master: 3 (30%) PhD: 0 (0%)
Heritage English	Mean: 40.8 SD: 15.5	M: 4 (40%) F: 6 (60%)	Spain: 9 (90%) UK: 1 (1%)	Secondary School: 0 (0%) High School: 1 (10%) 1st years of university/college: 1 (10%) University/college: 3 (30%) Master: 4 (40%) PhD: 1 (10%)
Totals	Mean: 41.6 SD: 16.5	M: 35 (36.5%) F: 59 (61.5%) O: 2 (2%)	Bolivia: 1 (1%) Canada: 14 (14%) Spain: 63 (66%) UK: 3 (3%) USA: 15 (16%)	Secondary School: 3 (3.1%) High School: 4 (4.2%) 1st years of university/college: 24 (25%) University/college: 29 (30.2%) Master: 26 (27.1%) PhD: 10 (10.4%)

Note. M=Male; F=Female; O=Other

In order to better understand our participants' sociolinguistic background, we added seven questions to the Biographical section of the original BLP. The first three questions elicited information about: (i) the country of birth, (ii) the home languages (until 16 years old of age), and (iii) the first language the participant was exposed to. With regards to the country of birth, our participants were born mainly in Spain (61%), followed by the English-speaking countries: the United States (16%), the UK (7%), and Canada (5%). There were other countries of origin, such as the Spanish-speaking Latin American countries Colombia (2%), Ecuador (1%), and Peru (1%),

along with other countries that have other official languages in addition to English, such as the Philippines (1%) and Zimbabwe (1%), and one last country, Romania (1%), whose official language is Romanian[53].

Regarding the questions about the first language(s) our participants were exposed to and the languages spoken at home, we found an equivalence in the CG and the L1Spa–L2Eng Immigrant group: only Spanish in the CG, and Spanish and 2L1 (2 first languages, Spanish and English, at birth) in the L1Spa–L2Eng Immigrant group. With respect to the first language(s), there are 62 participants (64.5%) whose first language is Spanish, 22 whose first language is English (23%) and 12 2L1 (12.5%) whose first language is English, too. In terms of the languages spoken at home, except for the Control group, all were bilinguals but for a small percentage of multilingual participants (3%). The original intention of this question was to get a better picture of all the languages our participants had been exposed to until the age of 16, beyond the critical period, but without further details on the acquisition of these languages (i.e., quantity and quality of input in this particular language or source of the language), along with the inconsistency between those languages and the first language(s) in most of the groups, we only considered the data pertaining the first language(s) to classify our participants in their respective groups. A summary of this information can be found in Table 9.

Table 9

Additional sociolinguistic data — Experimental Groups

Group	Country of Birth	First language(s)	Languages at home
Control	Spain: 20 (100%)	Spanish: 20 (100%)	Spanish: 20 (100%)

[53] For this specific case where English was not an official language in the participant's country of origin, English was noted as the participant's first language, and it was one of the participant's languages at home.

L1 Spa–L2 Eng Immigrants	Colombia: 2 (14.3%) Ecuador: 1 (7.2%) Perú: 1 (7.2%) Spain: 9 (64.1%) Venezuela: 1 (7.2%)	Spanish: 13 (93%) Spanish+English: 1 (7%)	Spanish: 13 (93%) Spanish+English: 1 (7%)
L1 Eng–L2 Spa Immigrants	Australia: 1 (10%) Romania: 1 (10%) UK: 6 (60%) US: 1 (10%) Zimbabwe: (10%)	English: 10 (100%)	English: 8 (80%) English+Spanish+French: 1% (10%) English+French+Romanian: 1% (10%)
L1 Spa–L2 Eng Students	Spain: 22 (100%)	Spanish: 20 (91%) Spanish+English: 2 (9%)	Spanish: 17 (77%) Spanish + English: 4 (18%) Spanish+Bulgarian: 1 (5%)
L1 Eng–L2 Spa Students	Canada: 1 (1%) Philippines: 1 (1%) USA: 8 (80%)	English: 10 (100%)	English: 8 (80%) Tagalog: 1 (10%) English + Italian: 1 (10%)
Heritage Spanish	Canada: 4 (40%) Spain: 1 (1%) USA: 4 (40%) Venezuela: 1 (1%)	Spanish: 6 (60%) English: 1 (10%) Spanish+English: 3 (70%)	Spanish: 2 (20%) Spanish+English: 8 (80%)
Heritage English	Spain: 7 (70%) UK: 1 (10%) USA: 2 (20%)	Spanish: 3 (30%) English: 1 (10%) Spanish+English: 6 (70%)	English: 1 (10%) English+Spanish: 8 (80%) English+Spanish+Galician: 1 (10%)
Totals	Australia: 1 (1%) Canada: 5 (5%) Colombia: 2 (2%) Ecuador: 1 (1%) Peru: 1 (1%) Philippines: 1 (1%) Romania: 1 (1%) Spain: 59 (61%) UK: 7 (7%) USA: 15 (16%) Venezuela: 2 (2%) Zimbabwe: 1 (1%)	Spanish: 62 (64.5%) English: 22 (23%) Spanish+English: 12 (12.5%)	Spanish: 52 (53%) English: 17 (18%) Spanish+English: 21 (22%) Spanish+other language: 1 (1%) English+other language: 1 (1%) Spanish+English+other language: 2 (2%) English+ 2 other languages: 1 (1%) Other language (Tagalog): 1 (1%)

In addition to the information described, we also asked whether the participant had siblings, the language spoken with siblings, the preferred language for reading, and the preferred language for audiovisual content. However, we ended up discarding these extra data because they did not add any information decisive in placing our participants into their corresponding groups. Table 10 provides a final summary of our participants.

Table 10

Summary of participants

Group	**Number of participants**	
	Task 1	**Task 2**
Control Group	20	NA
L1 Spa–L2 Eng Immigrants	14	14
L1 Eng–L2 Spa Immigrants	10	10
L1 Spa–L2 Eng Students	22	22
L1 Eng–L2 Spa Students	10	10
Heritage Spanish	10	10
Heritage English	10	10
Total of participants	96	76

4.3. Experimental Tasks

4.3.1. Spanish Grammaticality Judgment Task

The first part of the experimental study was centered on determining whether different types of bilinguals can differentiate between modal periphrases and nominal infinitival clauses. More specifically, we aimed to identify the potential influence that the type of bilingual, the language dominance, the grammaticality of the sentence, and the type of syntactic construction may have in the accurate differentiation of these constructions. For this purpose, we designed a

Grammaticality Judgment Task[54] to allow us to confront the essential syntactic differences between two superficially similar constructions. This type of data is quite common inside the generative syntax tradition (Sprouse, 2020) and has been broadly used in various language and population types (see Kail et al., 2012, for more details).

Participants were asked to judge untimed written sentences, 24 grammatical and 24 ungrammatical, by selecting one of two values (Good/Bad), expressed through an emoticon face (happy/sad[55]). These items introduced three different verb sequences in Spanish (infinitival constructions) headed by the modal verbs *poder* and *deber* and the lexical verbs *preferir* and *desear*. The three types of infinitival constructions came from the list of syntactic tests mentioned in Chapter 2 to identify a periphrasis and distinguish it from nominal infinitival clauses (García Fernández, 2006; Gómez Torrego, 1999; Topor, 2005), in particular, the pronominalization of the infinitive into an interrogative, the pronominalization of the infinitive into a demonstrative, and 'that' sentences. Those sentences whose inflected verbs were the auxiliaries *poder* and *deber* were the ungrammatical items, whereas those sentences whose inflected verb were the lexical verbs *preferir* and *desear* were the grammatical items. In addition, we included 24 distractors that introduced the auxiliary verbs *ser* (passive voice), *estar* (present continuous), and *haber* (present perfect) wrongly or accurately combined with clitics. The 48 experimental items and 24 distractors were randomly presented. Eight practice trials with corrective feedback were included at the beginning of the task. All questions were answered, so there were no missing data. See Table 11 for a sample of experimental items and distractors.

[54] On the difference of acceptability and grammaticality judgments see discussion in Cowart (1997), Goodall (2021), Zyzik & Sánchez (2018); A. o.
[55] Images from https//icons8.com

Table 11

Task 1 sample of experimental items and distractors

		Experimental items	
Type of sentence	# Items	Grammatical	Ungrammatical
Pronominalization of the infinitive with an interrogative (*Qué*)	8 + 8	¿Qué prefiere el arquitecto?	*¿Qué puede tu primo?
		¿Qué desea la rana?	* ¿Qué debe la mosca?
Pronominalization of the infinitive with a demonstrative (*Eso*)	8 + 8	Regalar su ropa, eso prefiere el músico.	*Conseguir un trabajo, eso puede el novelista.
		Cambiar de casa, eso desea mi hija.	*Centrar la atención, eso debe el empresario.
'That' sentences	8 + 8	Mi compañera prefiere que su amiga vea la película.	*Su padre puede que mi amigo lea el libro.
		El cliente desea que mi jefe tenga buenas ideas.	*La familia debe que mi abuela salga a pasear.
		Distractors	
Perfect tense	4 + 4	Las patatas, las he comido con gusto.	*El pájaro, he vístolo por la ventana.
Passive Voice	4 + 4	El edificio fue diseñado por mi hermana.	*El cuadro lo fue pintado por un artista.
Present continuous	4 + 4	El agua, la estoy bebiendo.	*La araña, la estoy viéndola.

When creating the experimental items, we took into consideration the number of words in every group of items per condition, in such a way that all experimental items had the same number of

words within each type of sentence. This did not apply to the distractors. The selection of words was based on Davies's frequency dictionary of Spanish core vocabulary for leaners (2006).

4.3.2. Code-switched Two-alternative Forced-Choice Task

The second part of the experimental study in this research was centered on determining whether there is a pattern of exclusion or inclusion of the non-finite marker *to* in code-switched infinitival constructions, such as modal periphrases and nominal infinitival constructions on the one hand, and on the exclusion or inclusion of the preposition *to/a* within the code-switched periphrastic future on the other. In addition, the research also intended to determine whether this potential pattern is conditioned by predictors such as the type of construction, the directionality of the switch (the language of the inflected verb), and the group of bilingual speakers. As stated by MacSwan (2014a), the use of experimental data to gather conclusions on CS has been both praised (Toribio, 2001) as well as rejected for those who only consider naturalistic data as a valid source (Mahootian & Santorini, 1996). In our study, we have opted for combining both types of data—the corpus introduced in Chapter 3 and a forced-choice task—an approach also welcomed in the field of CS research (MacSwan & McAlister, 2010).

In this second experimental task, participants were presented with 48 pairs of sentences (see Table 12 for a sample of experimental items) and had to choose which sentence was more acceptable.[56] Similar to Task 1, there was no time limit to perform this task. Each experimental item consisted of two identical code-switched infinitival constructions which only differed with respect to the absence or presence of the non-finite marker *to* between the inflected verb and the infinitive, the point of the sentence where the Spanish–English switch took place.

[56] We chose this type of task to determine the pattern on CS items because it is considered to have more statistical power than other acceptability judgment experiments, such as Likert tasks (Gigerenzer & Richter, 1990; Gigerenzer et al., 2004; Sprouse & Almeida, 2017; Stadthagen-González et al., 2018).

There were three types of infinitival constructions as experimental items: 16 modal periphrases (8 with the Spanish verbs *deber* and *poder* and 8 with the English verbs *must* and *can*), 16 nominal infinitival clauses (8 with the Spanish verbs *preferir* and *desear* and 8 with the English verbs *prefer* and *wish*), and 16 aspectual periphrases with the periphrastic future (8 with the Spanish construction *ir a* and 8 with the English construction *going to*). For the periphrastic future, the difference between the two pairs consisted of the inclusion/exclusion of the preposition *to/a*. Eight practice trials with corrective feedback were included at the beginning of the task.

The CS always occurred in the same place in the sentence, between the inflected verb and the infinitive, regardless of the directionality of the switch and the type of construction, a different approach to other studies that have used this type of forced-choice task when analysing CS preferences; in those cases, the contrast is between sentences where the code-switch happens at different points in the sentence (Stadthagen-González et al., 2018; Toribio, 2001). The divergence from these studies is because we are not focused on the acceptability of the switch, but on discerning the pattern of the switch that we have encountered in our naturalistic corpus data (and other studies) in relation to certain predictors, such as the type of bilingual (heritage, immigrant, or L2 student), the directionality of the switch (Spanish or English) and the type of infinitival construction (modal periphrasis, nominal infinitival clause, or periphrastic future).

Originally, the task was going to include the periphrastic future as a distractor. However, we ended up including the periphrastic future as a third type of construction because in both English and Spanish we need a particle between the inflected verb and the infinitive. In this way we confronted our participants with three types of constructions with three different linguistic scenarios: the absence of the particle within the modal periphrases, the presence of the particle within the periphrastic future, and the presence of the particle in English—the non-finite marker

to for the verbs *prefer* and *wish*—along with the absence of the particle in the Spanish verbs *preferir* and *desear*. The distractors were eight pairs of sentences with the inclusion and the exclusion of the Spanish definite article *el/la/los/las* preceding both plural and uncountable nouns, as in "(los) elephants are big" or "(la) truth is essential". The 48 experimental items and eight distractors were randomly presented. There were no missing data as participants answered all questions.

Table 12

Task 2 sample of experimental items

		Experimental items	
Type of sentence	# Items	Inflected verb (Eng) Infinitive (Spa)	Inflected verb (Spa) Infinitive (Eng)
Modal Periphrases	8 + 8	The partner must *aumentar el negocio* The partner must **to** *aumentar el negocio*	*La profesora debe* give clear examples *La profesora debe* **to** give clear examples
Nominal Infinitival Clauses	8 + 8	The worker prefers *salir del edificio* The worker prefers **to** *salir del edificio*	*El pescador prefiere* sell his fish *El pescador prefiere* **to** sell his fish
Periphrastic future	8 + 8	We´re going *comprar unas flores* We´re going **to** *comprar unas flores*	*Voy* drink juice *Voy* **a** drink juice

As in the case of the first task, we took into consideration the number of words in every set of items per condition, in such a way that all experimental items had the same number of words within each construction except for the periphrastic future that, initially, was going to be considered a distractor. More specifically, the difference within the experimental items of the periphrastic future was a maximum of two extra words compared to the smaller sentences. The selection of Spanish and English words was based on Davies's frequency dictionary of Spanish core vocabulary for leaners (2006) and Leech et al.'s (2014) corpus of most frequent English words.

4.4. Data collection procedure

The participants performed both tasks, along with the BLP, individually using the online survey tool Survey Monkey (www.surveymonkey.com) at their own leisure in their respective countries of residence (Spain, Bolivia, the United States, Canada, and the UK). Prior to the BLP and the two tasks, participants signed a consent form in their L1. For this purpose, we created two versions of the experiment, one with the consent form and the instructions in English and one in Spanish. As has been mentioned in the description of the tasks, the experimental items in Task 1 were in Spanish, and the experimental items in Task 2 were English–Spanish code-switched items. Given that the reading and signing of the consent form, the BLP, and the tasks were done in one session, participants were encouraged to take a rest between the tasks. It is recommended that online questionnaires take a maximum of 30 minutes to complete (Dörnyei & Taguchi, 2010); ours was designed to last approximately 30 minutes for bilingual participants and 20 minutes for the Control Group, as shown in Table 13.

Table 13

Experimental session for data collection

Control Group	Bilingual groups
1. Reading and signing of consent form (3 minutes) 2. Bilingual Language Questionnaire (7 minutes) 3. Task 1 — Grammaticality Judgment Task (10 minutes)	1. Reading and signing of consent form (3 minutes) 2. Bilingual Language Questionnaire (7 minutes) 3. Task 1 — Grammaticality Judgment Task (10 minutes) 4. Task 2 — Two-alternative Forced-Choice Task (10 minutes)
Approximate time: 20 minutes	Approximate time: 30 minutes

To encourage participation, we created a raffle under the following conditions: each participant was given a number for each finished task. At the end of the data collection process, 10 e-gift cards were sent to their respective winners. Winners were chosen using an online random number generator tool. Finally, to maintain the confidentiality of the participants, each was assigned a group code followed by a number that designated the chronological order in which the participant took the test. The study met the ethical standards of the University of Ottawa Research Ethics Board and was approved under the certificate number S-09-22-7930.

4.5. Data analysis overview

In order to determine, first, whether English–Spanish bilinguals were able to differentiate modal periphrases from nominal infinitival constructions, and, second, whether there is a pattern in the exclusion or inclusion of the English infinitival marker *to* in code-switched modal infinitival constructions, we carried out a mixed-effects logistic regression of both tasks due to the fact that they had binary categorical dependent variables: in the case of Task 1, we measured the accuracy

when judging the grammaticality of a set of sentences; for Task 2, participants had to select one sentence within a pair with the only difference being the absence or presence of the English infinitival marker *to*. The existence of categorical binary outcomes with two or more independent variables (predictors) suggests the use of mixed logit models to analyse the data (Jaeger, 2008; Linck & Cunnings, 2015).

For both tasks, analyses were conducted using mixed-effects logit models with crossed random effects for subjects and items. These models were run in R (version 4.3.1: R Core Team, 2023) using the RStudio interface with the *lme4* (Bates et al., 2015) package installed. The *nloptwrap2* optimizer was used to facilitate the model's convergence (*NLopt* package, version 2.7.1: S. G. Johnson, n.d.). We performed two analyses, one for each task. The first analysis, of the Grammaticality judgment task, included the fixed effects of the following categorical predictors and their corresponding interactions: *Grammaticality* (Grammatical, Ungrammatical), *Type of Sentence* (That, Demonstrative, and What), and *Group* (Control, Heritage Spanish, Heritage English, L1Eng–L2Spa Immigrants, L1Spa–L2Eng Immigrants, L1Eng–L2Spa Students, and L1Spa–L2Eng Students). The second analysis, of the Two-Alternative Forced-Choice task, included the following fixed effects (categorical predictors): *Directionality* (English, Spanish), *Type of Construction* (Modal Periphrasis, Periphrastic Future, and Nominal Infinitival Clause), and *Group* (Heritage Spanish, Heritage English, L1Eng–L2Spa Immigrants, L1Spa–L2Eng Immigrants, L1Eng–L2Spa Students, and L1Spa–L2Eng Students). The only continuous independent variable for both tasks, *Language Dominance*, was z-scored to reduce collinearity (de Bruin et al., 2018; Linck & Cunnings, 2015).

In addition, we checked the potential correlation between *Language Dominance* and *Group*, with the result that there was no correlation between them ($r = -6.618 \times 10^{-11}$, $p > .05$).

Despite the process of z-scoring and the lack of correlation between these variables, we ended up discarding *Language Dominance* as a fixed effect for both tasks because, during the process of configuration of the mixed logit models (detailed for each of the tasks in the following chapter), the presence of this variable lead to a lack of convergence, the presence of errors, and deficient models. Once we achieved the best-fitting models for both tasks, we checked the squared GVIF^(1/(2*Df)) values[57] for all the included predictors (Fox, 2014; MsGISRocker, 2020). For the first task, all these values were below 5 (low multicollinearity), except for the Grammaticality[58]. For the second task, all values were below 5. In addition to the inferential statistics already described, we also performed descriptive statistics with the statistical program JASP (version 0.18: JASP Team, 2023).

5. Data modelling and results

This chapter consists of the description of the modelling process for each set of data corresponding to the two experimental tasks, followed by their respective descriptive data and inferential statistical analyses.

5.1. Spanish Grammaticality Judgment Task

5.1.1. Modelling process

For the mixed logit model, we initiated the modelling process following the standard procedure in the field: starting with the maximal random-effects structure (maximal model) of our analysis (Barr et al., 2013; Linck & Cunnings, 2015; Winter, 2019). Our maximal model included random intercepts for participants and for items, along with by-participant random slopes for the

[57] Used for variables with more than one degree of freedom (Fox & Monette, 1992).
[58] A correlation matrix performed on the regression model for task 1 showed no correlation of Grammaticality with any of the predictors or their interactions.

Type of Sentence (What, Demonstrative, and That), *Grammaticality* (Grammatical NICs and Ungrammatical MPs), and *Language Dominance* (z-scored values from the BLP) and their interactions and by-item random slopes for *Group* (Control, L1Esp-L2Spa Immigrants, L1Eng–L2Spa Immigrants, L1Esp-L2Eng Students, L1Esp-L2Eng Students, Heritage Spanish and Heritage English).

Since the maximal model did not converge, we started a simplification process that consisted of the following steps. First, we removed the random slope interactions. When the model did not converge, we proceeded to decorrelate intercepts and slopes, followed by reducing the complexity of the item slope structure and the participant slope structure, in this order (see García-Alcaraz 2021 for a similar approach). Convergence was attained with the by-participant and by-item random intercepts. The next step was to simplify the fixed effects' structure, originally comprising the interaction of four predictors: *Type of Sentence, Grammaticality, Group,* and *Language Dominance*. This process of simplification started by breaking down the interactions between the fixed effects until convergence was attained, checking all possible combinations. The first model that reached convergence was the one that included the sum of the three following two-way interactions: *Type of Sentence*Group* + *Type of Sentence*Grammaticality* + *Group*Grammaticality*. *Language Dominance* was removed as a fixed effect since it prevented the model from converging in every two-way interaction. After attaining convergence, we continued the process of simplifying the fixed effects structure by checking whether removing any two-way interactions improved the model fit[59]; this was not the case. Thus, the best-fitting model that attained convergence was the one that included the aforementioned three two-way interactions. In terms of the model's explanatory power, the marginal R^2 was 31%, which implies

[59] Each new model was compared with the previous one(s) using the Akaike Information Criterion (AIC).

that the three two-way interactions included in the model explained 31% of the variability of our data. However, when the random-effects (by-participant and by-item effects) were added, the explanatory power of the mixed model increased to 49% ($x^2 = 257.92$, Df = 2, $p < 0.001$).

5.1.2. Results

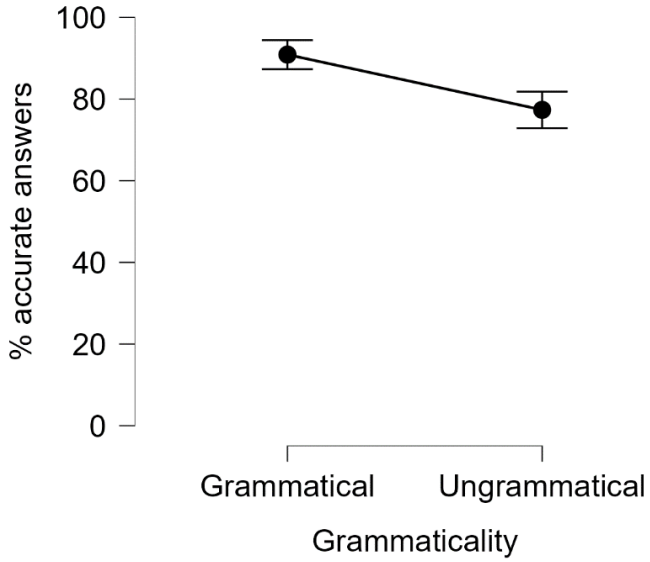
In the following sections, we reported the results from the descriptive data and the inferential statistics for the Grammaticality Judgment Task.

5.1.2.1. Descriptive results

The average percentage of accurate answers for all our participants was 84.1% (*SD* = 27.92). According to the descriptive data, in terms of *Grammaticality* (Figure 4), Grammatical NICs yielded 90.8% accurate answers (*SD* = 21.37), whereas Ungrammatical MPs generated 77.3% accurate answers (*SD* = 31.84). Therefore, it seems that participants, in general, were more accurate when assessing grammatical NICs than ungrammatical MPs.

Figure 4

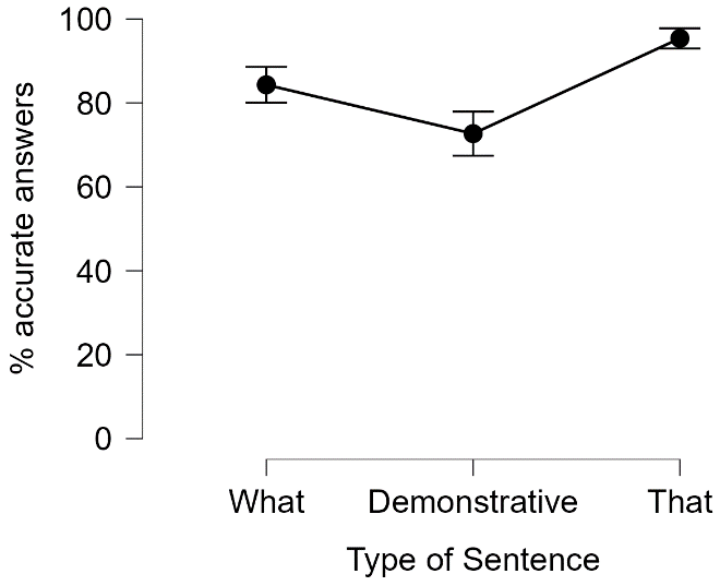
Descriptive Data Task 1 — Grammaticality



Focusing on the *Type of Sentence,* Figure 5 shows that participants obtained the highest accuracy rate, 95.4% (*SD* = 13.91) in 'That' sentences. The accuracy rate in 'What' sentences was 84.3% (*SD* = 27.97) and in 'Demonstrative' sentences was 72.7% (*SD* = 33.36). Thus, it seems that our participants performed better on 'That' sentences when compared to 'What' and 'Demonstrative' sentences.

Figure 5

Descriptive Data Task 1 — Type of Sentence

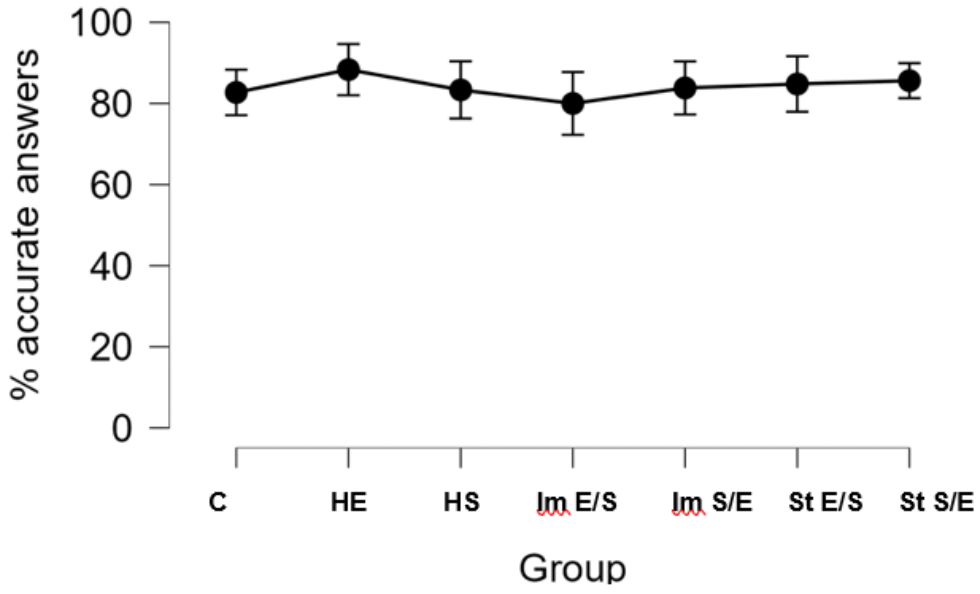


When we compared groups (Figure 6), the Heritage English group displayed the best performance, with 88% accuracy (*SD* = 24.45), followed by the L1Spa–L2Eng Students group, with 85.6% accuracy (*SD* = 25.07), the L1Eng–L2Spa Students group, with 84.7% accuracy (*SD* = 26.45), the L1Spa–L2Eng Immigrants group, with 83.8% accuracy (*SD* = 30.13), the Heritage Spanish group, with 83.3% accuracy (*SD* = 27.29), the Control Group, with 82.7% accuracy (*SD* = 31.05), and the L1Eng–L2Spa Immigrants group, with 80% accuracy (*SD* = 29.93). With a difference of 8% in accuracy between the best performer, the Heritage English group, and the worst performer, the

L1Eng–L2Spa Immigrants, the overall performance of all groups seems to be within a similarly high accuracy range with little differences between the groups.

Figure 6

Descriptive Data Task 1 — Group



Note: C = Control; HE = Heritage English; HS = Heritage Spanish; Im E/S = L1 English–L2 Spanish Immigrant; Im S/E = L1 Spanish–L2 English Immigrant; St E/S = L1 English–L2 Spanish Students; St S/E = L1 Spanish–L2 English Students.

When the interactions were explored, we found that within the interaction of *Type of Sentence* and *Grammaticality* (Table 14), results showed that *Grammaticality* could play a role in eliciting accurate responses in 'Demonstrative' and 'What' sentences, given the difference in accurate responses between their respective grammatical and ungrammatical items. However, 'That' sentences seem to elicit a high number of accurate responses regardless of the *Grammaticality* of the sentence. Therefore, participants seem to be consistent when judging the accuracy of 'That'

sentences in spite of their grammaticality, while ungrammatical 'What' and 'Demonstrative' sentences appeared to be more challenging to judge, especially 'Demonstrative' sentences (see Figure 7).

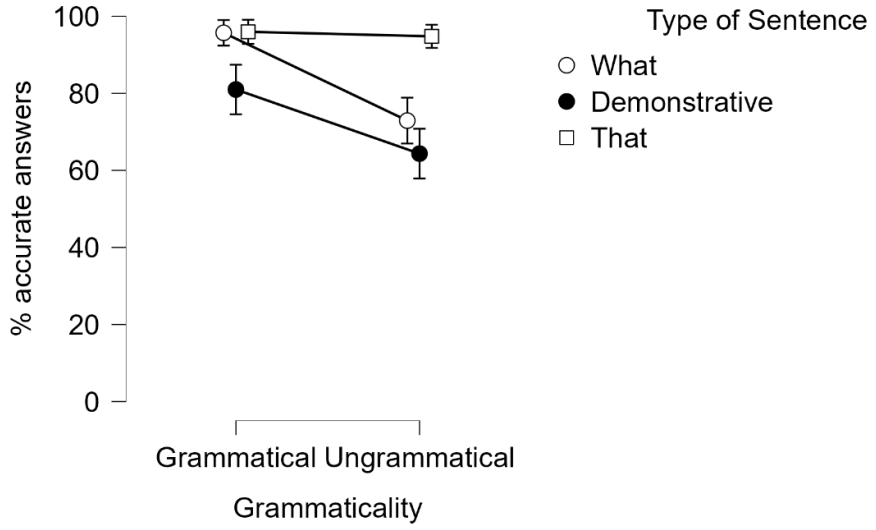
Table 14

*Task 1 — Percentage of accurate responses based on the interaction Grammaticality*Type of Sentence*

Type of Sentence	Grammatical Items	SD	Ungrammatical Items	SD
That	95.9%	12.8	94.8%	15
What	95.7%	14.7	72.9%	33.1
Demonstrative	80.9%	29.2	64.3%	35.3

Figure 7

Descriptive Data Task 1 — Interaction between Type of Sentence and Grammaticality



The interaction of *Grammaticality* and *Group* (Table 15) shows that all groups except the Heritage English participants were more accurate when judging grammatical sentences compared to

ungrammatical ones. The Heritage English group's behavior was the opposite, as they produced a higher number of accurate responses when judging ungrammatical rather than grammatical sentences (See Figure 8).

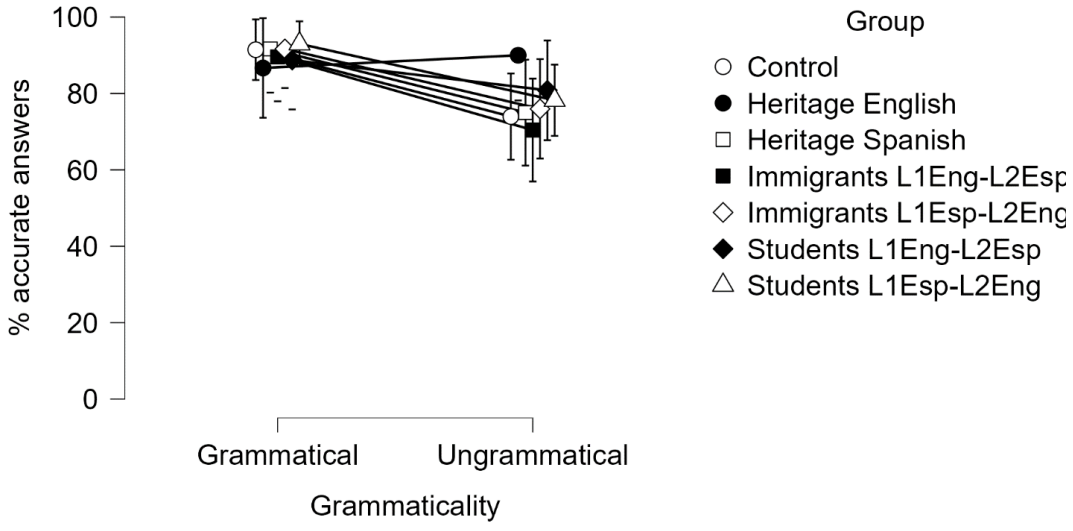
Table 15

*Task 1 — Percentage of accurate responses based on the interaction Grammaticality*Group*

Group	Gramm. items	SD	Ungram. items	SD
Control Group	91.4%	23.4	73.9%	35.2
Heritage English	86.6%	26.0	90%	23.1
Heritage Spanish	91.7%	18.4	75%	32.2
L1 Eng–L2 Spa Immigrants	89.6%	22.0	70.4%	33.9
L1 Spa–L2 Eng Immigrants	91.6%	19.8	75.9%	36.3
L1 Eng–L2 Spa Students	88.7%	25.3	80.8%	27.4
L1 Spa–L2 Eng Students	92.9%	17.2	78.2%	29.3

Figure 8

Descriptive Data Task 1 — Interaction between Group and Grammaticality



Focusing on the interaction between *Type of Sentence* and *Group* (Table 16), we found that the general tendency previously expressed in terms of the *Type of Sentence* (i.e., higher number of accurate answers for 'That' sentences, followed by 'What' sentences, and, finally, by 'Demonstrative' sentences) was consistent among groups, with one exception: the L1Eng–L2Spa Immigrants group, who performed better on 'Demonstrative' items than on 'What' items (see Figure 9).

Table 16

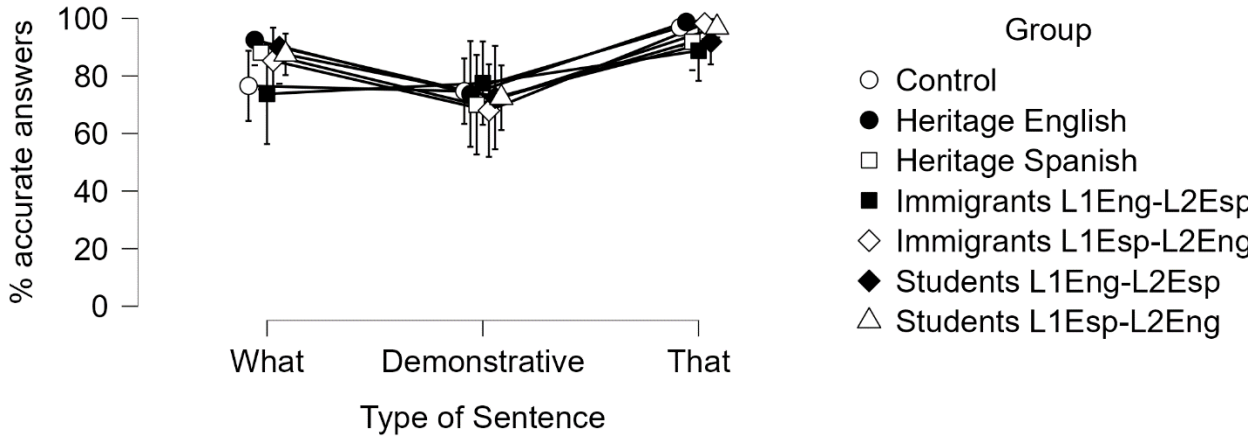
*Task 1 — Percentage of accurate responses based on the interaction Type of Sentence*Group*

Group	That	SD	What	SD	Demonstrative	SD
Control Group	96.9%	15.8	76.6%	35.6	74.7%	33.3
Heritage English	98.7%	3.8	92.5%	16.4	73.7%	34.9
Heritage Spanish	91.9%	18.7	88.1%	25.8	70%	31.8
L1 Eng–L2 Spa ImS	88.7%	24.3	73.7%	30.5	77.5%	30.5
L1 Spa–L2 Eng ImS	98.2%	4.4	85.3%	29.3	67.9%	37.9
L1 Eng–L2 Spa St	91.9%	13.6	90%	23.1	72.5%	34.8
L1 Spa–L2 Eng St	98.2%	4.4	85.3%	29.3	67.9%	37.9

Note. ImS = Immigrants; St = Students

Figure 9

Descriptive Data — Interaction between Group and Type of Sentence



For an overview of the whole descriptive data previously described see Appendix D.

5.1.2.2. Inferential statistics

In order to confirm what was observed in the descriptive data, we carried out a statistical analysis using a mixed logits model that included the three two-way interactions of *Type of Sentence*, *Grammaticality,* and *Group* along with *Participant* and *Item* as random effects. The outcome variable was *Accuracy*. Among the two levels of the dependent variable (*Accuracy*), accurate answers were coded as 1 and inaccurate answers were coded as 0. We chose '0' as our reference level, which implies that the output is given in terms of predicting the odds of observing '1' (accurate responses). Considering our fixed effects, the reference level (base case) for each of our categorical predictors was as follows: for *Type of Sentence* the reference level was the 'That' sentence, for *Group* it was the Control Group, and, for *Grammaticality*, the reference level was Ungrammatical. The performance of the model is shown in the Confusion Matrix[60] on Table 17.

[60] A Confusion Matrix is defined by Vanacore et al. (2022) as 'a cross table that counts the cases that are properly predicted or classified or not correctly predicted or classified' by a given model (364). In our experimental task, 3779 observations were accurate responses, and the model was successful in predicting them as such, but the model failed to predict 509 accurate responses, considering them inaccurate; 218 were inaccurate responses, and the model was

The specificity index (detection of true positives) was 0.88, the sensitivity index (detection of true negatives) was 0.68, which implies that the model was more accurate when predicting accurate responses on well-formed sentences than inaccurate responses on ill-formed sentences.

Table 17

Confusion Matrix — Grammaticality Judgment Task

		Predictions of the model	
		Success	Fail
Actual occurrences	Success	3779	102
	Fail	509	218

Thus, the results of our inferential statistical analyses are as follows: (i) the simple effects of the three predictors of the model (*Type of Sentence*, *Grammaticality,* and *Group*) on the dependent variable (*Accuracy*), (ii) the main effects of the aforementioned independent variables along with the main effects of the three two-way interactions pertaining to the mixed-model (*Type of Sentence*Group + Type of Sentence*Grammaticality + Group*Grammaticality*); these data were obtained from an ANOVA, and (iii) Bonferroni Post Hoc tests of the three main effects and the three two-way interactions.

successful in considering them as such, but the model failed to predict 102 inaccurate responses, considering them accurate.

With respect to the simple effects and starting with the *Type of Sentence*, the results in Table 18 (expressed in odds ratio[61]) indicate that the bilingual groups were significantly less accurate judging 'Demonstrative' sentences (Estimate: 0.05, SE = 0.40, z = -7.3, $p < 0.0001$) and 'What' sentences (Estimate: 0.03, SE = 0.41, z = -8.1, $p < 0.0001$) than the Control group when judging 'That' ungrammatical sentences. In other words, the odds of being accurate decreased in 'Demonstrative' and 'What' sentences. With respect to *Grammaticality*, Table 18 shows that the Control group performed similarly on 'That' sentences irrespective of the *Grammaticality* of the sentence (Estimate: 1.54, SE = 0.36, z = 1.17, $p > 0.05$), which implies that the sentence being (un)grammatical did not have a significant effect on the dependent variable.

Focusing on the simple effect of *Group*, Table 18 reveals that the Heritage English group performed significantly better than the Control group on ungrammatical sentences with 'That' (Estimate: 11.05, SE = 0.95, z = 2.53, $p < 0.05$); the opposite occurred in the case of the Heritage Spanish group (Estimate: 0.24, SE = 0.62, z = -2.28, $p < 0.05$) and the L1Eng–L2Spa Immigrants group (Estimate: 0.21, SE = 0.61, z = -2.53, $p < 0.05$), who were more likely to elicit inaccurate responses on ungrammatical 'That' sentences than the Control group. The rest of the groups did not significantly differ from the Control group in this respect.

[61] An odd ratio larger than 1 implies that the effect on the dependent variable increases. Conversely, if odds are lower than 1, the effect on the dependent variable decreases.

Table 18

Simple Fixed Effects and Random Effects — Grammaticality Judgment Task

							Random Effects	
	Fixed Effects						**Participant**	**Item**
Predictor	Estimate	SE	Z value	CI 2.5%	CI 97.5%	Pr (>|z|)	*SD*	*SD*
(Intercept)	48.10	0.44	8.75	20.22	114.44	**<.0001*** **	1.03	0.28
Type Sentence_Demonstrative	0.05	0.40	-7.28	0.02	0.11	**<.0001*** **	-	-
Type Sentence_What	0.03	0.41	-8.06	0.01	0.08	**<.0001*** **	-	-
Grammaticality_Grammatical	1.54	0.37	1.17	0.75	3.15	.240	-	-
Group_Heritage English	11.05	0.95	2.53	1.72	71.09	**.011 .**	-	-
Group_Heritage Spanish	0.24	0.62	-2.28	0.07	0.82	**.023 .**	-	-
Group_ImS L1Eng–L2Esp	0.21	0.61	-2.53	0.06	0.71	**.011 .**	-	-
Group_ImS L1Esp-L2Eng	1.29	0.69	0.37	0.34	4.96	.710	-	-
Group_Students L1Eng–L2Esp	0.35	0.63	-1.67	0.10	1.20	.096	-	-
Group_Students L1Esp-L2Eng	0.75	0.57	-0.50	0.25	2.30	.620	-	-

Note 1. Signif. codes: 0 '***'; 0.001 '**' ; 0.01 '*' ; 0.05 '**.**'

Note 2. ImS = Immigrants

As for the main effects of the three independent variables and their two-way interactions, the depiction of the ANOVA in Table 19 shows a significant main effect of *Grammaticality* ($F = 71.58$, $p < 0.0001$) and *Type of Sentence* ($F = 67.36$, $p < 0.0001$). With respect to the interactions, there were significant effects of *Type of Sentence*Group* ($F = 5.30$, $p < 0.0001$),

*Grammaticality*Group* (F = 7.99, p < 0.0001) and *Type of Sentence*Grammaticality* (F = 15.90, p < 0.0001).

Table 19

ANOVA — Main effects Grammaticality Judgment Task

Variables	Df	Sum of Squares	Mean Square	F	P-value
Grammaticality	1	71.58	71.58	71.58	**<.0001*** **
Type of Sentence	2	134.71	67.36	67.36	**<.0001*** **
Group	6	2.40	0.40	0.39	0.89
Type of Sentence * Group	12	63.65	5.30	5.30	**<.0001*** **
Grammaticality * Group	6	47.96	7.99	7.99	**<.0001*** **
Type of Sentence * Grammaticality	2	31.80	15.90	15.90	**<.0001*** **

Note. Signif. codes: <0.0001 '***'; <0.001 '**' ; <0.01 '*' ; <0.05 '.'

The Bonferroni Post Hoc test, expressed in log-odds[62] (Table 20) showed that, with respect to the *Type of Sentence*, 'Demonstrative' sentences were likely to elicit less accurate responses than 'What' sentences (Estimate: -1.21, SE = 0.16, z = -7.41, p < 0.001), whereas 'That' sentences were likely to elicit more accurate responses than 'What' sentences (Estimate: +1.13, SE = 0.21, z = 5.31, p < 0.001) and 'Demonstrative' sentences (Estimate: +2.34, SE = 0.20, z = 11.93, p < 0.001).

[62] A negative coefficient would imply that the specific predictor would lead to a decrease in the odds of eliciting an accurate response, whereas a positive coefficient would imply an increase in the odds of eliciting an accurate response.

In relation to *Grammaticality*, Table 7 shows that our participants performed significantly worse on ungrammatical MPs than grammatical NICs (Estimate: -1.04, SE = 0.15, z = -7.02, p < 0.001). Finally, the Post Hoc test revealed no significant differences between the overall performance of the groups.

Table 20

Bonferroni Post Hoc Tests — Fixed effects Grammaticality Judgment Task

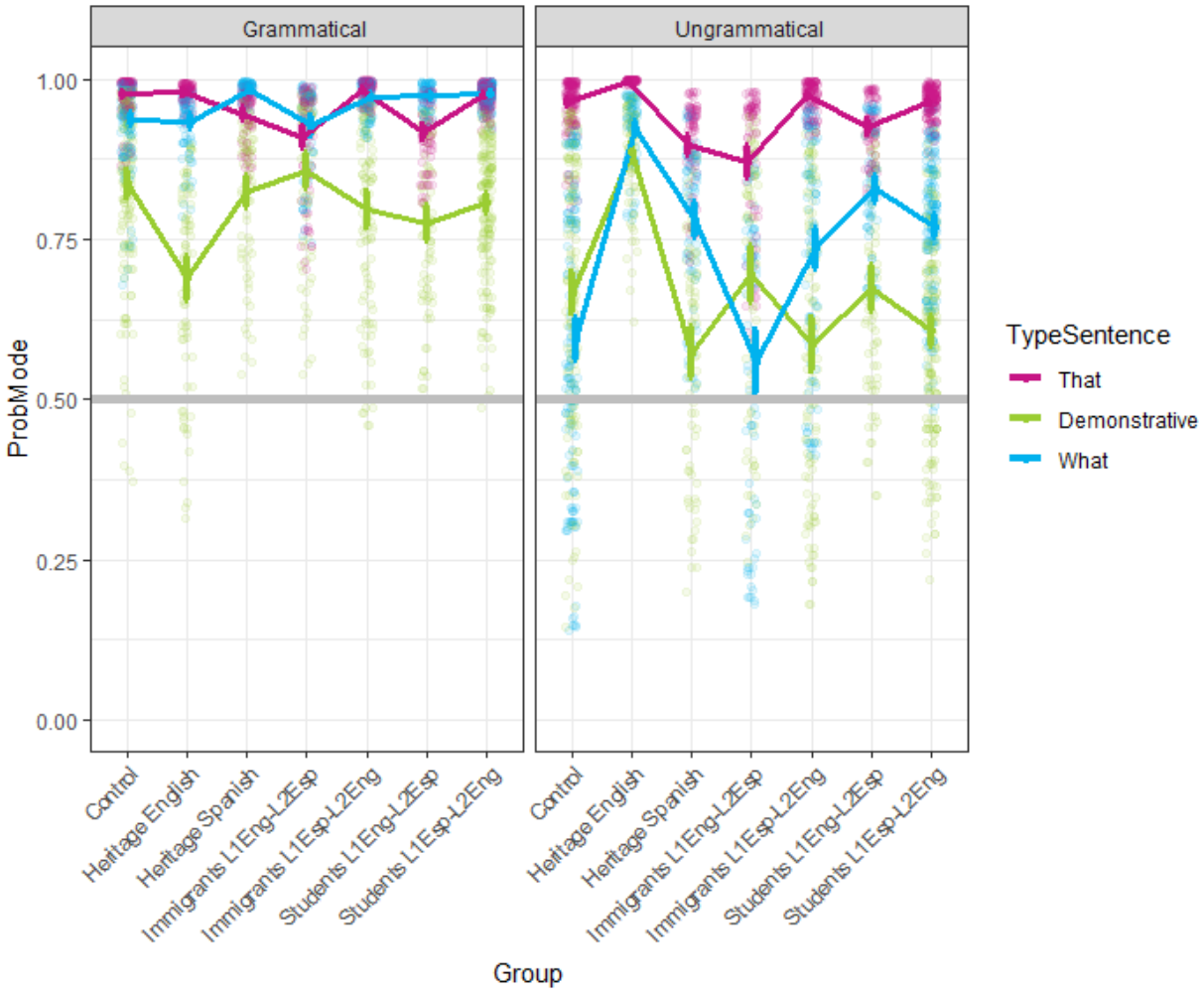
Comparisons	Estimate	SE	Z value	Pr (>|z|)
Demonstrative:What	-1.21	0.16	-7.41	**<.0001*** **
That:Demonstrative	2.34	0.20	11.93	**<.0001*** **
That:What	1.13	0.21	5.31	**<.0001*** **
Ungrammatical:Grammatical	-1.04	0.15	-7.02	**<.0001*** **

Note. Signif. codes: <0.0001 '***'; <0.001 '**' ; <0.01 '*' ; <0.05 '.'

Figure 10 displays a general overview of the data, which displays the probability of obtaining accurate responses given the three main predictors: *Grammaticality* (grammatical and ungrammatical), *Type of Sentence* ('That', 'Demonstrative' and 'What'), and *Group* (Control, Heritage English, Heritage Spanish, L1Esp-L2Eng Immigrants, L1Eng–L2Spa Immigrants, Students L1Esp-L2Eng, Students L1Eng–L2Spa).

Figure 10

Predicted mean probability of accurate responses — Grammaticality Judgment Task



Bonferroni Post Hoc tests of the three two-way interactions of the model (*Type of Sentence*Group* + *Group*Grammaticality* + *Type of Sentence*Grammaticality*) are explored separately in the following sections.

Interaction I: Type of Sentence*Group

Bonferroni Post Hoc Tests showed that there are no significant differences within the same *Type of sentence* between the groups; however, within the groups there are significant differences with respect to the *Type of sentence* (Table 21). These contrasts follow the pattern described in Table 20; namely, 'That' sentences elicited a higher number of accurate answers than 'What' and 'Demonstrative' sentences within most of the groups.

Table 21

*Bonferroni Post Hoc Tests — Two-way interactions of Type of Sentence*Group*

Group	Contrast	Estimate	SE	Z value	Pr (>|z|)
Control	That*Demonstrative	2.59	0.37	6.92	**<.0001*** **
	That*What	2.20	0.38	5.72	**<.0001*** **
	Demonstrative*What	-0.39	0.24	-1.61	0.32
Heritage	That*Demonstrative	3.57	0.75	4.75	**<.0001*** **
English	That*What	2.37	0.79	3.00	**.008***
	Demonstrative*What	-1.20	0.39	-3.10	**.006***
Heritage	That*Demonstrative	1.72	0.37	4.64	**<.0001*** **
Spanish	That*What	-0.16	0.43	-0.37	1.000
	Demonstrative*What	-1.88	0.37	-5.09	**<.0001*** **
Immigrants	That*Demonstrative	0.92	0.35	2.60	**.03.**
L1Eng–L2Spa	That*What	0.88	0.37	1.40	**.05.**
	Demonstrative*What	-0.04	0.33	-0.12	1.000
	That*Demonstrative	3.28	0.50	6.59	**<.0001*** **

Immigrants	That*What	1.68	0.52	3.22	**.004***
L1Spa–L2Eng	Demonstrative*What	-1.60	0.31	-5.14	**<.0001*****
Students	That*Demonstrative	1.57	0.36	4.34	**<.0001*****
L1Eng–L2Spa	That*What	-0.13	0.43	-0.31	1.000
	Demonstrative*What	-1.71	0.37	-4.64	**<.0001*****
Students	That*Demonstrative	2.69	0.35	7.76	**<.0001*****
L1Spa–L2Eng	That*What	1.06	0.38	2.82	**.015.**
	Demonstrative*What	-1.62	0.26	-6.30	**<.0001*****

Note. Signif. codes: <0.0001 '***'; <0.001 '**' ; <0.01 '*' ; <0.05 '.'

Interaction II: Group*Grammaticality

In relation to the differences within each group with respect to *Grammaticality*, the Bonferroni Post Hoc Test in Table 22 shows that for all groups except for the Heritage English group and the L1Eng–L2Spa Students, an inaccurate response was more likely when the sentence was ungrammatical than when it was grammatical. In other words, it was more probable that an ungrammatical sentence such as '**enseñar la lección, eso debe el maestro*' (an ill-formed MPs) was considered grammatical, than a grammatical one such as '*Adivinar el futuro, eso desea mi jefa*' (a well-formed NIC) considered ungrammatical. Within *Grammaticality*, there were no significant differences between the groups when the sentence was grammatical, but when the sentence was ungrammatical, the Heritage English group was more likely to give an accurate response (detect the ungrammaticality of the sentence) than the rest of the groups.

Table 22

*Bonferroni Post Hoc Tests — Two-way interactions Group*Grammaticality*

Comparisons	Estimate	SE	Z value	Pr (>|z|)
Control U: Control G	-1.43	0.25	-5.78	**<.0001*** **
Heritage English U: Heritage English G	1.09	0.37	2.94	0.30
Heritage Spanish U: Heritage Spanish G	-1.71	0.33	-5.20	**<.0001*** **
ImS L1Eng–L2Spa U: ImS L1Eng–L2Spa G	-1.45	0.30	-4.88	**<.001** **
ImS L1Spa–L2Eng U: ImS L1Spa–L2Eng G	-1.52	0.31	-4.97	**<.001** **
St L1Eng–L2Spa U: St L1Eng–L2Spa G	-0.86	0.31	-2.80	0.47
St L1Spa–L2Eng U: St L1Spa–L2Eng G	-1.39	0.245	-5.64	**<.0001*** **
Control U*Heritage English U	-2.02	0.57	-3.55	**.008***
Heritage English U*Heritage Spanish U	2.36	0.62	3.79	**.003***
Heritage English U*ImS L1Eng–L2SpaU	2.56	0.62	4.13	**.0008** **
Heritage English U*ImS L1Spa–L2EngU	1.82	0.60	3.01	**.05.**
Heritage English U*St L1Eng–L2SpaU	1.95	0.63	3.11	**.04.**
Heritage English U* St L1Spa–L2EngU	1.96	0.56	3.48	**.01***

Note 1. Signif. codes: <0.0001 '***'; <0.001 '**' ; <0.01 '*' ; <0.05 '.'

Note 2. ImS = Immigrants; St = Students; U = Ungrammatical; G = Grammatical

Interaction III: Type of Sentence*Grammaticality

With respect to the differences within each *Type of Sentence*, the output of the Bonferroni Post Hoc tests (Table 23) shows that there was a significant contrast in 'Demonstrative' (Estimate: -0.76, SE = 0.19, z = -3.89, p < .0001) and 'What' sentences (Estimate: -2.31, SE = 0.26, z = -8.75, p < .0001) in relation to *Grammaticality*. These results imply that ungrammatical 'Demonstrative' and 'What' sentences were likely to elicit less accurate responses than grammatical 'Demonstrative' and 'What' sentences. However, there were no significant differences with 'That' sentences, which implies that it was more likely that our participants detected both the ungrammaticality in MPs, as in '**mi madre puede **que** mi hermana regrese de Italia*', and the grammaticality in NICs, as in '*el abogado prefiere **que** el testigo venga más temprano*', when the sentence involved 'That' complementation.

With reference to what occurred in ungrammatical sentences, 'That' sentences were likely to elicit an accurate response in comparison to 'Demonstrative' sentences (Estimate: +2.69, SE = 0.26, z = 10.31, p < .0001) and 'What' sentences (Estimate: +2.26, SE = 0.26, z = 8.55, p < .0001). In grammatical sentences, 'That' sentences were more likely to elicit an accurate response in comparison to 'Demonstrative' sentences (Estimate: +1.98, SE = 0.27, z = 7.42, p < .0001); the opposite occurred between 'Demonstrative' and 'What' sentences (Estimate: -1.98, SE = 0.26, z = -7.69, p < .0001) where it was more likely that a grammatical 'Demonstrative' sentence elicited an inaccurate response as compared to grammatical 'What' sentences.

Table 23

*Bonferroni Post Hoc Tests — Two-way interactions of Type of Sentence*Grammaticality*

Comparisons	Estimate	SE	Z value	Pr (>|z|)
That U*G	-0.04	0.30	-0.14	0.888
Demonstrative U*G	-0.76	0.19	-3.89	**<.0001*** **
What U*G	-2.31	0.26	-8.75	**<.0001*** **
Ungrammatical That*Demonstrative	2.69	0.26	10.31	**<.0001*** **
Ungrammatical That*What	2.26	0.26	8.55	**<.0001*** **
Ungrammatical Demonstrative*What	-0.43	0.19	-2.22	0.079
Grammatical That*Demonstrative	1.98	0.27	7.42	**<.0001*** **
Grammatical That*What	-0.01	0.32	-0.02	1.000
Grammatical Demonstrative*What	-1.98	0.26	-7.69	**<.0001*** **

Note 1. Signif. codes: <0.0001 '***'; <0.001 '**' ; <0.01 '*' ; <0.05 '.'

Note 2. U = Ungrammatical; G = Grammatical

In conclusion, when addressing the significance of the main effects, our participants were more likely to elicit an accurate response when facing grammatical NICs than ungrammatical MPs. Also, there was a hierarchy in terms of the *Type of Sentence* in the sense that 'That' sentences were more likely to elicit an accurate response than 'What' and 'Demonstrative' sentences. The independent variable *Group* was not a significant predictor. However, two-way interactions between the aforementioned independent variables nuanced some of the previous conclusions. When we considered the two-way interaction of *Group* and *Grammaticality*, the Heritage English

group significantly elicited more accurate responses than the rest of the groups when judging ungrammatical MPs. The interaction between *Type of Sentence* and *Group*, and *Type of Sentence* and *Grammaticality* confirmed the main effect of the syntactic process: 'That' sentences were more likely to elicit accurate responses than 'What' and 'Demonstrative' sentences regardless of the group and the grammaticality of the sentence.

5.2. Code-switched Two-alternative Forced-Choice Task

5.2.1. Modelling process

Following the procedure used in Task 1, the model simplification process for the mixed logit model of Task 2 was the same as the one described for the first task. In this case, our maximal model included random intercepts for *Participant* and *Item*, in addition to by-participant random slopes for the *Infinitival Construction* (Modal Periphrasis, Nominal Infinitival Construction, and Periphrastic Future), *Directionality* (English and Spanish) and *Language Dominance* (z-scored values from the BLP) and their interactions, and by-item random slopes for *Group* (L1Esp-L2Spa Immigrants, L1Eng–L2Spa Immigrants, L1Esp-L2Eng Students, L1Esp-L2Eng Students, Heritage Spanish, and Heritage English). Similarly to the modelling process in the first task, the first model that converged was an only-intercepts-model with the fixed effects' structure that included the two-way interaction of all predictors but *Language Dominance*: *Directionality*InfinitivalConstruction + Directionality*Group + InfinitivalConstruction*Group*. Further simplification of the data did not improve the model fit. The marginal R^2 of this model was 69% and the conditional R^2 (fixed plus random effects) was 82%, which implies that the mixed model with random effects ($x^2 = 367.28$, Df = 2, $p < 0.001$) has higher explanatory effects than the simple model without them.

5.2.2. Results

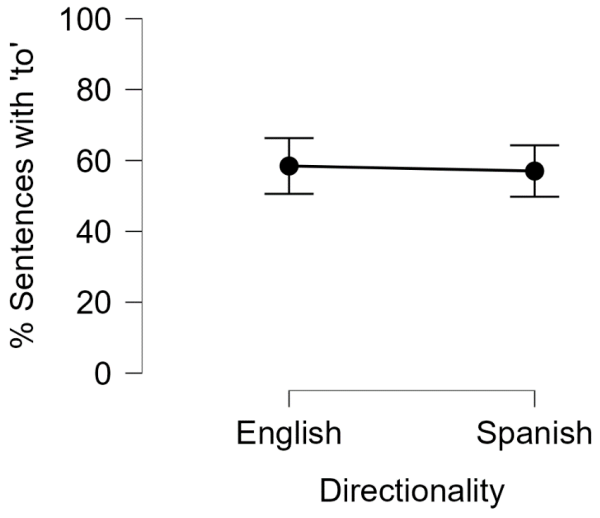
The quantitative results obtained from the descriptive data and the inferential statistics for the Code-switched Two-alternative Forced-Choice Task are shown in the following sections.

5.2.2.1. Descriptive results

According to the descriptive data, when our participants had to choose between two code-switched constructions, identical except for the presence/absence of the English infinitival marker *to*, an average of 57.7% of the overall choices (*SD* = 42.92) went to sentences including the English infinitival marker *to*. When investigating the effect of *Directionality* (see Figure 11), the means between both directionality options were very similar. When the inflected verb was in Spanish, 57% of the chosen sentences had *to* before the English infinitive (*SD* = 41.7) and when the inflected verb was in English, 58.4% of the chosen items had *to* before the Spanish infinitive (*SD* = 44.2). These figures seemed to indicate that our participants did not have a clear preference for the presence/absence of *to* independently of whether the sentence had English–Spanish or Spanish–English directionality.

Figure 11

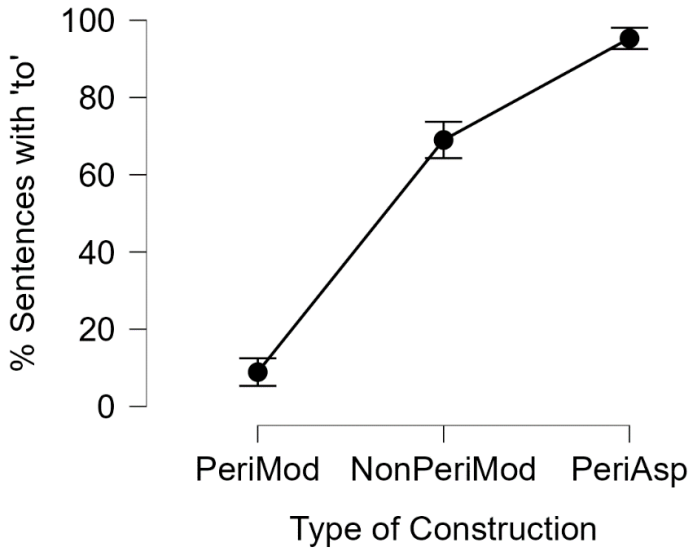
Descriptive Data Task 2 — Directionality



Regarding the *Infinitival Construction* (Figure 12), there seemed to be more differences. When participants judged MPs, 8.8% (*SD* = 20.52) of the chosen sentences had the infinitival marker, with NICs, 69% (*SD* = 31) of the chosen sentences included *to*, and with the PF construction, 95.3% (*SD* = 14.73) of the selected sentences had either *to* or *a*. These results pointed to a preference for including *to* in switches within the PF and in NICs, though to a lesser degree with respect to the latter, and a preference for excluding *to* in switches within MPs.

Figure 12

Descriptive Data Task 2 — Type of Construction



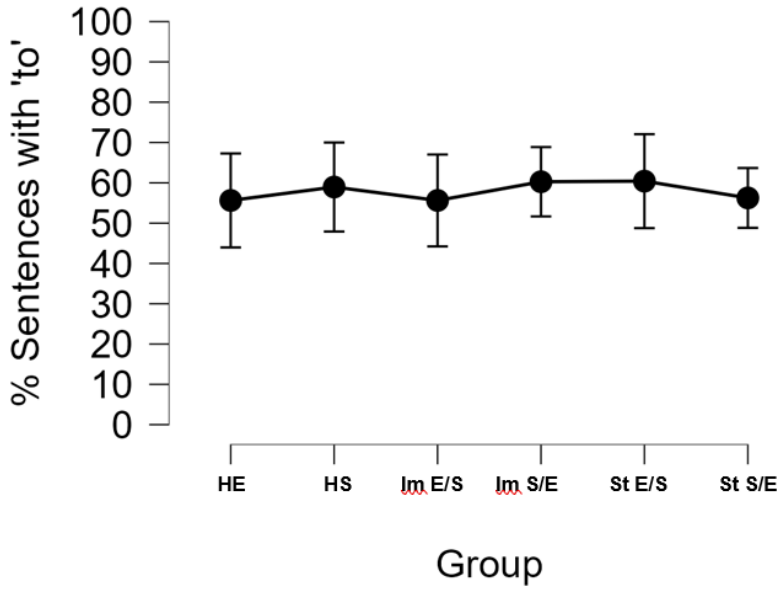
Note. PeriMod = Modal Periphrases; NonPeriMod = Nominal Infinitival Clause; PeriAsp = Periphrastic Future.

Considering the independent variable *Group*, Figure 13 indicates that the group of bilinguals who chose the highest number of sentences with the English infinitival marker *to* was the L1Eng–L2Spa Students group (60.4% of chosen sentences had *to*, *SD* = 45.1), closely followed by the L1Esp-L2Spa Immigrants group (60.3%, *SD* = 39.63), the Heritage Spanish group (58.9%, *SD* = 42.75),

the L1Esp-L2Eng Students group (56.2%, *SD* = 43.09), and with identical means the L1Eng–L2Esp Immigrants group (55.6%, *SD* = 44.11) and the Heritage English group (55.6%, *SD* = 45.12). These results seem to indicate that the preferences between the groups were similar.

Figure 13

Descriptive Data Task 2 — Group



Note: HE = Heritage English; HS = Heritage Spanish; Im E/S = L1 English–L2 Spanish Immigrant; Im S/E = L1 Spanish–L2 English Immigrant; St E/S = L1 English–L2 Spanish Students; St S/E = L1 Spanish–L2 English Students.

When we combined the two previously discussed independent variables (Table 24), we found that within the interaction between *Infinitival Construction* and *Directionality*, NICs and PF tended to increase the inclusion of *to* when the inflected verb was in English, whereas the opposite occurred

with MPs; however, the most notable differences seem to be due not to *Directionality* but to *Construction Type* (Figure 14).

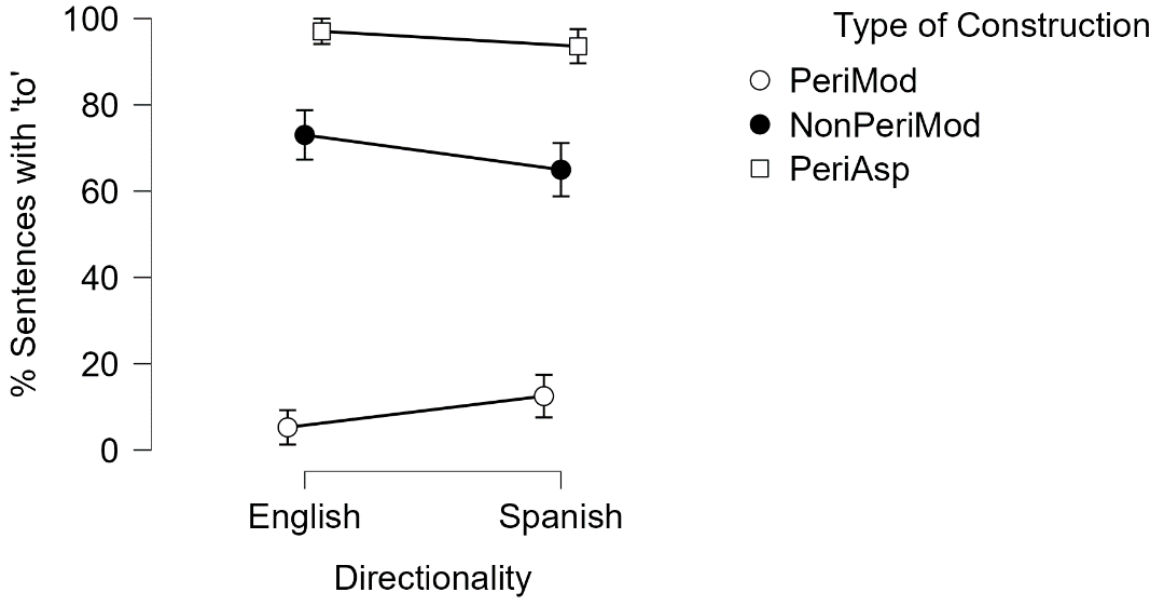
Table 24

*Task 2 — Percentage of sentences with 'to' based on the interaction Type of Construction*Directionality*

Type of Construction	Spanish	SD	English	SD
Modal Periphrasis	12.5%	22.5	5.3%	17.7
Nominal Infinitival Clause	64.9%	32.6	73.0%	29
Periphrastic Future	93.6%	16.3	97.0%	12.9

Figure 14

Descriptive Data Task 2 — Interaction between Type of Construction and Directionality



Note. PeriMod: Modal periphrasis; NonPeriMod: Nominal Infinitival Clauses; PeriAsp: Periphrastic Future

Focusing on the interaction between *Group* and *Directionality* (Table 25), descriptive data showed two different tendencies depending on the group of bilinguals: half of the groups (L1Spa–L2Eng Immigrants, Heritage Spanish, and L1Spa–L2Eng Students) included more instances of the English infinitival marker *to* when the inflected verb was in Spanish and the infinitive in English, while their mirror groups (L1Eng–L2Spa Immigrants, Heritage English, and L1Eng–L2Spa Students) showed a potential higher preference for the presence of the English infinitival marker *to* when the inflected verb was in English and the infinitive in Spanish (Figure 15).

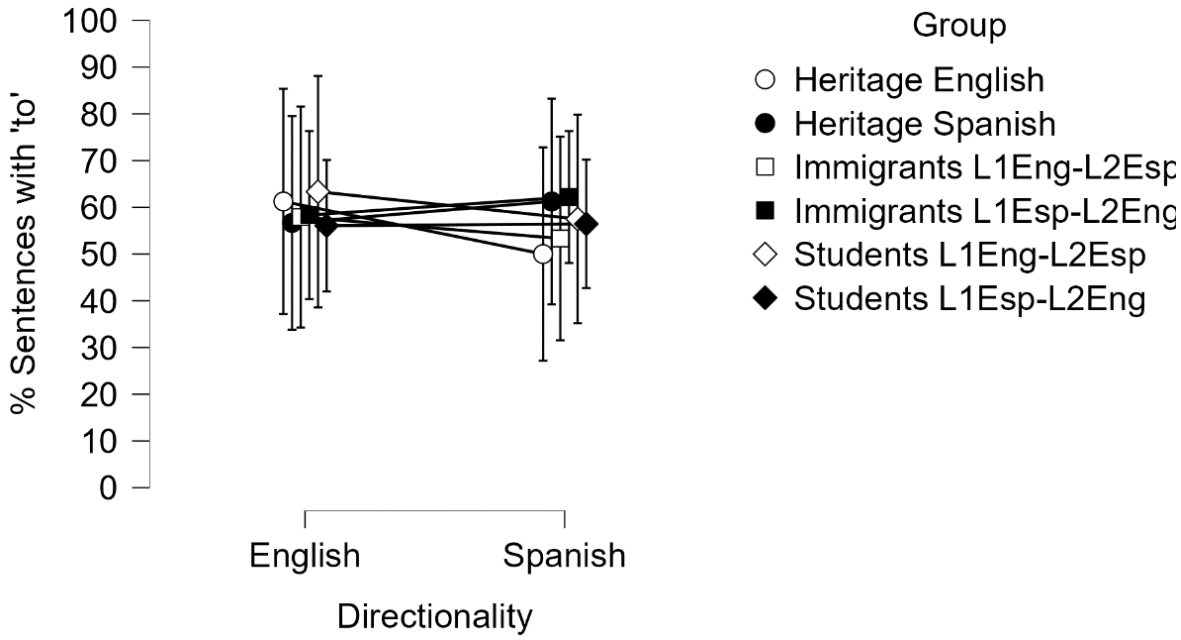
Table 25

*Task 2 — Percentage of sentences with 'to' based on the interaction Group*Directionality*

Group	Spanish	SD	English	SD
Heritage English	50%	44.2	61.2%	46.1
Heritage Spanish	61.2%	41.9	56.7%	44.1
L1 Eng–L2 Spa Immigrants	53.3%	43.3	57.9%	45.5
L1 Spa–L2 Eng Immigrants	62.2%	35.2	58.3%	43.4
L1 Eng–L2 Spa Students	57.5%	44.2	63.3%	46.5
L1 Spa–L2 Eng Students	56.4%%	42.9	56.1%	43.6

Figure 15

Descriptive Data Task 2 — Interaction between Group and Directionality



With regards to the interaction between *Group* and *Type of Construction* (Table 26), all groups showed the same tendency described before; namely, MPs seemed to elicit the lower rate of inclusion of the English infinitival marker *to*, followed by NICs and the PF, with higher averages of inclusion of *to*, especially with the Periphrastic Future. Despite this tendency conditioned by the *Infinitival Construction*, we could also observe potential differences between groups within each construction. In MPs, there seemed to be a noticeable difference between the Heritage English group, with the lowest preference for including *to*, and the L1Spa–L2Eng Immigrants group, with the highest preference for including *to* (Figure 16).

Table 26

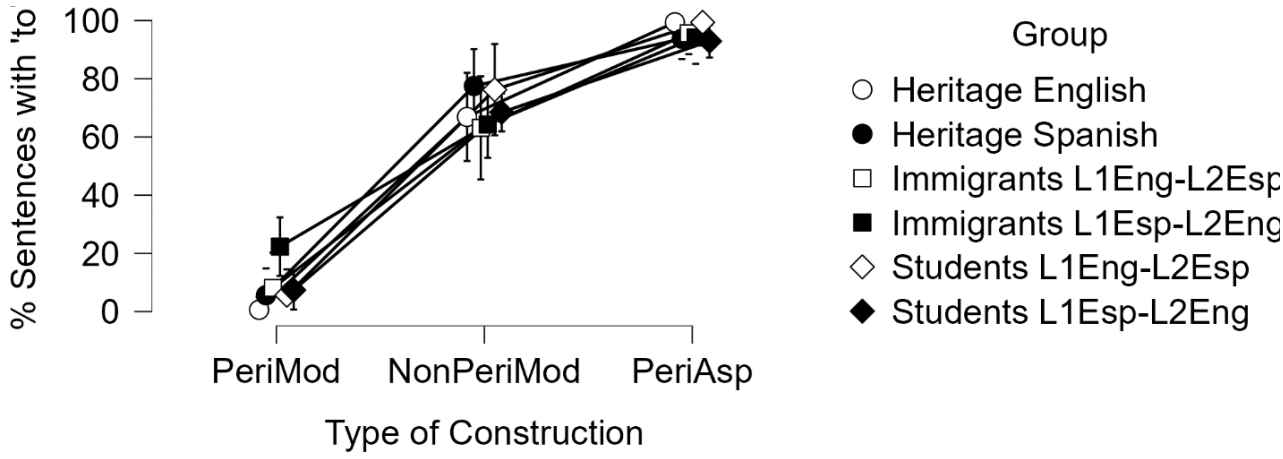
*Task 2 — Percentage of sentences with 'to' based on the interaction Group*Type of Construction*

Group	PeriMod	SD	NonPeriMod	SD	PeriAsp	SD
Heritage English	0.6%	2.8	66.9%	31.2	99.4%	2.8
Heritage Spanish	5.6%	13.7	77.5%	26.8	93.7%	11.8
L1 Eng–L2 Spa ImS	8.1%	17.8	63.1%	38.1	95.6%	12.3
L1 Spa–L2 Eng ImS	22.3%	31.4	64.3%	31.1	94.2%	12.9
L1 Eng–L2 Spa St	5.6%	15.4	76.2%	32.4	99.4%	2.8
L1 Spa–L2 Eng St	7.4%	19.1	68.5%	28.7	92.9%	22.3

Note. PeriMod = Modal periphrasis; NonPeriMod = Nominal Infinitival Clauses; PeriAsp = Periphrastic Future; ImS = Immigrants, St = Students

Figure 16

Descriptive Data Task 2 — Interaction between Type of Construction and Group



Note: PeriMod: Modal periphrasis; NonPeriMod: Nominal Infinitival Clauses; PeriAsp: Periphrastic Future

For an overview summary of the whole descriptive data previously described see Appendix D.

5.2.2.2. Inferential statistics

In order to determine the potential patterns of presence/absence of the English infinitival marker *to* in the three different code-switched infinitival constructions, we performed a logit mixed model with three two-way interactions of three independent variables—*Directionality* of the switch (English or Spanish), *Infinitival Construction* (MP, NIC, PF), and *Group* (Heritage English, Heritage Spanish, L1Esp-L2Eng Immigrants, L1Eng–L2Spa Immigrants, L1Esp-L2Eng Students, and L1Eng–L2Spa Students)—*Participant* and *Item* as random effects, and the dependent variable, presence/absence of *to* between the inflected verb and the infinitive (*ToChoice*). Among the two levels of the dependent variable (*ToChoice*), the presence of *to* was coded as 1 and the absence of *to* was coded as 0. We chose '0' as our reference level, which implies that the output is given in terms of predicting the odds of observing '1' (Presence of *to*). Considering our fixed effects, the reference level for each of our categorical predictors was as follows: for *Directionality*, Spanish was the reference variable, for *Group*, the Heritage Spanish group, and for *Infinitival Construction*, the reference level was Nominal Infinitival Clauses. The performance of the model can be seen in the Confusion Matrix on Table 27. The specificity index (detection of true positives) was 0.91, the sensitivity index (detection of true negatives) was 0.90, which implies that the model was equally and highly effective in predicting the probability of both the inclusion and the exclusion of the English infinitival marker *to*[63].

[63] According to the Confusion Matrix, the model successfully predicted 1972 observations with *to*, but wrongly predicted 217 observations without *to* (when in our data they included *to*); on the other hand, the model successfully predicted 1327 observations without *to*, but wrongly predicted 132 observations with *to* (when in our data they excluded *to*).

Table 27

Confusion Matrix — Two-alternative Forced-Choice Task

		Predictions of the model	
		Present *to*	Absent *to*
Actual occurrences	Present *to*	1972	132
	Absent *to*	217	1327

Following the same structure of the first task, the results of the inferential statistical analyses pertaining the data on CS patterns are as follows: (i) the simple effects of the three predictors of the mixed-model (*Infinitival Construction*, *Directionality,* and *Group*) on the outcome (absence/presence of *to*), (ii) the main effects of the aforementioned independent variables along with the main effects of the three two-way interactions of the mixed-model (*Directionality*InfinitivalConstruction + Directionality*Group + InfinitivalConstruction*Group*) —these data were gathered by ANOVA, and (iii) Bonferroni Post Hoc tests of the three main effects and the three two-way interactions.

Starting with the simple effects of the independent variables, in terms of *Directionality*, Table 28 (depicted in odds ratio) shows that there was no significant effect when the inflected verb was in English for the Heritage Spanish group in NICs (Estimate: 0.57, SE = 0.39, z = -1.42, $p > 0.05$), thus revealing that the Heritage Spanish group showed no specific pattern for including or excluding *to* in NICs connected to the language of the inflected verb. In relation to the *Infinitival Construction*, Table 28 shows that there was a significant positive effect of the PF (Estimate: 2.93,

SE = 0.47, z = 2.29, p < 0.05) and a significant negative effect of the MP (Estimate: 0.01, SE = 0.51, z = -8.55, p < 0.0001). This implies that the Heritage Spanish group tended to include *to* in the PF much more than in NICs, when the inflected verb was in Spanish. The opposite occurred when the construction was an MP; the Heritage English group tended to exclude *to* when the directionality was in Spanish. With respect to the independent variable *Group*, the only significant effect, as shown in Table 28, was observed in the Heritage English group (Estimate: 0.18, SE = 0.76, z = -2.23, p < 0.05) in favour of excluding the English infinitival marker *to* when compared to the Heritage Spanish on NICs when the inflected verb was in Spanish.

Table 28

Simple Fixed Effects and Random Effects — Two-alternative Forced-Choice Task

							Random Effects	
	Fixed Effects						**Participant**	**Item**
Predictor	Estimate	SE	Z value	CI 2.5%	CI 97.5%	Pr (>|z|)	SD	SD
(Intercept)	5.81	0.56	3.13	1.93	17.47	**.001****	1.46	0.36
Directionality_English	0.57	0.39	-1.42	0.26	1.23	.153	-	-
IC_PeriAsp	2.93	0.47	2.29	1.16	7.35	**.022 .**	-	-
IC_PeriMod	0.01	0.51	-8.55	0.00	0.03	**<.0001*****	-	-
Gp_Heritage English	0.18	0.76	-2.23	0.04	0.81	**.026 .**	-	-
Gp_ImS L1Eng–L2Esp	0.24	0.75	-1.87	0.05	1.07	.061	-	-
Gp_ImS L1Esp-L2Eng	0.42	0.70	-1.22	0.11	1.67	.219	-	-
Gp_St L1Eng–L2Esp	0.45	0.76	-1.05	0.10	1.99	.291	-	-

Gp_St L1Esp-L2Eng	0.48	0.65	-1.11	0.13	1.75	.269	-	-

Note 1. Signif. codes: <0.0001 '***'; <0.001 '**' ; <0.01 '*' ; <0.05 '.'

Note 2. IC = Infinitival Construction; Gp = Group; PeriAsp = Periphrastic Future; PeriMod = Modal Periphrasis; ImS = Immigrants; St = Students.

As for the main effects of the independent variables and their three two-way interactions, the ANOVA in Table 29 shows a significant main effect of *Infinitival Construction* ($F = 282.5$, $p < 0.0001$) and of the following interactions: *Directionality*InfinitivalConstruction* ($F = 11.1$, $p < 0.0001$), *Directionality*Group* ($F = 4.57$, $p < 0.001$) and *InfinitivalConstruction*Group* ($F = 6.91$, $p < 0.0001$).

Table 29

ANOVA — Main effects Two-alternative Forced-Choice Task

Variables	Df	Sum of Squares	Mean Square	F	P-value
Directionality	1	0.19	0.19	0.19	0.66
Infinitival Construction	2	565.03	282.51	282.51	**<.0001*** **
Group	5	1.53	0.31	0.31	0.91
Directionality*Infinitival Construction	2	22.27	11.14	11.14	**<.0001*** **
Directionality*Group	5	22.85	4.57	4.57	**<.001** **
Infinitival Construction*Group	10	69.09	6.91	6.91	**<.0001*** **

Note. Signif. codes: <0.0001 '***'; <0.001 '**' ; <0.01 '*' ; <0.05 '.'

Bonferroni Post Hoc tests (expressed in Log-odds) (Table 30) revealed a significant effect of Spanish compared to English (Estimate: -0.43, SE = 0.2, z = -2.14, $p < 0.05$) relating to *Directionality*, which implies that our participants tended to exclude the English infinitival marker *to* when the inflected verb was in Spanish. With respect to *Infinitival Construction*, Table 17 reveals an overall significant difference between the three constructions; namely, our participants were more likely to include *to* in the PF and exclude *to* in MPs; NICs were in an intermediate position at a significant distance between the other two constructions. Finally, Post Hoc tests revealed no significant differences between the overall performance of the groups.

Table 30

Bonferroni Post Hoc Tests — Fixed effects Two-alternative Forced-Choice Task

Comparisons	Estimate	SE	Z value	Pr (>|z|)
Spanish:English	-0.43	-0.20	-2.14	**<.032 .**
NonPeriMod:PeriAsp	-3.51	0.33	-10.59	**<.0001*****
NonPeriMod:PeriMod	4.73	0.28	16.72	**<.0001*****
PeriAsp:PeriMod	8.24	0.42	19.75	**<.0001*****

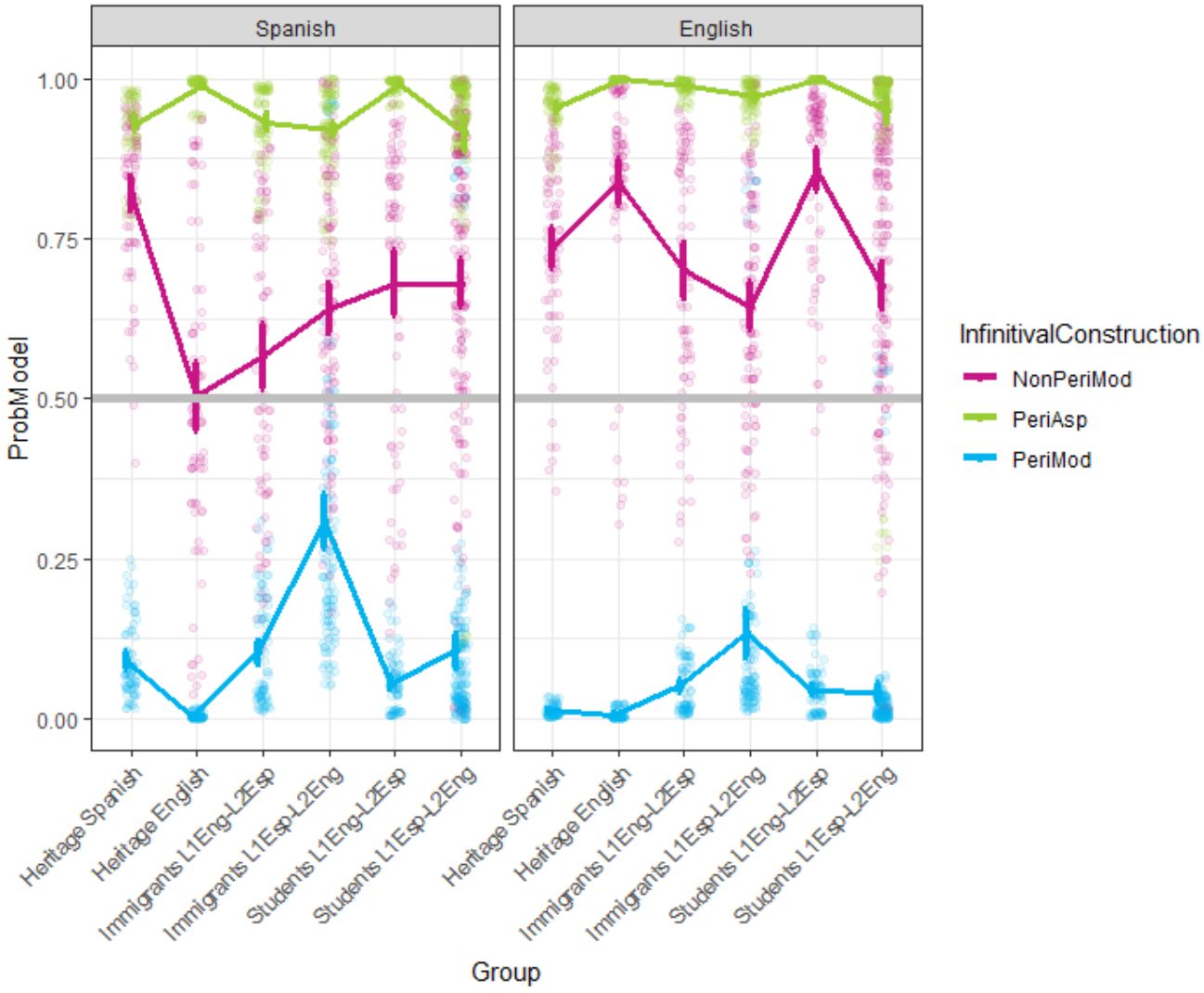
Note 1. Signif. codes: <0.0001 '***'; <0.001 '**' ; <0.01 '*' ; <0.05 '.'

Note 2. PeriMod = Modal periphrasis (MP); NonPeriMod = Nominal Infinitival Clauses (NIC); PeriAsp = Periphrastic Future (PF).

A general overview of the data are shown in Figure 17 where the probability of including *to* is detailed considering the three main predictors: *Directionality* (English or Spanish), *Infinitival Construction* (MP, NIC or PF), and *Group* (Heritage English, Heritage Spanish, L1Esp-L2Eng Immigrants, L1Eng–L2Spa Immigrants, L1Esp-L2Eng Students, L1Eng–L2Spa Students).

Figure 17

Predicted mean of probability of including to *— Two-alternative Forced-Choice Task*



Note 1. PeriMod = Modal periphrasis (MP); NonPeriMod = Nominal Infinitival Clauses (NIC); PeriAsp = Periphrastic Future (PF).

Bonferroni Post Hoc tests of the three two-way interactions of the model (*Directionality*Infinitival Construction + Directionality*Group + Infinitival Construction*Groups*) are explored separately in the following sections.

Interaction I: Directionality*Infinitival Construction

When we focused on the differences within the same constructions, Bonferroni Post Hoc test on Table 31 revealed several significant differences in terms of *Directionality*: English NonPeriMod*Spanish NonPeriMod (Estimate: +0.59, SE = 0.24, z = -2.47, $p < 0.05$), English PeriMod*Spanish PeriMod (Estimate: -1.65, SE = 0.41, z = -4.06, $p < 0.01$) and English PeriAsp*Spanish PeriAsp (Estimate: 1.65, SE = 0.41, z = -4.06, $p < 0.0001$). These results show that, with respect to the NICs, there was a higher probability of our participants including *to* when the inflected verb was in English than when it was in Spanish. In the MPs the opposite occurs; it was more likely that our participants excluded *to* when the inflected verb was in English. Finally, concerning the PF, it was more likely that our participants included *to* when the inflected verb was in English than when the inflected verb was in Spanish. Within the same *Directionality* level (inflected verb in Spanish or in English), Table 31 shows that there were significant contrasts between the constructions that mirrored the same pattern on Table 30: there was a higher probability that our participants included to/a in the PF, followed by the NICs and the MPs, where it was less likely that our participants included *to*, irrespective of the *Directionality* level.

Table 31

*Bonferroni Post Hoc Tests — Two-way interactions Directionality*Infinitival Construction*

Comparisons	Estimate	SE	z. ratio	p. value
English NonPeriMod*Spanish NonPeriMod	0.59	0.24	-2.47	**.014.**
Spanish NonPeriMod*English NonPeriMod	-0.59	0.24	-2.47	0.20
English PeriMod*Spanish PeriMod	-0.95	0.33	2.83	**.005***
Spanish PeriMod*English PeriMod	0.95	0.33	2.83	0.07
English PeriAsp*Spanish PeriAsp	1.65	0.41	-4.06	**<.0001****
Spanish PeriAsp*English PeriAsp	-1.65	0.41	-4.06	**0.0007****
Spanish NonPeriMod : Spanish PeriAsp	-2.98	0.35	-8.41	**<.0001****
Spanish NonPeriMod : Spanish PeriMod	3.96	0.33	12.13	**<.0001****
Spanish PeriAsp : Spanish PeriMod	6.94	0.43	16.09	**<.0001****
English NonPeriMod : English PeriAsp	-4.04	0.45	-9.03	**<.0001****
English NonPeriMod : English PeriMod	5.50	0.37	14.98	**<.0001****
English PeriAsp : English PeriMod	9.54	0.54	17.55	**<.0001****

Note 1. Signif. codes: <0.0001 '***'; <0.001 '**' ; <0.01 '*' ; <0.05 '.'

Note 2. PeriMod = Modal periphrasis; NonPeriMod = Nominal Infinitival Clauses; PeriAsp = Periphrastic Future.

Interaction II: Directionality*Group

The Bonferroni Post Hoc test revealed no significant differences between the groups within the same *Directionality*; however, there were significant differences within three groups with respect to the language of the inflected verb. These differences are displayed in Table 32: the Heritage English group (Estimate: -1.90, SE = 0.48, z = -3.99, p < .0001), the Heritage Spanish group (Estimate: +0.72, SE = 0.36, z = 1.99, p < .05) and the L1Eng–L2Spa Students group (Estimate: -1.15, SE = 0.27, z = -2.76, p < .01). These results show that Heritage English group and the L1Eng–L2Spa Students group were more likely to exclude *to* when the inflected verb was in Spanish; contrary to the Heritage Spanish group that was more likely to include *to* when the inflected verb was in Spanish.

Table 32

*Bonferroni Post Hoc Tests — Two-way interactions Directionality*Group*

Comparisons	Estimate	SE	Z value	Pr (>|z|)
Heritage English S: Heritage English E	-1.90	0.48	-4.00	**.0001*** **
Heritage Spanish S: Heritage Spanish E	0.72	0.36	1.99	**.05.**
ImS L1Eng–L2Spa S: ImS L1Eng–L2Spa E	-0.57	0.34	-1.66	0.10
ImS L1Spa–L2Eng S: ImS L1Spa–L2Eng E	0.14	0.27	0.53	0.60
St L1Eng–L2Spa S: St L1Eng–L2Spa E	-1.15	0.42	-2.76	**.006***
St L1Spa–L2Eng S: St L1Spa–L2Eng E	0.19	0.27	0.70	0.49

Note 1. Signif. codes: <0.0001 '***'; <0.001 '**' ; <0.01 '*' ; <0.05 '.'

Note 2. ImS = Immigrants; St = Students; S = Spanish; E = English

Interaction III: Infinitival Construction*Group

When we explore the differences between the groups within the same *Infinitival Constructions*, the Bonferroni Post Hoc test revealed no significant differences with reference to the NICs and the PF; however, Table 33 shows two significant differences within the MPs: Heritage English * ImS L1Spa–L2Eng (Estimate: -4.41, SE = 1.26, z = -3.50, p < 0.01) and ImS L1Eng–L2Spa * St L1Spa–L2Eng (Estimate: +2.09, SE = 0.61, z = 3.40, p < 0.01). These results indicate that it was more likely that the L1Spa–L2Eng Immigrants group include *to* in these constructions (MPs) when compared to the Heritage English and the L1Spa–L2Eng Students group. If we consider the differences in terms of *Infinitival Construction* within each group, there were significant differences that indicated how the PF was more likely to elicit the inclusion of *to/a* followed by the NICs and the MPs.

Table 33

*Bonferroni Post Hoc Tests — Two-way interactions Infinitival Construction*Group*

	Comparisons	**Estimate**	**SE**	**Z value**	**Pr (>|z|)**
Modal Periphrases	Heritage English * ImS L1SpaL2Eng	-4.41	1.26	-3.50	**.007***
	ImS L1SpaL2Eng * St L1SpaL2Eng	2.09	0.61	3.40	**.01***
Heritage Spanish	NonPeriMod*PeriAsp	-1.61	0.42	-382	**.0004****
	NonPeriMod*PeriMod	5.07	0.50	10.19	**<.0001*****
	PeriAsp*PeriMod	6.67	0.57	11.75	**<.0001*****
Heritage English	NonPeriMod*PeriAsp	-5.86	1.31	-5.15	**<.0001*****
	NonPeriMod*PeriMod	7.18	1.48	6.26	**<.0001*****

	PeriAsp*PeriMod	13.04	1.62	8.05	**<.0001*** **
L1Eng–L2Spa	NonPeriMod*PeriAsp	-3.25	0.50	-6.50	**<.0001*** **
Immigrants	NonPeriMod*PeriMod	3.72	0.42	8.85	**<.0001*** **
	PeriAsp*PeriMod	6.97	0.60	11.62	**<.0001*** **
L1Spa–L2Eng	NonPeriMod*PeriAsp	-2.62	0.37	-6.99	**<.0001*** **
Immigrants	NonPeriMod*PeriMod	2.59	0.30	8.54	**<.0001*** **
	PeriAsp*PeriMod	5.21	0.42	12.53	**<.0001*** **
L1Eng–L2Spa	NonPeriMod*PeriAsp	-4.85	1.08	-4.48	**<.0001*** **
Students	NonPeriMod*PeriMod	5.02	0.50	10.08	**<.0001*** **
	PeriAsp*PeriMod	9.86	1.17	8.44	**<.0001*** **
L1Spa–L2Eng	NonPeriMod*PeriAsp	-2.87	0.37	-7.80	**<.0001*** **
Students	NonPeriMod*PeriMod	4.80	0.38	12.69	**<.0001*** **
	PeriAsp*PeriMod	7.67	0.49	15.61	**<.0001*** **

Note 1. Signif. codes: <0.0001 '***'; <0.001 '**' ; <0.01 '*' ; <0.05 '.'

Note 2. ImS = Immigrants; St = Students; PeriMod = Modal periphrasis (MP); NonPeriMod = Nominal Infinitival Clauses (NIC); PeriAsp = Periphrastic Future (PF).

In conclusion, the clearest pattern provided by the experimental data was determined by the *Infinitival Construction* as a predictor: it was highly probable that bilingual speakers included to/a within a code-switched PF, it was not probable that they included *to* within a code-switched MP, and it was somewhat probable that they included *to* within a code-switched NIC. There were no

significant differences between the bilingual groups and the significance of the *Directionality* of the switch was only revealed though the Bonferroni Post Hoc test that determined that it was more likely to exclude *to* when the inflected verb was in Spanish. Two-way interactions confirmed the previously mentioned pattern in terms of the *Infinitival Construction* in such a way that this pattern occurred regardless of the language of the inflected verb and the *Group*. However, within the same *Directionality* level (English or Spanish) and within each *Group* there were significant differences: when the inflected verb was in English, it was more likely that *to* were absent in MPs, but present in NICs and PF; in terms of *Group*, the L1Spa–L2Eng Immigrants were more likely to include *to* in MPs in comparison to other bilingual groups. Finally, the two-way interaction *Directionality* and *Group* only revealed significant differences between some groups: Heritage English and L1Eng–L2Spa Students were more likely to exclude *to* when the inflected verb was in Spanish, whereas the Heritage Spanish group was more likely to include it.

6. Discussion and conclusion

This chapter commences with the discussion of the two main topics raised in this thesis: first, the distinction between Modal Periphrases (MP) and Nominal Infinitival Clauses (NIC), and second, the search for code-switching patterns displayed by three infinitival constructions, (Modal Periphrases, Nominal Infinitival Clauses, and the Periphrastic Future). This discussion will be followed by a summary of the main findings gathered from our research and contributions to the field. The chapter will end by stating the limitations of the study along with suggestions for future research.

6.1. Distinguishing between MPs and NICs

The first part of this thesis focused on studying the status of two superficially similar constructions, MPs and NICs, in the mind of the Spanish native speakers and of a broad spectrum of Spanish–English bilinguals with differing linguistic backgrounds, namely, differences in timing, setting, mode, quantity, and quality of input. Our first research question focused on whether both the control group and the different types of bilinguals would differentiate the two constructions regardless of their linguistic background. Our second research question aimed at determining whether two independent predictors, such as the grammaticality of the sentence and the type of the syntactic process ('That', 'What', and 'Demonstrative' sentences) played a role in the accurate distinction of MPs and NICs.

Research question 1

In relation to the first research question, our hypothesis was, first, that our Control group would be able to differentiate both infinitival constructions, based on the results gathered by Alonso González (2021), and, second, that there would be significant differences between our six groups of bilinguals in relation to the Control group for two reasons: the first, because of the specific characteristics of these constructions—the variety of grammatical constraints (syntactic processes that differentiate NICs from MPs (García Fernández, 2006; Gómez Torrego, 1999; Topor, 2005)), the absence of distinguishable features between them (both constructions have a similar superficial structure of inflected verb plus infinitive), and their semantic complexity (see section 2.1 on 'Modality and meaning'). These traits could prove to be problematic for three of our experimental groups: the Heritage Spanish group, the L1Eng–L2Spa Immigrants, and the L1Eng–L2Spa Students, due to the fact that they are all English dominant. The second reason lies in the format of the experimental items (untimed, written, (un)grammatical sentences) which could

yield the use of explicit knowledge (conscious and potentially verbalizable) when judging the grammaticality of the sentence (Ellis, 2005; Gutiérrez, 2013; Vafaee et al., 2017; Zyzik & Sánchez, 2019); this could go against those groups of bilinguals with English as a dominant language that have acquired Spanish in a naturalistic environment, namely, the Heritage Spanish group and the L1Eng–L2Spa Immigrant group.

Our descriptive results showed a high accuracy rate for the GJT for all groups, with an average score of 84.1% accurate answers, higher than the 75.8% accuracy of the adult native speakers in Alonso González (2021). If we specifically compare the results of the Spanish native speakers in both studies, the difference increases in favour of the current study, given that in the previous one the participants yielded lower scores, 69.57% accuracy on average on this task. It should be noted that there were differences between both tasks with respect to the syntactic processes included: in Alonso González (2021) we did not include the 'Demonstrative' sentence but did include 'Cleft' sentences and reflexive passives, both absent in our current study.

With regards to the differences between the groups, inferential statistical analyses showed that, contrary to our hypothesis, the independent variable *Group* had no significant effect on *Accuracy*—all groups performed similarly when compared to the Control group and to each other. One possible explanation for this lack of significant differences between the overall performance of the various groups despite their differences in terms of linguistic background could be accounted for by appealing to one of the core issues of the innate view of language acquisition: the Poverty of the Stimulus argument (Chomsky, 1980)[64], also referred to as Plato's Problem and the Logical Problem of Language Acquisition. In relation to L1 acquisition, this argument defends that the

[64] For a detailed overview on L1 acquisition, see Crain & Thornton (1998); on L2 acquisition, see Schwartz & Sprouse (2000, 2013).

child is able to acquire grammatical structures and their constraints despite limited input and the absence of negative evidence in the linguistic data children are exposed to, due to the innate knowledge of the properties of their language (i.e., Universal Grammar (Chomsky, 1965)).

When applied to L2 acquisition, the Poverty of the Stimulus argument sustains that L2 learners could acquire grammatical structures characterized by certain features that, in principle, may hinder their learnability. Among these factors, Heil & López (2020) mention lack of saliency, an intricate web of grammatical constraints, semantic complexity, infrequency, and absence of communicative intent. In this respect, the constructions included in this study could qualify for the first three factors, which, along with the absence of explicit instruction related to these constructions in second and foreign language contexts, may provide evidence in favour of L2 acquisition despite lack of sufficient input (Poverty of the Stimulus). However, there were significant differences between the groups in the interaction with the other two variables, *Grammaticality* and *Type of Sentence*, something that we address in the second research question.

Research question 2

With respect to the second research question, we hypothesized that both the Control group and the bilingual speakers would differentiate both infinitival constructions regardless of the *Type of Sentence* ('That', 'What', or 'Demonstrative' sentence) whilst the *Grammaticality* of the sentence would condition the degree of *Accuracy* in the sense that ungrammatical sentences would elicit lower levels of accuracy when compared to grammatical sentences, in accordance with previous literature, especially in the case of L2 learners (Bialystok, E. 1979, 1986; Ellis, 1991; Godfroid et al., 2015; Gutiérrez, 2013; Hedgcock, 1993; Kim & Nam, 2017; Loewen, 2009; Murphy, 1997; Shiu et al., 2018; Vafaee et al., 2017; Zyzik & Sánchez, 2019).

Concerning the *Type of Sentence* and contrary to our initial hypothesis based on the results gathered by Alonso González (2021)[65], *Type of Sentence* was a significant predictor of *Accuracy*. Our current results show a significant difference with respect to accuracy for each type of process. In fact, even though the overall performance was high, 'That' sentences, as in (102–103), yielded the highest levels of accuracy regardless of the *Grammaticality* of the sentence (95.4% accurate answers), followed by 'What' sentences, as in (104–105), where the infinitive was pronominalized by *qué* (84.3% accuracy), and 'Demonstrative' sentences (72.7% accuracy), as in (106–107), where the infinitive was pronominalized by *eso*. Inferential results confirmed that this hierarchy of syntactic processes was significant.

(102) Mi hijo prefiere que su tía compre unos tomates.

'Mi son prefers that his aunt buy some tomatoes'.

(103) *La familia debe que mi abuela salga a pasear.

*'The family must that my grandmother goes for a walk'.

(104) ¿Qué prefiere el arquitecto?

'What does the architect prefer?'

(105) *¿Qué puede la viuda?

*'What can the widow?'

(106) Ganar ese premio, eso desea mi compañero.

'Winning that award, my colleague wishes that.'

(107) *Comprar un regalo, eso puede esa señora.

*'Buying a present, that lady can that.'

[65] In Alonso González (2021), the syntactic processes (infinitive pronominalization by 'what', 'that' sentences, cleft sentences, and the reflexive passive) did not have a significant effect on accuracy, although 'that' sentences also elicited the highest number of accurate responses.

In order to explain these differences, previous studies on task stimuli and grammaticality state that an error occurring in a simple structure could be easier to detect than an error in a complex sentence (Sutter & Johnson, 1990). This could explain the significant difference between 'What' sentences, a Wh-construction as in (104–105), and 'Demonstrative' sentences, a topicalized construction as in (106–107). It may be the case that this topicalization, contrary to the canonical position in the sentence where the elements of the sentence appear further to the right, may have conditioned the processing and the interpretation of these sentences placing them at the bottom of this hierarchy, at a significant distance from 'What' and 'That' constructions.

With respect to 'What' sentences, an element that may have contributed to the inaccurate identification of their ungrammaticality was the connotation of the verb *deber* as 'to owe', which may have induced our participants to deem grammatical those sentences that would be otherwise ungrammatical when *deber* has a root meaning. To avoid this occurrence, during the design of the task we used an animal noun in all the experimental items with *deber* in 'What' sentences, as shown in (108). Despite this decision, the sentence may have been interpreted as in (109), a well-formed sentence with an unusual meaning:

(108) *¿Qué debe el perro?

*What must the dog?

(109) ¿Qué debe el perro?

What does the dog owe?

There are two factors that may have specifically contributed to placing 'That' sentences at the top of this hierarchy of syntactic processes: frequency, and the light verb *hacer*/do. On the one hand, 'That' sentences are more common in written and oral Spanish than 'What' or 'Demonstrative' sentences, which may have contributed to emphasizing the error in 'That'

sentences and making it more detectable to our participants. On the other hand, our participants may have included the light verb *hacer* (do) after the modal verb while processing 'Demonstrative' and 'What' sentences, as in (110–111) and (112–113), respectively:

(110) *Enseñar la lección, eso puede el maestro.

(111) Enseñar la lección, eso puede *hacer* el maestro.

(112) *¿Qué puede el arquitecto?

(113) ¿Qué puede *hacer* el arquitecto?

Regarding the interaction between *Group* and *Type of Sentence*, most groups followed the previously described pattern where 'That' sentences yielded the best results followed by 'What' and 'Demonstrative' sentences; however, there were two observable differences within the groups in the descriptive data: first, the L1Eng–L2Esp Immigrants group was the only group that performed better in 'Demonstrative' than in 'What' sentences; second, both the L1Eng–L2Esp Immigrants and Control groups' performance on 'What' sentences was worse than the rest of the groups. However, inferential analysis showed that there was neither a significant difference between 'Demonstrative' and 'What' sentences within the L1Eng–L2Esp Immigrant group nor between groups within any of the syntactic processes.

With respect to the *Grammaticality* of the sentence, descriptive and inferential analyses corroborated the hypothesis that ungrammatical sentences significantly yielded a lower level of accurate answers (77.3% ungrammatical vs 90.8% grammatical), in other words, it was more likely that our participants detected the well-formedness of the NICs than the ungrammaticality of the MPs. These results agree with previous literature (Bialystok, E. 1979, 1986; Ellis, 1991; Godfroid et al., 2015; Gutiérrez, 2013; Hedgcock, 1993; Kim & Nam, 2017; Loewen, 2009; Murphy, 1997; Shiu et al., 2018; Vafaee et al., 2017; Zyzik & Sánchez, 2019); however, this result is more

nuanced when taking into account the interactions with the other two predictors, *Group* and *Type of Sentence*.

Regarding the *Grammaticality* of the sentence and the interaction with *Group*, all groups but the Heritage English speakers performed better on grammatical than on ungrammatical sentences. The Heritage English group, with the highest average in overall accurate responses (88%: 90% accurate answers for ungrammatical sentences, 86.6% accurate answers for grammatical sentences), performed significantly better in ungrammatical sentences than the rest of the groups. One possible explanation of these results could be the linguistic background that characterizes this group: with a Language Dominance average of -25.4, this was the most balanced group in our study[66], thus benefitting from both an academic education in Spanish (where the differences between NICs and MPs are explicitly taught) and a positive influence from English. English (a language of prestige in Spain, where our Heritage English participants were recruited) clearly differentiates both constructions since the distinction between modal auxiliary verbs and modal lexical verbs is clear, as in 'I *can* go' versus 'I *prefer* to go'. In addition, there are studies in L2 that showed higher levels of accuracy in ungrammatical sentences among adult subjects (Bley-Vroman et al., 1998; Gass, 1983); however, as pointed out by Hedgcock (1993), these results may be due to familiarity with the stimuli "on the basis of extensive instructions, input, and/or direct negative evidence" (8).

Finally, within the interaction of *Grammaticality* and the *Type of Sentence*, descriptive data confirmed the general trend in which grammatical sentences were likely to elicit a higher number

[66] Another aspect that we can take into consideration when explaining this balanced bilingualism is the status of the English language in Spain, the country were our Heritage English speakers were born. As we mentioned in Chapter three, English is a language of prestige in this country, which may contribute to the balanced bilingualism of these speakers.

of accurate responses than ungrammatical sentences; however, this difference was only significant in 'What' and 'Demonstrative' sentences, there was no significant difference within 'That' sentences, the highest performer of the three syntactic processes regardless of its grammaticality. As we mentioned before when we analyzed this hierarchy of structures, where 'That' sentences (followed by 'What' and 'Demonstrative' sentences) are at the top in terms of the probability of eliciting an accurate response, the higher frequency of 'That' sentences along with inclusion of the light verb *hacer* in the processing of 'What' and 'Demonstrative' ungrammatical sentences may have played a role in eliciting an inaccurate judgment when facing ill-formed 'What' and 'Demonstrative' MPs.

6.2. Code-switching patterns

The second part of this thesis is dedicated to the patterns of inclusion and exclusion of the English infinitival marker *to* within three types of code-switched infinitival constructions. Our first research question was centered on potential patterns of Modal Periphrases (MP), the second research question, on the Periphrastic Future (PF), and the third research question, on Nominal Infinitival Clauses (NICs). Finally, our fourth research question was whether the potential patterns detected in the previous constructions were consistent with the naturalistic data gathered from our longitudinal case study involving two Heritage Spanish children.

Research question 3

With respect to the MPs, we hypothesized that our participants would omit *to* between the inflected modal verb (*poder/deber*/can/must) and the infinitive regardless of the *Group* and the *Directionality* of the switch (inflected verb in English or Spanish), selecting switches such as (114–115), over switches such as (116–117):

(114) The partner must *aumentar el negocio*

'The partner must increase the business.'

(115) *La profesora debe* give clear examples

'The teacher must give clear examples.'

(116) The partner must **to** *aumentar el negocio*

*The partner must to increase the business

'The partner must increase the business.'

(117) *La profesora debe* **to** give clear examples

*The teacher must **to** give clear examples.

'The teacher must give clear examples.'

Descriptive and inferential data confirmed our initial hypothesis on the exclusion of *to* in MPs regardless of the *Directionality* of the switch and the *Group*, given that neither of these predictors had a significant main effect on the inclusion/exclusion of *to*. Overall, when our participants were confronted with two pairs of code-switched sentences where the inflected verb was a modal verb, 8.8% of the chosen sentences had the English infinitival marker *to* between the inflected verb and the infinitive. Compared to the other two infinitival constructions, MPs were the one with the lowest inclusion of *to*, inferential statistics showed that these differences were significant. The main explanation to this pattern is that the lack of *to* (or an equivalent in Spanish) between the modal verb and the infinitive in both languages may have played a role in the general preference for excluding the infinitival *to*. However, there were significant differences in the two-way interaction with the other predictors: *Directionality* and *Group*.

In relation to *Directionality*, the descriptive data showed that when the inflected verb was in Spanish (*poder* or *deber*), our participants selected a higher percentage of sentences including

the English infinitival marker *to* (12%) than when the inflected verb was in English (must or can) with an average of 5.3% selected sentences; inferential data showed that this difference was significant, there was a lower probability that our participants included *to* when the inflected verb was 'must' or 'can' than when it was *poder* or *deber*. With reference to *Group*, descriptive data pinpointed the Immigrants L1Spa–L2Eng as the group with the highest average of selected sentences with *to* in modal periphrases (22.3%), inferential statistics showed that this was significant in comparison to the Heritage English group (0.6%) and the Students L1Spa–L2Eng (7.4%). In both results, it may be the case that English may have had a positive crosslinguistic influence given that the English modal system clearly singles out modal verbs that, with some exceptions such as *ought to*, are not followed by the infinitival marker, something that is learnt explicitly in L2 English classrooms.

Research question 4

With respect to the Periphrastic Future (PF), we hypothesized that our participants would include *to* or *a* between *ir/*going and the infinitive regardless of the *Group* and the *Directionality* of the switch (English or Spanish), thus selecting switches such as (118–119), over switches such as (120–121):

(118) *Voy* ***a*** drink juice

'I'm going to drink juice.'

(119) I'm going **to** *viajar a Estados Unidos*

'I'm going to travel to the United States.'

(120) *Vamos* visit my parents

*We're going visit my parents.

'We're going to visit my parents.'

(121) We're going *comprar unas flores*

*We're going buy some flowers.

'We're going to buy some flowers.'

Descriptive and inferential analyses confirmed the hypothesis of the existence of a pattern of inclusion of *to/a* regardless of *Directionality* and *Group*, in comparison to the other two infinitival constructions. Descriptive data showed that when the infinitival construction was the PF, our participants chose 95.3% of the sentences with *to/a*, compared to 69% in the Nominal Infinitival Clauses (NIC) and the aforementioned 8.8% in MPs; these differences were significant. In the same way that the absence of a word between a modal verb and the infinitive in English and Spanish could explain the preference of excluding it when there is a switch within the MP, the presence of *to/a* in the PF in both languages could explain the preference of including it.

There was, however, a significant difference in the two-way interaction with respect to the *Directionality* of the switch. Descriptive data showed that, despite the high percentage of inclusion of *to/a* when the inflected verb was in Spanish, this percentage was lower (93%) than when the inflected verb was in English (97%); inferential statistical analysis showed that this difference was significant. To explain this preference in terms of *Directionality*, one possible explanation could be found in the experimental items. Within the eight experimental items that constituted the condition of PF with an inflected verb in Spanish, there were two where the final letter of the inflected verb was the vowel *a*, as in (122), and another two with the semi-vowel *y*, as in (123). Although it was a written task, the phonetic similarities between the last vowel of the inflected verb and the preposition *a* may have played a role in excluding it when the inflected verb was in Spanish:

(122) *Va (a)* send the letter.

'She/He is going to send the letter.'

(123) *Voy (a)* drink juice.

'I'm going to drink juice.'

Concerning the interaction of *Infinitival Construction* and *Group*, the significant differences were only perceived within the groups. In contrast to the other two infinitival constructions, there were no significant differences between the groups within the PF where our participants were consistent in the pattern of including *to/a*.

Research question 5

Considering the NICs, we did not expect the same consistency as the previous infinitival constructions, given that there is no equivalent to the English non-finite marker *to*, which is the *-r* suffix in Spanish (Klein, 1984). Due to this, we hypothesized two potential scenarios when the inflected verb is in English: (i) the inclusion of the English infinitival marker *to*, as in (124), the only acceptable switch according to MacSwan (1999); or (ii) the exclusion of the English infinitival marker *to*, as in (125), given that the Functional Head Constraint rejects "code-switching between functional heads that articulate INFL and their complements" (Toribio, 2001, p. 209). When the inflected verb is in Spanish and the infinitive is in English, we hypothesized that our participants would follow the pattern perceived in our corpus data, namely, the pattern without the non-finite marker *to*, as in (126), while discarding the sentence in (127):

(124) The worker prefers **to** *salir del edificio*

'The worker prefers to leave the building.'

(125) The worker prefers *salir del edificio*

*The worker prefers leave the building.

'The worker prefers to leave the building.'

(126) *El científico prefiere* clean the laboratory

*The scientist prefers clean the laboratory.

'The scientist prefers to clean the laboratory.'

(127) *El científico prefiere* **to** clean the laboratory

'The scientist prefers to clean the laboratory.'

Overall descriptive results showed that our participants were more likely to include the *to* than to exclude it, given that 69% of the chosen sentences had the infinitival marker and the probability of including *to* was above 0.5. Inferential results showed that there was a significant difference with the other two infinitival constructions, in the sense that our participants were more likely to exclude *to* in NICs when compared to the PF but were more inclined to include *to* when compared to MPs. Similarly to the PF, there were significant differences in terms of *Directionality* but not in terms of *Group*. Descriptive data showed that when the inflected verb was in English, 73% of the chosen sentences included *to*, compared to 65% of the chosen sentences when the inflected verb was in Spanish. Inferential statistics showed that this difference was significant in the sense that our participants were more likely to include *to* when the inflected verb was 'wish' or 'prefer' than when it was *desear* or *preferir*.

This confirms the hypothesis stated by MacSwan (1999) according to which their preferred option is (23), thus going against the Functional Head Constraint (Toribio, 2001) that disallows switches between functional heads and their complements. However, we could not confirm the

opposite pattern when the inflected verb is in Spanish, as in (25), given that the low likelihood of excluding *to* with *desear* or *preferir* was not significant. With respect to the groups, descriptive data showed that the Heritage Spanish group was the one with the highest percentage of inclusion of *to* in these constructions (77.5%). These data were significant when compared to the Heritage English group. In particular, inferential statistics showed that the latter was more likely to exclude *to* in NICs when the inflected verb was in Spanish in comparison to the Heritage Spanish group. We will discuss these results in the following research question where we contrast these results with the naturalistic data.

Research question 6

Our last research question was formulated around contrasting the previously described experimental data with the conclusions gathered from our naturalistic data, the longitudinal corpus of two Spanish heritage children. The naturalistic data revealed a consistent pattern in the three infinitival constructions. These balanced bilinguals, whose utterances only involved switches with the inflected verb in Spanish (there were no switches with the inflected verb in English), included *a* in the PF, as in (128), excluded *to* within MPs, as in (129), and excluded *to* in NICs, as in (130):

(128) *Voy a* move on (7:11)

'I'm going to move on.'

(129) ¿*Puedes* bring it up? (3;11)

'Could you bring it up?'

(130) *Necesitas* get off the floor (3;7)

'You need to get off the floor.'

We hypothesized agreement between the experimental data and the naturalistic data. Our hypothesis was confirmed with respect to MPs and the PF, given that our participants included *a* in the PF but excluded *to* within MPs. This was not the case for the NICs, since the consistent pattern that we observed in our naturalistic data, namely, the exclusion of *to* when the inflected verb was in Spanish, was not confirmed by our inferential statistics. Although our participants did include the infinitival marker more often when the inflected verb was in English, the lack of significance of the exclusion of *to* when the inflected verb was in Spanish did not coincide with the consistent pattern of exclusion of *to* in NICs gathered from our corpus.

Inferential statistical analyses showed that, in general, the Heritage Spanish speakers were more likely to include *to* when the inflected verb was in Spanish than when the inflected verb was in English, whilst the opposite was the case with the Heritage English group and the L1Eng–L2Spa Students group. This was the case, broadly, for all constructions; however, when we focus on NICs, inferential statistics showed no significant differences between the overall performance of the groups within NICs. Conversely, when analyzing the simple effect of the predictor *Group*, inferential statistics showed that the Heritage English group was more likely to exclude *to* in NICs when the inflected verb is in Spanish in comparison to the Heritage Spanish group in the same construction. This showed that the children and the adult Heritage Spanish group reacted differently, but the Heritage Spanish children and Heritage English adults performed similarly.

If we attempt to explain these differences within Heritage speakers in terms of Language Dominance (i.e., the Heritage English group and our Spanish Heritage children are balanced bilinguals while the Spanish Heritage group is English dominant) we find a contradiction: the experimental group with the highest English dominance, the L1Eng–L2Spa Students, was also likely to exclude *to* when the inflected verb was in Spanish. Thus, a possible explanation could be

a combination of two factors: first, the explicit instruction of Spanish in the Heritage English and the L1Eng–L2Spa groups, which is divergent from the naturalistic acquisition of Spanish that characterizes the Heritage Spanish group; second, the age difference between the Heritage Spanish adults and children, given that the former group has received more English in an instructed setting.

6.3. Summary of findings

This thesis has investigated modal infinitival constructions with a twofold aim, first, to determine whether Spanish dominant native speakers as well as Spanish–English bilinguals with different linguistic backgrounds and varying Spanish–English dominance were able to distinguish between two superficially similar constructions in Spanish: MPs, as in *podemos salir,* and NICs, as in *deseamos salir*. Results showed a positive overall performance, with no significant differences between the groups, but with a significant main effect of the *Grammaticality* of the sentence and the *Type of Sentence*. In particular, ungrammatical sentences significantly decreased the probability of eliciting accurate responses, in line with previous research (Godfroid et al., 2015; Zyzik & Sánchez, 2019), which implies that it is easier to identify a well-constructed NIC than an ill-constructed MP, mainly when the syntactic process implies the pronominalization of the infinitive with the interrogative *qué* and the demonstrative *eso*. This did not occur with 'That' complementation, as our participants were likely to be accurate regardless of the grammaticality of the sentence.

With respect to the second objective, this thesis has moved beyond the debate around whether switches between functional and lexical categories (i.e., AUX/MOD + V) are acceptable; instead, we have aimed to detect potential patterns within switches of three types of English–Spanish code-switched infinitival constructions (MPs, NICs, and PF) with respect to the inclusion/exclusion of the English infinitival marker *to* (preposition *a* for the Spanish PF). Our

results showed that the main predictor for including *to* was the *Infinitival Construction*, more specifically, a bilingual speaker would rarely include *to* within a code-switched MP; the opposite is expected when the switch occurs within the PF: it is highly probable that *to/a* is included between both verbs within a switch between *ir/*going and an infinitive.

That which happens in switches within NICs is less categorical and only one pattern has emerged in our experimental data: when the inflected verb is in English, it is more likely that the bilingual speaker includes the English infinitival marker *to* than when the inflected verb is in Spanish. These results partly corroborate the pattern observed in the code-switched infinitival constructions produced by two Heritage Spanish children in our longitudinal study containing naturalistic data. This corpus showed a clear pattern of exclusion of *to* when the inflected verb is in Spanish (there were no utterances with English *Directionality*). With regards to the previously described patterns of MPs and the PF, there was an alignment between both the experimental and the naturalistic data, specifically when Spanish contributed the inflected verb.

6.4. Contributions to the field

To the best of our knowledge, this thesis is the first empirical attempt to investigate two different but connected phenomena in the broad spectrum of Spanish–English bilingualism: first, the status of MPs and NICs in the mind of different types of bilingual speakers as they compare to native speakers, and second, the potential code-switching patterns favoured by those speakers in relation to three types of infinitival constructions, Modal Periphrases, Nominal Infinitival Clauses, and the Periphrastic Future.

The research on the differentiation of Modal Periphrases and Nominal Infinitival Clauses will contribute novel data, both naturalistic and experimental, to an area of Spanish grammar that

has received little attention in the literature. It will also contribute an empirical perspective regarding the theoretical debate on the auxiliary status of *deber* and *poder*.

The investigation of code-switching patterns in modal and periphrastic constructions constitutes a different perspective in this field if we take into consideration that most studies are focused on determining the acceptability of certain switches that involve functional and lexical categories; however, few studies have attempted to track the specific pattern of inclusion/exclusion of the infinitival *to* in different types of code-switched modal infinitival constructions using both, naturalistic and experimental data. As such, these results contribute new data to discerning potential patterns that govern AUX/MOD–V and V–V code-switched sequences and provide a new perspective on the debate concerning switches between functional and lexical categories.

6.5. Limitations and future research

One of the limitations of this thesis is the low number of participants within some of the experimental groups.[67] In our attempt to keep our groups as homogeneous as possible, a number of participants that did the tasks were not considered for the data analysis because they did not fulfill the linguistic characteristics of their respective groups. As a result, we ended up with lower numbers in some of the groups. To compensate for the low numbers of participants we increased the number of experimental groups. We must emphasize that we had the required number of experimental items per condition (8 tokens per condition in a 3 x 2 design) plus distractors, with the added limitation that an online questionnaire is recommended to last less than 20 minutes (Dörnyei & Taguchi, 2020). Both tasks took 20 minutes on average to complete but the BLP added 10 minutes at the beginning of the questionnaire, leaving the average time devoted to finishing the

[67] It was difficult to recruit participants in site of the fact that an economic compensation was offered.

whole experimental procedure as half an hour. In keeping this balance between the appropriate number of items within our 3 x 2 design and the minimum time suggested for the design of online questionnaires we also lost participants that started the online questionnaire but did not go beyond the BLP. In future research, it may be advisable to simplify the design of the experiment to guarantee its completion by a higher number of participants, thus increasing the numbers per experimental group. Another aspect that could be explored in the prospective design of these tasks is 'gamification', with the aim of increasing the level of engagement and motivation when performing the tasks (Sailer et al., 2017).

Returning to our experimental design, there are two variables whose incorporation would enrich the study: the first is Language Dominance, which we failed to include as a predictor due to convergence issues during the modelling process; the second is Reaction Time (RT), both experimental tasks would benefit from adding this independent variable. In the case of the first task, it would serve to determine whether certain syntactic processes need more time to be processed than others, which could help in determining the reason why there were significant differences between them. In addition, ungrammaticality could occur at different places in the sentence depending on the syntactic process; consequently, higher reaction times would be expected in 'What' sentences compared to 'Demonstrative' or 'That' sentences as participants would be faster when detecting syntactic violations that occur towards the end of the sentence (Kail, 2004; Kail et al., 2012). Relating to the forced-choice task, lower processing times could indicate more certainty in the degree of acceptance of a particular switch, whilst the contrary could imply a less definite pattern. This could add an insight perspective in the second task regarding those predictors without a defined significant effect, *Directionality* and *Group*.

If we take into consideration the specific design of each of the experimental tasks, the first task would benefit from introducing other syntactic processes that distinguish MPs from NICs, given that our results were in contradiction with those of Alonso González (2021) where the syntactic processes did not play a significant role in distinguishing both constructions. In addition, in the current research, grammatical sentences were NICs and ungrammatical sentences were MPs (in the previous study, that was not the case due to the introduction of the reflexive passive, not included in this research) which may have nuanced the role of *Grammaticality* independently from the type of construction, given that our participants had to contrast grammatical NICs with ungrammatical MPs, not grammatical/ungrammatical NICs versus grammatical/ungrammatical MPs. Thus, the inclusion of this former option could be considered for future research.

With regards to the second task, it should be mentioned that we excluded a third possible switch where the particle and the infinitive in the PF are in the same language, as in (131–132):

(131) *Vamos* **to** visit my parents

We're going to visit my parents.

(132) We're going ***a*** *comprar unas flores*

We're going to buy some flowers.

The main reason for excluding this third option was that, when we designed the experiment, the periphrastic future was considered a distractor, not an experimental item. It was during the data analysis that we treated the periphrastic future as a third level in our independent variable *Infinitival Constructions*, along with Modal Periphrases and Nominal Infinitival Clauses. In addition, the inclusion of this third option within the PF would have complicated an already complex design by adding a fourth independent variable (the language of *to/a* in accordance with the inflected verb or the infinitive) for each infinitival construction, not only the PF. Finally, our

naturalistic data did not include this option, whenever the PF was uttered, the language of the particle coincided with the language of the inflected verb, as in, '*Lo voy* ***a*** kick', not with the language of the infinitive, as in '*Lo voy* **to** kick'. Thus, it would be advisable to explore this third option in future research.

Beyond the task design, centering on the topic of study—modal constructions—one issue that could be considered is the analysis of modal constructions in terms of meaning. In our study we have focused on the root meanings of *poder* and *deber* but have left unexplored their epistemic meanings. In her study on modal verbs in Catalan, Picallo (1990) stated that those modals with *epistemic* meaning should be considered VP adjuncts, whilst those modals with *root* meaning have to be analyzed as constituents of INFL. In the particular case of Catalan, verbs such as *deure*/must can only have *epistemic* meaning, while *poder*/can have both deontic and epistemic meanings. Investigating the role of meaning in the syntactic structure of these infinitival constructions, even in contrast to other Romance languages, could provide more data intended to clarify the auxiliary status of *poder* and *deber*. It should also be pointed out that this thesis has explored the differences between the two modal infinitival constructions using a Grammaticality Judgment task. Other written and oral experimental tasks could contribute to further investigating the status of these constructions in the bilingual mind.

At the applied level, there could also be room for a pedagogical approach with a twofold aim: first, it seems clear that learning how to convey modal meaning through MPs headed by *poder* and *deber* could be useful for Spanish L2 learners due to their semantic similarities with English modal verbs, so, following Pihler Ciglic´s (2014) study on the development of communicative competencies, the idea would be to analyze a selection of L2 Spanish textbooks to discern how modality in general and MPs in particular are presented (quantitatively and qualitatively) in order

to determine their relevance in the Spanish L2 classroom curriculum; second, to create a didactic proposal about Spanish infinitival constructions that would provide information to grammarians and language teachers about two types of Spanish constructions that have not been compared before.

Within the code-switching approach, there could also be scope for exploring the different types of switches that were noted in the naturalistic data, especially those that occurred more frequently such as those in (133), where the language of the complement coincides with the language of the infinitive verb, and (134), similar to the previous one but with the Determiner, a functional category, expressed in the same language as the inflected verb.

(133) '¿Puedo *let it go*?' (6;8)

Can I let it go?

(134) 'Yo voy a *make* un *rocket ship*' (3;7)

I'm going to make a rocket ship.

This would open the door to exploring, from a different angle, those changes between lexical and functional categories that have been accounted for in naturalistic data within AUX/MOD + V switches (Di Sciullo et al., 1986; Gawlitzek-Maiwald & Tracy, 1996; Nishimura & Yoon, 1998; A.o.), despite those theories against them (Belazi et al., 1994: González-Vilbazo, 2005; A. o.). It could be the case that what occurs in functional/lexical codeswitching is not a matter of the legitimacy of certain changes, given that the trajectory of code-switching research is full of counterexamples, but of frequency conditioned by the grammars involved in the switch (Lau et al., 2017), which would require further contrasting between experimental and naturalistic

data, along with the potential incorporation of large language models as a complementary (not determining) input.

Furthermore, the code-switching patterns favoured in this research have left us with unanswered questions, especially in regards to the lack of a confirmed pattern within NICs when the inflected verb was in Spanish, precisely the infinitival construction where English–Spanish grammars differ the most. *Directionality* played a significant role only when the inflected verb was in English, not when the inflected verb was in Spanish, which does not allow us to decide whether there is a pattern of excluding *to* when this is the case, as opposed to the distinguishable pattern that we detected in our naturalistic data.

In connection to this, age is another factor that could be further explored, given the contradiction between the performance of our adult Heritage Spanish speakers, more inclined to include *to*, and our naturalistic data from two Heritage Spanish children who omit *to* in every switch within NICs. It is for this reason that further experimental studies along with more naturalistic data, both from written and oral sources[68] and from speakers with different ages and linguistic backgrounds, could be explored and developed not only in the language pair Spanish–English but also in other language pairs with modal constructions. This exploration could go even further by examining other restructuring verbs such as *querer*/want or *necesitar*/need, quite frequent not only in our naturalistic data but also in daily life. This would add more insight into an academic debate on code-switching restrictions that started more than 50 years ago and is still active today.

[68] We suggest the study of written works in Spanglish and the analysis of naturalistic data on different types of adult balanced bilinguals (given that our data came from Spanish heritage children).

In summary, our incursion into modality, the so-called 'soul of the sentence', has not been exempt of limitations, like any other research endeavour, but it has contributed to the existing body of knowledge by presenting the first empirical attempt, using both experimental and naturalistic data, to investigate the status of two superficially similar constructions (NICs and MPs) in the mind of a broad spectrum of Spanish–English bilingual speakers. Ultimately, this study has provided significant insights into the understanding of Spanish modal infinitival constructions from two different perspectives within the field of Hispanic linguistics, paving the way for future research on the characterization of lexical/functional categories and their relation to code-switching patterns in languages with distinct modality systems.

References

Aissen, J., & Perlmutter, D. M. (1976) Clause Reduction in Spanish. *Proceedings of the Second Annual Meeting of the Berkeley Linguistics Society* (University of California, Berkeley).

Alonso González, V. (2021). *On the status of modal infinitival constructions in child and adult native Spanish.* [Doctoral Comprehensive Examination]. University of Ottawa.

Amengual, M., & Simonet, M. (2020). Language dominance does not always predict cross-linguistic interactions in bilingual speech production. *Linguistic Approaches to Bilingualism*, *10*(6), 847–872. https://doi.org/10.1075/lab.18042.

American Councils for International Education (2017). *The National K-12 Foreign Language Enrollment Survey Report*. https://www.americancouncils.org/sites/default/files/FLE-report-June17.pdf

Anderson, G. D. S. (2006). *Auxiliary verb constructions*. Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199280315.001.0001

Aronin, L. (2022). *An advanced guide to multilingualism*. Edinburgh University Press.

Azuma, S. (1991). *Processing and intrasentential code-switching.* [Doctoral dissertation, University of Texas]. ProQuest Dissertations & Theses Global Closed Collection; ProQuest One Literature.

Baker, M. C. (1995). *The polysynthesis parameter*. Oxford University Press.

Bally, C. (1950) *Linguistique générale et linguistique française*. Berne: Francke.

Bartlett, L., & González-Vilbazo, K. E. (2013). The structure of the Taiwanese DP in Taiwanese–Spanish bilinguals: evidence from code-switching. *Journal of East Asian Linguistics, 22*(1), 65–99. https://doi.org/10.1007/s10831-012-9098-3

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bassano, D., Laaha, S., Maillochon, I., & Dressler, W. U. (2004). Early acquisition of verb grammar and lexical development: Evidence from periphrastic constructions in French and Austrian German. *First Language, 24*(1), 33–70. https://doi.org/10.1177/0142723704040704

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Belazi, H., Rubin, E., & Toribio, J. (1994). Code switching and X-Bar theory: The Functional Head Constraint. *Linguistic Inquiry, 25*(2), 221–237.

Benson, E. J. (2001). The neglected early history of codeswitching research in the United States. *Language & Communication, 21*(1), 23–36.

Bentahila, A., & Davies, E. E. (1983). The syntax of Arabic–French code-switching. *Lingua, 59*(4), 301–330. https://doi.org/10.1016/0024-3841(83)90007-4

Bernardini, P. & Schlyter, S. (2004). Growing syntactic structure and code-mixing in the weaker language: The Ivy Hypothesis. *Bilingualism: Language and Cognition, 7*, 49-70.

Bialystok, E. (1979). Explicit and implicit judgments of L2 grammaticality. *Language Learning*, *29*(1), 81–103. https://doi.org/10.1111/j.1467-1770.1979.tb01053.x

Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child Development*, *57*(2), 498–510. https://doi.org/10.2307/1130604

Birdsong, D. (2024, May 28). *Dominance in bilingualism: concept – evaluation – research – theory*. Language Acquisition Laboratory. Universidad de Valladolid.

Birdsong, D., Gertken, L.M., & Amengual, M. (2012). *Bilingual Language Profile: An easy-to-use instrument to assess bilingualism*. COERLL, University of Texas at Austin. https://sites.la.utexas.edu/bilingual/

Bley-Vroman, R., Felix, S. & Ioup, G. (1988). The accessibility of Universal Grammar in adult language learning. *Second Language Research, 4*, 1–32.

Bosque, I. (2015). *Las categorías gramaticales. Relaciones y diferencias* (2nd ed.). Madrid: Síntesis.

Boztepe, E. (2003). Issues in code-switching: Competing theories and models. *Studies in Applied Linguistics & TESOL, 3*(2). https://doi.org/10.7916/salt.v3i2.1626

Bravo, A., García Fernández, L., Krivochen, D. G. (2015). On auxiliary chains: Auxiliaries at the syntax-semantics interface. *Borealis: An International Journal of Hispanic Linguistics, 4*(2), 71–101. http://dx.doi.org/10.7557/1.4.2.3612

Bueno Hudson, R. (2021) *El español en Estados Unidos. El español en el mundo*. Anuario del Instituto Cervantes. Retrieved from CVC. Anuario 2021. El español en el mundo 2021. El español en Estados Unidos. Richard Bueno Hudson. (cervantes.es)

Bybee, J. L., Pagliuca, W., & Perkins, R. D. (1991). Back to the future. In E. C. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization: Volume II. Types of grammatical markers* (pp. 17–58). John Benjamin.

Camacho, J. (2018) *Introducción a la sintaxis en español.* Cambridge University Press.

Cantone, K. F. (2007). *Code-switching in bilingual children*. Dordrecht: Springer.

Cantone, K. F., & Müller, N. (2005). Codeswitching at the interface of language-specific lexicons and the computational system. *The International Journal of Bilingualism : Cross-Disciplinary, Cross-Linguistic Studies of Language Behavior, 9*(2), 205–225. https://doi.org/10.1177/13670069050090020501

Cardinaletti, A. & G. Giusti (2006). 'Semi-lexical' Motion Verbs in Romance and Germanic. In N. Corver & H. van Riemsdijk (Eds.), *Semi-lexical categories. On the function of content words and content of function* (pp. 371–414). Berlin, Mouton de Gruyter.

Carrasco Gutiérrez, A. (2019). On the actuality entailment of Spanish root modals. *Borealis (Tromsø), 8*(2), 111–141. https://doi.org/10.7557/1.8.2.4842

Castillo, C. (2003). *English syntax for Spanish speakers: A comparative introduction*. Peter Lang.

Carlisle, R. (1980) Spanish Equivalents of English Modal Verbs. In R. Nash & D. Belaval (Eds.), *Readings in Spanish–English contrastive linguistics* (Vol. 2, pp. 40–97). Inter American University Press.

Carnie, A. (2013). *Syntax: a generative introduction* (3rd ed.). Wiley-Blackwell.

Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English. A comprehensive guide: Spoken and written English grammar and usage*. Cambridge University Press.

Cervantes Institute (2006). *Plan curricular del Instituto Cervantes. Niveles de referencia para el español*. Madrid: Instituto Cervantes, Biblioteca nueva.

Cestero Mancera, A. M., Forment Fernández, M., Gelabert Navarro, M. J. & Martinell Gifre, E. (2020) *Hablar español sin palabras. Diccionario audiovisual de gestos españoles*. Universidad de Alcalá. Diccionario audiovisual de gestos españoles (uah.es)

Chan, B. H. S. (2003). *Aspects of the syntax, the pragmatics, and the production of code-switching: Cantonese and English*. Peter Lang.

Chan, B. H. S. (2008). Code-switching, word order and the lexical/functional category distinction: Formal syntactic approaches to bilingual code-switching. *Lingua, 118*(6), 777–809.

Choi, J., (1991). Korean-English code-switching: Switch-alpha and Linguistic constraints. *Linguistics* (29), 877–902.

Chomsky, N. (1957). *Syntactic structures*. Mouton.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA : MIT Press.

Chomsky, N. (1980). *Rules and representations*. Columbia University Press.

Chomsky, N., Mukherji, N., Patnaik, B., & Agnihotri, R. (2000). *Noam Chomsky: The Architecture of Language*. Oxford University Press: Oxford.

Cournane, A. (2020). Learning modals: A grammatical perspective. *Language and Linguistics Compass, 14*(10), 1–22. https://doi.org/10.1111/lnc3.12399

Cowart, W. (1997). *Experimental syntax.* Thousand Oaks, CA: Sage.

Crain, S., & Thornton, R. (1998). *Investigations in universal grammar a guide to experiments on the acquisition of syntax and semantics*. MIT Press.

Criado de Diego, C., & San Mateo Valdehíta, A. (2020). Las perífrasis modales con deber en el español actual: Estudio de corpus. *Spanish in Context, 17*(1), 84–107. https://doi.org/10.1075/sic.18011.die

Díaz-Campos, M. (2014). *Introducción a la Sociolingüística Hispánica*. Wiley-Blackwell.

Dietrich S. and Hernandez, E. (2022) *Language use in the United States: 2016*. United States Census Bureau. US Department of Commerce. Retrieved from https://www.census.gov/library/publications/2022/acs/acs-50.html

De Bruin, A., Samuel, A. G., & Duñabeitia, J. A. (2018). Voluntary language switching: When and why do bilinguals switch between their languages? *Journal of Memory and Language, 103*, 28–43. https://doi.org/10.1016/j.jml.2018.07.005

De Hower, A. (1990). *The acquisition of two languages from birth: a case study*. Cambridge, MA: Cambridge University Press.

Di Sciullo, A.M., Van Ameringen, A., Cedergren, H., & Pupier, P. (1976). Étude de l'interaction verbale chez des montréalais d'origine italienne. *Cahier de linguistique (Montréal), 6*(6), 127–153. https://doi.org/10.7202/800044ar

Di Sciullo, A. M., Muysken, P., & Singh, R. (1986). Government and code-mixing. *Journal of Linguistics, 22*(1), 1–24. https://doi.org/10.1017/S0022226700010537

Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge University Press.

Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research: construction, administration, and processing* (2nd ed.). Routledge.

Dussias, P. E. (2003). Spanish–English code-mixing at the auxiliary phrase: Evidence from eye-movement data. *Revista Internacional de Lingüística Iberoamericana: RILI, 1*(2 (2)), 7–34.

Eberhard, D. M., Simons, G. F., & Fennig Ch. D. (Eds.). (2021). *Ethnologue: Languages of the World. Twenty-fourth edition.* Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Edwards, J. (2014) Bilingualism and Multilingualism: Some Central Concepts. In T. K. Bhatia & W. C. Ritchie (Eds.), *The Handbook of Bilingualism and Multilingualism* (pp. 5–25). Oxford, UK: Blackwell Publishing.

Ellis, R. (1991). Grammaticality Judgments and Second Language Acquisition. *Studies in Second Language Acquisition*, *13*(2), 161–186. https://doi.org/10.1017/S0272263100009931

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language. *Studies in Second Language Acquisition, 27*(2), 141–172. https://doi.org/10.1017/S0272263105050096

Emonds, J. (2006). The flat structure economy of semi-lexical heads. In N. Corver & H. van Riemsdijk (Eds.), *Semi-lexical categories. On the function of content words and content of function* (pp. 23–66). Berlin, Mouton de Gruyter.

European Commission (2012). *Europeans and their Languages–Factsheets Spain*. Retrieved from https://europa.eu/eurobarometer/surveys/detail/1049

Fábregas, A. (2005). *The definition of the grammatical category in a syntactically oriented morphology: the case of nouns and adjectives*. [Unpublished doctoral dissertation]. Universidad Autónoma de Madrid.

Farrall, M. L. (2012). Oral Language. In M.L. Farrall (Ed.), *Reading Assessment* (pp. 27–46). John Wiley & Sons. https://doi-org.proxy.bib.uottawa.ca/10.1002/9781118092668.ch3

Fernández Alonso, R., Lloro Javierre, J., Montes Garcés, E. & Pascual-Salcedo Sáez, M. (2021) *El español en Canadá. El español en el mundo*. Anuario del Instituto Cervantes. Retrieved from CVC. Anuario 2021. El español en el mundo 2021. El español en Canadá. Rodolfo Fernández Alonso, Javier Lloro Javierre, Elisabeth Montes Garcés y Melania Pascual-Salcedo Sáez. (cervantes.es)

Fernández Fuertes, R., & Liceras, J. M. (2018) Bilingualism as a first language: language dominance and crosslinguistic influence. In A. Cuza & P. Guijarro-Fuentes (Eds.), *Language Acquisition and Contact in the Iberian Peninsula* 57, (pp. 159–186). De Gruyter. https://doi.org/10.1515/9781501509988-007

Fisherman, J. (2014). Three Hundred-Plus Years of Heritage Language Education in the United States. In J. K. Peyton, T.G. Wiley, N. Liu, D. Christian & S. C. K. Moore (Eds.), *Handbook of Heritage, Community, and Native American Languages in the United States* (pp. 50–58). Routledge. https://doi.org/10.4324/9780203122419-11

Flynn, S., Foley, C., & Vinnitskaya, I. (2005). New Paradigm for the Study of Simultaneous v. Sequential Bilingualism. In J. Cohen, K. Mcalister, K. Rolstad, & J. MacSwan (Eds.), *ISB4: Proceedings of the 4th International Symposium on Bilingualism* (pp. 768–774). Somerville, MA: Cascadilla Press.

Fontanella de Weinberg, M. B. (1970). Los auxiliares españoles. *Anales del instituto de lingüística*, (10), 62–73.

Fox, J., & Monette, G. (1992). Generalized Collinearity Diagnostics. *Journal of the American Statistical Association*, *87*(417), 178–183. https://doi.org/10.1080/01621459.1992.10475190

Fox, J. (2014). *Which variance inflation factor should I be using: GVIF or GVIF^(1/(2*Df))*, https://stats.stackexchange.com/users/45067/john-fox

García Alcaraz, E. (2021). *The cognitive and linguistic abilities of bilinguals with genetic disorders: the Prader-Willi Syndrome population.* [Unpublished doctoral dissertation]. University of Ottawa.

García Fernández, L. (dir.) (2006). *Diccionario de perífrasis verbales*. Madrid: Gredos.

García Fernández, L., Krivochen, D. G. & Bravo Martín, A. (2017). Aspectos de la semántica y sintaxis de las cadenas de verbos auxiliares en español. *Moenia: Revista license de lingüística & literatura, 27*, 21-28.

Gass, S. (1983). The development of L2 intuitions. *TESOL Quarterly, 17*, 273–91.

Gawlitzek-Maiwald, I., & Tracy, R. (1996). Bilingual boot-strapping. *Linguistics, 34*(5), 901–926.

Genesee, f. (1989). Early bilingual development: One language or two? *Journal of Child Language 16*, 161–179.

Genesee, F. (2000). Syntactic aspects of bilingual acquisition. *Bilingualism, 3*(3), 167–172.

Gertken, L. M., Amengual, M., & Birdsong, D. (2014). Assessing language dominance with the Bilingual Language Profile. In P. Leclercq, A. Edmonds, & H. Hilton (Eds.), *Measuring L2 Proficiency: Perspectives from SLA*, (pp. 208–225). Multilingual Matters.

Giancaspro, D. (2015). Code-switching at the auxiliary-VP boundary: A comparison of heritage speakers and L2 learners. *Linguistic Approaches to Bilingualism, 5*(3), 379–407. https://doi.org/10.1075/lab.5.3.04gia

Gigerenzer, G., & Richter, H. R. (1990). Context effects and their interaction with development: Area judgments. *Cognitive Development, 5*(3), 235–264. https://doi.org/10.1016/0885-2014(90)90017-N

Gigerenzer, G., Krauss, S. & Vitouch, O. (2004). The Null Ritual: What you always wanted to know about significance testing but were afraid to ask. In D. Kaplan (Ed.), *The SAGE Handbook of Quantitative Methodology for the Social Sciences*. SAGE Publications, Incorporated. https://doi.org/10.4135/9781412986311.n21

Godfroid, A., Loewen, S., Jung, S., Park, J.H., Gass, S., & Ellis, R. (2015). Timed and untimed grammaticality judgments measure distinct types of knowledge: Evidence from Eye-Movement Patterns. *Studies in Second Language Acquisition, 37*(2), 269–297. https://doi.org/10.1017/S0272263114000850

Gómez Torrego, L. (1999). Los verbos auxiliares. Las perífrasis verbales infinitivo. In I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española* (Vol. 2, pp. 3323–3390). Espasa Calpe.

Gómez Torrego, L. (2011). *Gramática didáctica del español*. Madrid: Ediciones SM

González Vilbazo, K. (2005) *Die Syntax des Code-Switching. Esplugisch: Sprachwechsel an der Deutschen Schule Barcelona*. [Unpublished Doctoral dissertation]. Universität zu Köln.

Goodall, G. (2021). Sentence acceptability experiments: What, How, and Why. In G. Goodall (Ed.), *The Cambridge Handbook of Experimental Syntax* (pp. 7–38). Cambridge University Press. https://doi.org/10.1017/9781108569620.002

Goodz, N. S. (1989). Parental language mixing in bilingual families. *Infant Mental Health Journal, 10*(1), 25–44.

Goto Butler, Y. (2014) Bilingualism / Multilingualism and Second-Language Acquisition. In T.K. Bhatia & W.C. Ritchie (Eds.), *The Handbook of Bilingualism and Multilingualism* (pp. 109–136). Blackwell Publishing. https://doi-org.proxy.bib.uottawa.ca/10.1002/9781118332382.ch2

Grano, T. (2015). Explaining the mixed properties of want. In *Control and Restructuring* (Vol. 56). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198703921.003.0004

Grosjean, F. (1982) *Life with two languages: An introduction to bilingualism*. Cambridge, Mass.: Harvard University Press.

Grosjean, F. (2021). *Life as a Bilingual*. Cambridge, MA: Cambridge University Press.

Gutiérrez, X. (2013). The construct validity of grammaticality judgment tests as measures of implicit and explicit knowledge. *Studies in Second Language Acquisition, 35*(3), 423–449. https://doi.org/10.1017/S0272263113000041

Guzzardo Tamargo, R. (2012). *Linking comprehension costs to production patterns during the*

processing of mixed language. [Unpublished doctoral dissertation]. Pennsylvania State University.

Hacquard, V. (2011). Modality. In C. Maienborn, K. Von Heusinger, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning* (pp. 1484–1515). Berlin, Germany: De Gruyter Mouton.

Hadlich, R. (1971). *A transformational grammar of Spanish*. Prentice-Hall.

Hakutani, Y. (1978). The features of modals in English syntax. *Linguistics, 16*(209), 59–70. https://doi.org/10.1515/ling.1978.16.209.59

Haspelmath, M. (2015). Defining vs diagnosing linguistic categories: a case study of clitic phenomena. In J. Błaszczak, D. Klimek-Jankowska and K. Migdalski (Eds.), *How categorical are categories?* Berlín: De Gruyter Mouton.

Hedgcock, J. (1993). Well-formed vs ill-formed strings in L2 metalingual tasks: specifying features of grammaticality judgements. *Second Language Research*, *9*(1), 1–21. https://doi.org/10.1177/026765839300900101

Heil, J., & López, L. (2020). Acquisition without evidence: English infinitives and poverty of stimulus in adult second language acquisition. *Second Language Research, 36*(4), 415–443.

Hernanz Carbó, M. L. (1980) Las perífrasis verbales de infinitivo en español: hacia una posible solución transformacional. *Revista Española de Lingüística, 10*(2), 411–444.

Hoye, L. (2013). *Adverbs and modality in English*. Routledge. https://doi.org/10.4324/9781315844015

Huddleston, R. D. (1976a). *An introduction to English transformational syntax*. Longman.

Huddleston, R. D. (1976b). Some theoretical issues in the description of the English verb. *Lingua, 40*(4), 331–383. https://doi.org/10.1016/0024-3841(76)90084-X

Iordăchioaia, G., I. Roy & K. Takamine (Eds.). (2013). *Categorization and Category Change*. Cambridge: Cambridge Scholars Publishing.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

JASP Team (2023). JASO (version 0.18) [Computer software] http://jasp-stats.org/

Jespersen, O. H. (1940). *A modern English grammar on historical principles*. (Vol 7). Copenhagen: Ejnar Munksgaard.

Johnson, J. & Newport, E. (1991). Critical period effects on universal principles of language. The status of subjacency in the acquisition of a second language. *Cognition, 39*, 215–258.

Johnson, S. G. *The NLopt nonlinear-optimization package*, http://github.com/stevengj/nlopt

Joshi, A. (1985). Processing of sentences with intrasentenial code-switching. In D. R. Dowty, L. Kattunen, & A. M. Zwicky (Eds.), *Natural Language Parsing: Psychological, Computational and Theoretical Perspectives* (pp. 145–150) Cambridge: Cambridge University Press.

Kail, M. (2004). On-line grammaticality judgments in French children and adults: a crosslinguistic perspective. *Journal of Child Language*, *31*(3), 713–737. https://doi.org/10.1017/S030500090400649X

Kail, M., Lemaire, P., & Lecacheur, M. (2012). Online grammaticality judgments in French young and older adults. *Experimental Aging Research*, *38*(2), 186–207. https://doi.org/10.1080/0361073X.2012.660031

Kastronic, L. & Poplack, S. (2021). Be that as it may: The Unremarkable Trajectory of the English Subjunctive in North American Speech. *Language Variation and Change, 33*(1), 107–134. https://doi.org/10.1017/S095439452100003X

Kayne, R. (1989). Null Subjects and Clitic Climbing. In O. Jaeggli & K. Safir (Eds.), *The Null Subject Parameter* (pp. 239-261). Kluwer, Dordrecht.

Kim, J., & Nam, H. (2017). Measures of implicit knowledge revisited: Processing Modes, Time Pressure, and Modality. *Studies in Second Language Acquisition*, *39*(3), 431–457. https://doi.org/10.1017/S0272263115000510

Klein, P. W. (1968). *Modal auxiliaries in Spanish*. [Master's thesis, University of Washington]. University Microfilms.

Klein, P. W. (1982). On defining auxiliary verbs. *Spanish. Hispanic Journal, 4*(1), 113–128.

Klein, P. W. (1985). Apparent Correspondences in Spanish to English Infinitival to. *Hispania 67*(3), 416–419.

Kocher, A. (2017). From verum to epistemic modality and evidentiality: On the emergence of the Spanish Adv+C construction. *Journal of Historical Linguistics, 7*(1–2), 77–110. https://doi.org/10.1075/jhl.7.1-2.04koc

Köpke, B. & Schmid, M. (2004). First language attrition: The next phase. In M. Schmid, B. Köpke, M. Keijzer, & L. Weilmar (Eds.), *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues* (pp. 1–43). Amsterdam: John Benjamins.

Köppe, R. & Meisel, J. M. (1995). Code-Switching in bilingual first language acquisition. In L. Milroy & P. Muysken (Eds.), *One speaker, two languages. Cross-disciplinary perspectives on code-switching* (pp. 276–301). Cambridge University Press.

Kovacci, O. (1990). *El comentario gramatical: teoría y práctica*. Arco/Libros.

Kratzer, A. (1981). The Notional Category of Modality. In H. Eikmeyer & H. Rieser (Eds.), *Words, Worlds, and Contexts: New Approaches in Word Semantics* (pp. 38–74). Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110842524-004

Krug, M. G. (2000). *Emerging English modals: a corpus-based study of grammaticalization.* Mouton de Gruyter.

Lacomba, C. (2023). Mapa hispano de los Estados Unidos 2022. *Estudios del Observatorio / Observatorio Studies, 84*, 1–123. https://cervantesobservatorio.fas.harvard.edu/es/informes

Lambert, W. E. (1955). Measurement of the linguistic dominance in bilinguals. *Journal of Abnormal and Social Psychology, 50*, 197–200.

Lanza, E. (2007). Multilingualism and the family. In L. Wei & P. Auer (Eds.), *Handbook of Multilingualism and Multilingual Communication* (pp. 45–68). Berlin: Mouton de Gruyter.

Lau, J. H., Clark, A., & Lappin, S. (2017). Grammaticality, Acceptability, and Probability: A Probabilistic View of Linguistic Knowledge. *Cognitive Science*, *41*(5), 1202–1241. https://doi.org/10.1111/cogs.12414

Leech, G. (2003). Modality on the move: The English modal auxiliaries 1961-1992. In R. Facchinetti, F. Palmer & M. Krug (Eds.), *Modality in Contemporary English* (pp. 223–240). Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110895339.223

Leech, G. (2013) Where have all the modals gone? An essay on the declining frequency of core modal auxiliaries in recent standard English. In J. I. Marín-Arrese, M. Carretero, J. Arús Hita, & J. Van der Auwera (Eds.), *English Modality* (pp. 95–116). De Gruyter, Inc.

Leech, G. N., Rayson, P., & Wilson, A. (2014). *Word frequencies in written and spoken English: based on the British National Corpus*. Routledge. https://doi.org/10.4324/9781315840161

Lenneberg, E. (1967). *Biological Foundations of Language*. New York NY: Wiley.

Liceras, J. M., Spradlin, K.T., Senn, C., Sikorska, M., Fernández Fuertes, R., de la Fuente, & E.A., (2003). *Second language acquisition and bilingual competence: the Grammatical Features Spell-out Hypothesis.* European Association of Second Language Acquisition (EuroSLA-13), Edinburgh, United Kingdom, September 19–21, 2003.

Liceras, J. M., Spradlin, K.T., & Fernández Fuertes, R. (2005). Bilingual early functional–lexical mixing and the activation of formal features. *International Journal of Bilingualism, 9*(2), 227–252.

Liceras, J. M, Martínez, C., Pérez-Tattam, R., Perales, S., & Fernández Fuertes, R. (2006). L2 acquisition as a process of creolization: insights from child and adult code-mixing. In C.

Lefebvre, L. White & Ch. Jourdan (Eds.), *L2 Acquisition and Creole Genesis: Dialogues*. John Benjamins, Amsterdam and Philadelphia.

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

Liceras, J. M., & Senn, C. (2009). Linguistic Theory and the Analysis of Minority Languages: Native, Immigrant and Heritage Spanish. *Lengua y migración, 1*(1), 39–73.

Liceras, J. M. (2017) Herramientas para abordar el análisis de la gramática no nativa de los inmigrantes. In D. G. Níkleva (Ed.), *La formación de los docentes de español para inmigrantes en distintos contextos educativos* (pp. 93–138). Bern: Peter Lang AG

Linck, J. A., & Cunnings, I. (2015). The Utility and Application of Mixed-Effects Models in Second Language Research. *Language Learning*, *65*(S1), 185–207. https://doi.org/10.1111/lang.12117

Lipski, J. M. (1978). Code-switching and the problem of language competence. In M. Paradis (Ed.), *Aspect of bilingualism* (pp. 250–264). Durham, SC: Hornbeam Press.

Lloyd-Smith, A., Gyllstad, H., Kupisch, T., & Quaglia, S. (2021). Heritage language proficiency does not predict syntactic CLI into L3 English. *International Journal of Bilingual Education and Bilingualism, 24*(3), 435–451. https://doi.org/10.1080/13670050.2018.1472208

Looney, D., & Lusin, N. (2019). Enrollments in languages other than English in United States institutions of higher education. *Modern Language Association* (June). https://www.mla.org/content/download/110154/file/2016-Enrollments-Final-Report.pdf

López, L., Alexiadou, A., & Veenstra, T. (2017). Code switching by phase. *Languages, 2*(3), 9. https://doi.org/10.3390/languages2030009

López, L. (2020). *Bilingual grammar: toward an integrated model*. Cambridge University Press.

López Ornat, S., Fernández, A., Gallo, P., & Mariscal, S. (1994) *La adquisición de la lengua española.* Siglo Veintiuno de España Editores, S.A.

Love, R., & Curry, N. (2021). Recent change in modality in informal spoken British English: 1990s–2010s. *English Language and Linguistics, 25*(3), 537–562. https://doi.org/10.1017/S1360674321000265

Loewen, S. (2009). Grammaticality judgment tests and the measurement of implicit and explicit L2 knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 94–112). Tonawanda, NY: Multilingual Matters. Mandell, O.

MacSwan, J. (1999). *A minimalist approach to intrasentential code switching.* New York: Garland Press.

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

MacSwan, J. (2009) Generative approaches to code-switching. In B. E. Bullock & A. J. Toribio (Eds.), *Cambridge Handbook of Linguistic Code-Switching* (pp. 309–335). New York: Cambridge University Press

MacSwan, J. (2014a) Code switching and Grammatical Theory. In T. K. Bhatia & W. Ritchie (Eds.), *Handbook of bilingualism and multilingualism* (pp. 322–350). Oxford: Blackwell. https://doi-org.proxy.bib.uottawa.ca/10.1002/9781118332382.ch2

MacSwan, J. (2014b) Programs and proposals in codeswitching research: unconstraining theories of bilingual language mixing. In J. MacSwan (Ed.), *Grammatical theory and bilingual codeswitching* (pp. 1–53). Cambridge: MIT Press.

MacSwan, J., & McAlister, K. T. (2010). Naturalistic and Elicited Data in Grammatical Studies of Codeswitching. *Studies in Hispanic and Lusophone Linguistics*, *3*(2), 521–532. https://doi.org/10.1515/shll-2010-1085

Mahootian, S. (1993). *A null theory of codeswitching*. [Doctoral dissertation, Northeastern Illinois University]. ProQuest Dissertations Publishing.

Mahootian, S., & Santorini, B. (1996). Code Switching and the Complement/Adjunct Distinction. *Linguistic Inquiry, 27*(3), 464–479.

Mahootian, S. (2019). *Bilingualism*. Taylor and Francis.

McCarty, T. L. (2002). *A place to be Navajo: Rough Rock and the struggle for self-determination in indigenous schooling*. Lawrence Erlbaum Associates. https://doi.org/10.4324/9781410602503

McClure, E. (1981). Formal and functional aspects of the codeswitched discourse of bilingual children. In R. P. Durán (Ed.), *Latino language and communicative behavior*. Norwood, NJ: ABLEX Publishing Corporation.

Mendoza Zambrano, M. (2002). *Satanás*. Seix Barral.

Meisel, J. M. (1989). Early differentiation of languages in bilingual children. In K. Hyltenstam & L. Obler (Eds.), *Bilingualism across the lifespan. Aspects of acquisition, maturity, and loss* (pp. 13–40). Cambridge University Press.

Meisel, J. M. (1994). Code-switching in young bilingual children: The acquisition of grammatical constraints. *Studies in Second Language Acquisition, 16*(4) 413–440.

Meisel, J. M. (2001). The simultaneous acquisition of two first languages. *Early differentiation and subsequent development of grammars*. In J. Cenoz & F. Genesee (Eds.), Trends in Bilingual Acquisition (pp. 11–42). Amsterdam: John Benjamins

Migge, B. (2015). The role of discursive information in analyzing multilingual practices. In *Code-Switching Between Structural and Sociolinguistic Perspectives, 43*, (185–206). De Gruyter. https://doi.org/10.1515/9783110346879.185

Montrul, S. (2008). *Incomplete acquisition in bilingualism: re-examining the age factor*. John Benjamins Pub. Co.

Montrul, S. (2014). Bilingualism and the Heritage Language Speaker. In T. K. Bhatia & W. C. Ritchie (Eds.), *The Handbook of Bilingualism and Multilingualism* (pp. 168–189). Oxford, UK: Blackwell Publishing.

MsGISRocker, (2020). https://stats.stackexchange.com/users/44862/msgisrocker *Which variance inflation factor should I be using: GVIF or GVIF^(1/(2*Df))* https://stats.stackexchange.com/users/44862/msgisrocker

Mubarak, R. (2015). *Politeness strategies in different varieties of English: De- and encoding of modality in speech acts from a translational, editorial and transcultural perspective considering methods of automatic text processing*. [Doctoral dissertation, University of Vienna]. University Library.

Müller, N., & Cantone, K. F. (2009). Language mixing in bilingual children: code-switching? In B. E. Bullock & A. J. Toribio (Eds.), *The Cambridge Handbook of Linguistic Code-switching* (pp. 199–220). Cambridge University Press. https://doi.org/10.1017/CBO9780511576331.013

Murphy, V. A. (1997). The effect of modality on a grammaticality judgement task. *Second Language Research*, *13*(1), 34–65. https://doi.org/10.1191/026765897671676818

Myers-Scotton, C. (1993). *Duelling Languages*. Oxford: Clarendon Press.

Myers-Scotton, C. (2002). *Contact Linguistics*. Oxford: Oxford University Press.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development, 38*, 1–135. University of Chicago Press for the Society for Research in Child Development.

Nicoladis, E., & Secco, G. (1998). The role of translation equivalents in a bilingual family's code-switching. In A. Greenhill, M. Hughes, H. Littlefield, & H. Walsh (Eds.), *Proceedings of*

the 22nd annual Boston University Conference on Language Development. Somerville, MA: Cascadilla Press.

Nishimura, M., & Yoon, K. K. (1998). Head directionality and intrasentential code-switching: a study of Japanese Canadians and Korean Americans' bilingual speech. In D. Silva (Ed.), *Japanese/Korean Linguistics 8* (pp. 121–130). CSLI Publications.

Olarrea, A. (2010). La estructura de la oración: sintaxis. In J. I. Hualde, A. Olarrea, A. M. Escobar, & C. E. Travis (Eds.), *Introducción a la Lingüística Hispánica.* (2nd ed., pp. 201–278). Cambridge University Press.

Otaola Olano, C. (1988). La modalidad (con especial referencia a la lengua española). *Revista de filología española, 68*(1), 97–117. https://doi.org/10.3989/rfe.1988.v68.i1/2.414

Oxford Advanced Learner's Dictionaries. (n.d.). https://www.oxfordlearnersdictionaries.com/definition/english/maybe_1

Padilla, A., & Liebmann, E. (1975). Language acquisition in the bilingual child. *The Bilingual Review/La Revista Bilingüe 2*, 34–55.

Padilla, A., & Lindholm, K. (1984). Child bilingualism: the same old issues revisited. In J. L. Martinez, & R. H. Mendoza (Eds.), *Chicano Psycology* (2nd ed., pp. 369–408). New York: Academic Press.

Palmer, F. (1965). *A linguistic study of the English verb*. London: Longman, Green and Co. Ltd.

Palmer, F. (2001). *Mood and modality* (2nd ed.). Cambridge University Press.

Panagiotidis, P. (2015). *Categorial features. A generative theory of word class categories*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139811927

Panhwar, F., & G. A. Buriro. (2020). An Overview of Theories and Approaches to Code-Switching. *International Research Journal of Arts & Humanities, 48*(48), 265.

Penfield, W. (1953). A consideration of the neurophysiological mechanisms of speech and some educational consequences. *Proceedings of the American Academy of Arts and Sciences*, 82, 199–214.

Penfield, W., & Lamar R. (1959). *Speech and Brain Mechanisms*. Princeton NJ: Princeton University Press.

Pfaff, C. W. (1979). Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish–English. *Language (Baltimore), 55*(2), 291–318. https://doi.org/10.2307/412586

Picallo, M. C. (1990). Modal Verbs in Catalan. *Natural Language and Linguistic Theory, 8*(2), 285–312. https://doi.org/10.1007/BF00208525

Pihler Ciglic, B. (2014). Los marcadores verbales de modalidad epistémica y su papel en el desarrollo de las competencias comunicativas. *Linguistica (Ljubljana), 54*(1), 381–395. https://doi.org/10.4312/linguistica.54.1.381-395

Polinsky, M., & Kagan, O. (2007). Heritage Languages: In the "Wild" and in the Classroom. *Language and Linguistics Compass, 1*(5), 368–395. https://doi.org/10.1111/j.1749-818X.2007.00022.x

Polinsky, M. (2018). *Heritage languages and their speakers.* Cambridge University Press.

Poplack, S. (1980). Sometimes I´ll start a sentence in Spanish y termino en español: toward a typology of code-switching. *Linguistics, 18*(7), 581–818.

Poplack, S. (1981). Syntactic structure and social function of codeswitching. In R. P. Duran (Ed.) *Latino language and communicative behavior* (pp. 169–184). Norwood, N.J.: Ablex.

Portner, P. (2009). *Modality*. Oxford University Press.

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rabadán, R. (2006) Modality and Modal Verbs in contrast: Mapping Out a Translation(Ally) Relevant Approach English–Spanish. *Languages in Contrast: International Journal for contrastive Linguistics 6*(2), 261–306.

Radford, A. (1988) *Transformational Grammar: A First Course*. Cambridge University Press.

RAE-ASALE. (2005, July 26) *Diccionario panispánico de dudas*. Retrieved from https://www.rae.es/dpd/deber

RAE-ASALE. (2011). *Nueva gramática básica de la lengua española*. Espasa Libros.

RAE-ASALE. (2014) Puede que. In RAE-ASALE, *Diccionario de la lengua española (DLE)*. Retrieved from poder | Definición | Diccionario de la lengua española | RAE–ASALE

Ridruejo Alonso, E. (1999). Modo y modalidad. El modo en las subordinadas sustantivas. In I. Bosque & V. Demonte (Eds.) *Gramática descriptiva de la lengua española* (Vol. 2., pp. 3209–3252). Madrid: Espasa Calpe.

Rizzi, L. (1976) Ristrutturazione. *Rivista di grammatica generativa, 1*, 1–54.

Rizzi, L. (1978) A Restructuring Rule in Italian Syntax. In. S. J. Keyser (Ed.), *Recent Transformational Studies in European Languages*, (pp.113–158). Cambridge (Mass.), MIT Press.

Rizzi, L. (1982) *Issues in Italian Syntax*. Dordrecht, Foris

Rizzi, L. (2000). *Comparative syntax and language acquisition*. Routledge.

Romaine, S., (1995). *Bilingualism* (2nd ed.). Blackwell.

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, *69*, 371–380. https://doi.org/10.1016/j.chb.2016.12.033

Sankoff, D., Poplack, S., & Vanniarajan, S. (1990). The case of the nonce loan in Tamil. *Language Variation and Change* (2), 71–101.

Schmid, M., & Köpke, B. (2007). Bilingualism and attrition. In M. Schmid, B. Köpke, M. Keijzer, & L. Weilmar (Eds.), *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues* (pp. 1–8). Amsterdam: John Benjamins.

Schroten, J. (1974). En torno a los verbos perifrásticos del español: un análisis sintáctico transformacional. *Revista de Filología Española 57*(1/4), 35–63.

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the full transfer/full access hypothesis. *Second Language Research, 12*, 40–72.

Schwartz, B. D., & Sprouse, R. A. (2000). When syntactic theories evolve: Consequences for L2 acquisition research. In J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp. 156–186). Malden, MA: Blackwell.

Schwartz, B. D., & Sprouse, R. A. (2012). Generative approaches and the poverty of the stimulus. In J. R. Herschensohn & M. Young-Scholten (Eds.), *The Cambridge Handbook of Second*

Language Acquisition (pp. 137–158). Cambridge University Press. https://doi.org/10.1017/CBO9781139051729.010

Selinger, H. (1996) Primary language attrition in the context of bilingualism. In W. Ritchie & T. Bhatia (Eds.), *The Handbook of Second Language Acquisition* (pp. 605–625). New York: Academic Press.

Serratrice, L. (2014) The Bilingual Child. In T.K. Bhatia and W.C. Ritchie (Eds.), *The Handbook of Bilingualism and Multilingualism* (pp. 87–108). Blackwell Publishing. https://doi-org.proxy.bib.uottawa.ca/10.1002/9781118332382.ch2

Shiu, L. J., Yalçın, Ş., & Spada, N. (2018). Exploring second language learners' grammaticality judgment performance in relation to task design features. *System*, *72*, 215–225.

Silva-Corvalán, C. (1994). *Language Contact and Change. Spanish in Los Angeles*. Oxford, UK: Oxford University Press.

Silva-Corvalán, C. (2014*). Bilingual Language Acquisition: Spanish and English in the First Six Years*. Cambridge University Press.

Spanish National Institute of Statistics (2016, November 30). *Encuesta sobre la Participación de la Población Adulta en las Actividades de Aprendizaje*. [Press release] Retrieved from Encuesta sobre la Participación de la Población Adulta en las Actividades de Aprendizaje. Año 2016. (ine.es)

Spanish Ministry of Education and Professional Formation. (2020) *Estadística de las enseñanzas no universitarias curso 2021-2022. Enseñanza de lenguas extranjeras*. [Press release] Retrieved from GRUPO TÉCNICO DE ESTADÍSTICA (educacionyfp.gob.es)

Sprouse, J. (2020). A user´s view of the validity of acceptability judgments as evidence for syntactic theories. In S. Schindler, A., Drożdżowicz, & K. Brøcker (Eds.), *Linguistic Intuitions: Evidence and method* (pp. 215-232). Oxford University Press.

Sprouse, J., & Almeida, D. (2017). Design Sensitivity and Statistical Power in Acceptability Judgment Experiments. *Glossa (London), 2*(1), 1. https://doi.org/10.5334/gjgl.236

Stadthagen-González, H., López, L., Parafita Couto, M. C., & Párraga, C. A. (2018). Using two-alternative forced choice tasks and Thurstone's law of comparative judgments for code-switching research. *Linguistic Approaches to Bilingualism, 8*(1), 67–97. https://doi.org/10.1075/lab.16030.sta

Statistics Canada (2022, August). *While English and French are still the main languages spoken in Canada, the country's linguistic diversity continues to grow*. Retrieved from https://www150.statcan.gc.ca/n1/daily-quotidien/220817/dq220817a-eng.htm

Statistics Canada (2023, March). *Indigenous languages across Canada*. Retrieved from https://www.census.gc.ca/census-recensement/2021/as-sa/98-200-X/2021012/98-200-X2021012-eng.cfm

Stevens, G. (2006). The Age-Length-Onset Problem in Research on Second Language Acquisition Among Immigrants. *Language Learning, 56*(4), 671–692. https://doi.org/10.1111/j.1467-9922.2006.00392.x

Sutter, J. C., & Johnson, C. J. (1990). School-Age Children's Metalinguistic Awareness of Grammaticality in Verb Form. *Journal of Speech, Language, and Hearing Research*, *33*(1), 84–95. https://doi.org/10.1044/jshr.3301.84

Swain, M. (1972). *Bilingualism as a First Language*. [Unpublished Doctoral dissertation, University of California].

Taeschner, T. (1983). *The sun is feminine: A study on language acquisition in bilingual children*. Berlin: Springer

Timm, L. A. (1975). Spanish–English Code-Switching: El Porqué y How-Not-To. *Romance Philology, 28*(4), 473–482.

Topor, M. (2005). Criterios identificadores de las perífrasis verbales en español. *Sintagma, 17*, 51–69.

Toribio, A. (2001). On the emergence of bilingual code-switching competence. *Bilingualism: Language and Cognition, 4*(3), 203-231. doi:10.1017/S1366728901000414

Twaddell, W. F. (1963). *The English verb auxiliaries* (2nd ed.). Providence: Brown University Press.

Unsworth, S., Chonogianni, V., & Skarabela, B. (2018). Experiential measures can be used as a proxy for language dominance in bilingual language acquisition research. *Frontiers in Psychology, 9*, 1809. https://doi.org/10.3389/fpsyg.2018.01809

Vanacore, A., Pellegrino, M. S., & Ciardiello, A. (2022). Fair evaluation of classifier predictive performance based on binary confusion matrix. *Computational Statistics, 39*(1), 363–383. https://doi.org/10.1007/s00180-022-01301-9

Vafaee, P., Suzuki, Y., & Kachisnke, I. (2017). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition, 39*(1), 59–95. https://doi.org/10.1017/S0272263115000455

Vargas Llosa, M. (2000). *La Fiesta del Chivo*. Grupo Santillana de Ediciones.

Vivanco, M. (2019) To be or not to be an auxiliary verb: the case of Spanish Poner(se) a + infinitive. *Borealis: An International Journal of Hispanic Linguistics, 8*(1), 35–54. http://dx.doi.org/10.7557/1.8.1.4655

Volterra, V., & Taeschner, T. (1978). The acquisition and the development of language by bilingual children. *Journal of Child Language, 5*, 311–326.

Wright, G. H. Von. (1951). *An essay in modal logic*. North-Holland Pub. Co.

Wapole, C. (2000) The Bilingual child: One system or two? In E. V. Clark (Ed.), *The Proceedings of the 30th Annual Child Language Research Forum* (pp. 187–194). Stanford: CSLI.

Wei, L. (2000). Dimensions of bilingualism. In L. Wei (Ed.), *The bilingualism reader* (pp. 3–25). London & New York: Routledge.

Wei, L. (2014). Conceptual and Methodological Issues in Bilingualism and Multilingualism Research. In T.K. Bhatia & W.C. Ritchie (Eds.), *The Handbook of Bilingualism and Multilingualism* (pp. 26–51). Blackwell Publishing. https://doi-org.proxy.bib.uottawa.ca/10.1002/9781118332382.ch2

Winter, B. (2019). *Statistics for linguists: an introduction using R*. Routledge.

Zyzik, E. C., & Sánchez, R. A. (2019). Beyond accuracy: Heritage speakers' performance on two kinds of acceptability judgment tasks. *Applied Psycholinguistics, 40*(3), 645–671. https://doi.org/10.1017/S0142716418000760

Appendix A. Modal infinitival constructions: longitudinal corpus

Child A

Age	Sentence
6;5	*Hay que* taper *esto*
6;6	*Vamos a* don´t forget *papá*.
6;7	*Vamos a* play *aquí*
6;7	*Podemos* share, don´t be rude.
6;7	*Hay que* bend it.
6;8	*Necesitamos* finish all my math
6;8	¿*Y se puede* bend?
6;8	*Lo voy a* kick
6;8	*Lo voy a* pressar
6;8	¿*Puedo* let it go?
7;1	*Ha dicho que los puedes* keep
7;1	*Yo quiero* write *eso*
7;2	*Lo quiero* sharpen
7;2	*Voy a sharpen los* pencils
7;2	*Necesitas* leave un gap
7;2	¿*Vas a* skate *después de yo*?
7;2	*No puede* bounce *aquí*
7;2	*Necesito* sharpen *mi* pencil
7;2	*No lo puedo* sharpen
7;2	*Necesitas* sweat
7;3	*Nadie va a* score *a mí ahora*
7;3	*Vamos a* play *la pelota*
7;3	*Solo hay que* touch the floor *una sola vez*
7;3	*La pelota no puede* touch the floor
7;4	¿*Puedo* keep?
7;4	¿*Puedes* press ‘go’?

7;4 *Hay que* sharpen *eso*

7;4 ¿*Necesito* sharpen *los* pencils?

7;4 *Mira lo que necesito* sharpen

7;4 *Yo quiero* sharpen

7;4 *Tengo que* play it

7;4 ¿*Podemos* keep it? *Es muy bonito*

7;4 ¿*Vas a* check *cuanto vale*?

7;5 ¿*Puedo* try it?

7;5 *Yo quiero* play

7;7 *Hay que* cook something *para ti*

7;7 *Si quieres, puedes* have a drink

7;8 *Ni lo puedo* handle it

7;8 *Hay que* stretch the neck

7;10 *Yo quiero* check

7;10 *Necesitas* hurry up

7;10 *Puedes* switch directions

7;10 *Puedes* change directions

7;11 *Voy a* pop-ar-las (explode bubbles)

7;11 *Voy a* sharpen *tu* pencil

7;11 *Voy a* move on

7;11 *El mío necesita* sharpen, *mira*

7;11 *No puedes* change directions

8;0 *Puede* spin all the way around

8;1 ¿*Puedes* keep *eso*?

8;3 ¿*Puedes* stand up?

8;3 *No puedes* hold on

8;3 ¿*Puedo* mix it?

8;3 *Podemos* play algo

8;3 *Intentamos* shoot it *con* guns

8;5 *Voy a* be nice

8;5 *Esto es lo que no me gusta, cuando esto y esto quieren* switch spots

8;5 ¿*Puedes* fix *mi* reloj?

8;5 *Intenta* score

8;5 *Vamos* a play

8;5 ¿*Me puedes* save *esto*?

8;5 Solo quiero fix esto

8;5 *Yo quiero* climb *ese* tree

8;6 *No puedes* change directions

8;6 *Vamos a* run in circles

8;6 ¿*Puedes* check *este* box?

8;6 *Voy a* hold it

8;6 *Yo quiero* keep *esa oruga*

8;6 *Hay que* pomp *este* ball

8;6 *No hay que* rush

8;6 *Puedes* breathe *ahora*

8;6 *Voy a* shake it

8;6 *Hay que* press hard

8;6 *Voy a* hold it

8;6 ¿*Puedes* untie *este* knot?

8;6 *Quiero* ask *algo*

8;6 *No hay que* flat*lo*

8;6 *Voy a* turn it on

8;6 *No puedes* step *en ese cojín*

8;6 ¿*Puedes* hold on *esto*?

8;6 *Puedes empezar a* catch me

8;6 *Vamos a* play tag

8;6 *Lo podemos* keep

Child B

Age	Sentence
2;7	¿*Puedes* sit down?
2;7	(Yo) *quiere* sit
2;10	*Quiero* press the button
3;0	*Yo quiero* wait
3;5	¿*Puedo* play *con esto, por favor*?
3;5	*Yo no quiero* to sleep
3;5	*Yo quiero* drink *agüita*
3;5	*Voy a* clean up my mess *rápido*
3;5	*Yo quiero* help
3;5	*Quiero* water *las plantas*
3;6	*Vamos a* mix it, ¿*te parece*?
3;6	*Tú puedes* touch the floor
3;6	*Yo no quiero* touch the floor
3;6	*Después de esto*, ¿*puedo* have *esto*?
3;7	*Mamá*, ¿*puedo* sit next to you?
3;7	¿*Puedo* mix it?
3;7	*Yo quiero* cut *eso*
3;7	*Necesitas* get off the floor
3;7	*Yo voy a* make *un* rocket ship
3;7	*Quiero* play *con las cartas*
3;7	¿*Puedo* play *con las cartas contigo*?
3;7	¿*Quieres* watch me?
3;7	¿*Puedo* sit next to _____?
3;7	¿*Puedo* have *eso*?
3;8	¿*Puedo* have *el* toy?
3;8	*Quiero* sit *en tu* lap
3;8	¿*Puedes* lift up?
3;8	*Puedo* try *uno y ya está*

Age	Sentence
2;7	¿*Puedes* sit down?
3;8	*Yo quiero* clean it
3;8	*Puedes* make *el* puzzle
3;8	¿*Puedo* take *un* bite?
3;8	¿*Me puedes* lift up *a tus* eyes?
3;8	¿*Puedo* cut *tu* hair?
3;8	*Lo voy a* take it off
3;9	¿*Puedo* have a turn?
3;10	*Estoy* hiding *donde puedo* hide
3;11	*Puedes* bring it up
3;11	¿*Puedo* mix it?
4;0	*Voy a* erase it
4;2	*Yo quiero* bring *los* gloves *que no son míos*
4;3	*Quiero* smell it again
4;3	*Yo quiero* pick some 'bichis' (stuffed toys)
4;3	*No puedo* hang it *en ningún sitio*
4;3	*Podíamos* watch movies
4;3	*Si quieren* come *a mi* house
4;3	¿*Cómo me* puedo stand up?
4;4	¿*Cómo puedo* climb?
4;5	*Yo quiero* smell mine
4;5	*El lobo quiere* **manger des cochons**
4;5	*Voy a* buy more bunny-bunnies *para* _____
4;5	*Quiero* check it out
4;5	*Voy a* change directions
4;5	¿*Puedo* take off the jacket?
4;6	¿*Podemos* bring *algo*?
4;6	¿*Puedes* grab *este* side?
4;6	¿*Puedes* move *tu* leg?
4;6	*Quiero* climb *esas* rocks

Age	Sentence
2;7	¿*Puedes* sit down?
4;6	*El teacher puede* find me *pero no mis* friends
4;6	*Voy a* check
4;6	*Podemos* play *un* game
4;6	¿*Puedes* back up?
4;7	¿*Me puedes* hold *el* bebé?
4;7	*Yo quiero* take esto
4;7	¿*Puedes* sharpen this?
4;7	*Puedes* take eso
4;7	¿*Puedo* sing *un* song?
4;7	*Quiero* sing *eso todo el día*
4;7	*No quiero* guess
4;7	¿*Cómo puedo* bend it?
4;7	*Necesitas* take *eso*
4;8	¿*Cómo lo podemos* hold?
4;8	*Hay un* chicken *que podemos* pet it
4;8	*Yo quiero* talk *de mi campamento*
4;8	¿*Puedes* erase*lo*?
4;8	*Puedes* bounce
4;8	¿*Puedo* press *el* yellow?
4;8	*Voy a* hug it
4;8	*No puedes* draw kiwis
4;8	*En ese papel vamos a* draw kiwis *y uvas*
4;8	*Voy a* share
4;8	¿*Puedes* tickle me?
4;8	¿*Puedo yo* tie-*las sola*?
4;8	*Puedo* cut *estos* baloons
4;8	¿*Puedes* stop messing *mi* pelo?

Appendix B. Sample items of the Grammaticality Judgment task

Practice items

Grammatical sentences

1. La jirafa ha comido.
2. La rana es verde.
3. Los gatos duermen mucho.
4. El abuelo compra pan.

Ungrammatical sentences

5. Los niños están jugado.
6. El escarabajo es vivo.
7. Los chica tiene frío.
8. El rinocerontes tiene hambre.

Experimental items

Grammatical sentences

Infinitive replacement (what)

9. ¿Qué desea la cabra?
10. ¿Qué prefiere el mono?
11. ¿Qué prefiere el conejo?
12. ¿Qué desea la rana?
13. ¿Qué prefiere el zorro?
14. ¿Qué desea la gaviota?

15. ¿Qué prefiere el arquitecto?
16. ¿Qué desea la actriz?

Infinitive replacement (demonstrative)

17. Empezar el trabajo, eso desea el cliente.
18. Ganar ese premio, eso desea mi compañero.
19. Escribir las cartas, eso prefiere la secretaria.
20. Beber la leche, eso prefiere mi hija.
21. Adivinar el futuro, eso desea mi jefa.
22. Regalar su ropa, eso prefiere el músico.
23. Cambiar de casa, eso desea mi hija.
24. Evitar el peligro, eso prefiere el inspector.

'That' sentences

25. Mi hijo prefiere que su tía compre unos tomates.
26. Mi hija desea que su novio encuentre las entradas.
27. Mi compañera prefiere que su amiga vea la película.
28. La vecina desea que su hijo limpie la habitación.
29. El gobierno prefiere que las elecciones sean cuanto antes.
30. El cliente desea que mi jefe tenga buenas ideas.
31. El abogado prefiere que el testigo venga más temprano.
32. El turista desea que la habitación tenga más camas.

Ungrammatical sentences

Infinitive replacement (what)

33. ¿Qué puede la viuda?
34. ¿Qué debe la serpiente?
35. ¿Qué puede tu primo?
36. ¿Qué debe el perro?
37. ¿Qué puede el arquitecto?
38. ¿Qué debe la mosca?
39. ¿Qué puede la economista?
40. ¿Qué debe el rebaño?

Infinitive replacement (demonstrative)

41. Contratar más personal, eso debe la vendedora.
42. Conseguir un trabajo, eso puede el novelista.
43. Comer un pescado, eso debe ese señor.
44. Comprar un regalo, eso puede esa señora.
45. Elaborar un informe, eso debe esa persona.
46. Enseñar la lección, eso puede el maestro.
47. Centrar la atención, eso debe el empresario.
48. Apagar la computadora, eso puede el portero.

‘That’ sentences

49. Mi madre puede que mi hermana regrese de Italia.
50. Mi hermano debe que su novia deje de fumar.
51. Su padre puede que mi amigo lea el libro.
52. Mi tía debe que mi prima haga los deberes.
53. El equipo puede que el director apruebe el proyecto.
54. La familia debe que mi abuela salga a pasear.
55. La profesora puede que el alumno busque los apuntes.
56. Mi hermana debe que su hija viva su vida.

Distractors

Grammatical sentences

Perfect tense

57. Las patatas, las he comido con gusto.
58. Los deberes, los he terminado esta mañana.
59. El examen, lo he aprobado sin problema.
60. La ciudad, la he visitado con mi hermana.

Passive voice

61. El edificio fue diseñado por mi hermana.
62. El pan fue horneado por mi padre.
63. La enferma fue tratada por el médico.
64. El ascensor fue arreglado esta mañana.

Present continuous

65. El agua, la estoy bebiendo.
66. El regalo, lo estoy abriendo.
67. El pulpo, lo estoy cocinando.
68. El perro, lo estoy cuidando.

Ungrammatical sentences

Perfect tense

69. El pájaro, he vístolo por la ventana.
70. La perla, he escondídola en mi bolso.
71. La comida, he comprádola en el supermercado.
72. Los guantes, he perdídolos en casa.

Passive voice

73. El cuadro lo fue pintado por un artista.
74. El libro lo fue escrito por un poeta.
75. La flor la fue plantada por mi abuela.
76. El mueble lo fue trasladado sin problema.

Present continuous

77. La araña, la estoy viéndola.
78. La pared, la estoy pintándola.
79. Los huevos, los estoy friéndolos.
80. La música, la estoy escuchándola.

Appendix C. Sample items of the Two-alternative Forced-Choice task

Practice items

1. I need to *visitar el medico.*

   I need *visitar al medico.*

2. I need *arreglar el reloj.*

   I need to *arreglar el reloj.*

3. *Necesita* choose a present.

   Necesita to choose a present.

4. *Necesita* to write the email.

   Necesita write the email.

5. He has to *ver a su madre.*

   He has *ver a su madre.*

6. She has *cambiar de coche.*

   She has to *cambiar de coche.*

7. *Él tiene que* see the house.

   Él tiene see the house.

8. *Ella tiene* water the plants.

   Ella tiene que water the plants.

Experimental items

Modal periphrases

English directionality

9. The partner must to *aumentar el negocio*

The partner must *aumentar el negocio*.

10. The gentleman can to *cenar en casa.*

    The gentleman can *cenar en casa.*

11. The expert must to *describir el experimento.*

    The expert must *describir el experimento.*

12. The candidate can to *ganar las elecciones.*

    The candidate can *ganar las elecciones.*

13. The boss must to *organizar la empresa.*

    The boss must *organizar la empresa.*

14. The student can to *sacar buenas notas.*

    The student can *sacar buenas notas.*

15. The doctor must to *rellenar un documento.*

    The doctor must *rellenar un documento.*

16. The girl can to *escalar la montaña.*

    The girl can *escalar la montaña.*

Spanish directionality

17. *La profesora debe* to give clear examples.

    La profesora debe give clear examples.

18. *La ganadora puede* to celebrate the victory.

    La ganadora puede celebrate the victory.

19. *La pianista debe* to buy that piano.

    La pianista debe buy that piano.

20. *El dueño puede* to sell the house.

El dueño puede sell the house.

21. *El arquitecto debe* to follow the protocol.

    El arquitecto debe follow the protocol.

22. *La diseñadora puede* to finish the dress.

    La diseñadora puede finish the dress.

23. *El abogado debe* to convince the jury.

    El abogado debe convince the jury.

24. *La niña puede* to keep her dolly.

    La niña puede keep her dolly.

Nominal infinitival clauses

English directionality

25. The worker prefers to *salir del edificio.*

    The worker prefers *salir del edificio.*

26. The model wishes to *proponer un cambio*.

    The model wishes *proponer un cambio*.

27. The scientist prefers to *seleccionar su equipo*.

    The scientist prefers *seleccionar su equipo*.

28. The president wishes to *celebrar su cumpleaños.*

    The president wishes *celebrar su cumpleaños.*

29. The director prefers to *superar el reto.*

    The director prefers *superar el reto.*

30. The queen wishes to *viajar al espacio.*

    The queen wishes *viajar al espacio.*

31. The pupil prefers to *mejorar su nota.*

    The pupil prefers *mejorar su nota.*

32. The speaker wishes to *hablar más claro.*

    The speaker wishes *hablar más claro.*

Spanish directionality

33. *El compañero prefiere* to call her friends.

    El compañero prefiere call her friends.

34. *El observador desea* to know the team.

    El observador desea know the team.

35. *El científico prefiere* to clean the laboratory.

    El científico prefiere clean the laboratory.

36. *La vecina desea* to design a dress.

    La vecina desea design a dress.

37. *La extranjera prefiere* to learn new languages.

    La extranjera prefiere learn new languages.

38. *El comprador desea* to lower the price.

    El comprador desea lower the price.

39. *El pescador prefiere* to sell his fish.

    El pescador prefiere sell his fish.

40. *La princesa desea* to see a dragon.

    La princesa desea see a dragon.

Periphrastic future

English directionality

41. I´m going to *viajar a Estados Unidos.*

    I´m going *viajar a Estados Unidos.*

42. We´re going to *comprar unas flores.*

    We´re going *comprar unas flores.*

43. He´s going to *cortar el césped.*

    He´s going *cortar el césped.*

44. I´m going to *leer un libro.*

    I´m going *leer un libro.*

45. We´re going to *ir a la fiesta.*

    We´re going *ir a la fiesta.*

46. She´s going to *comprar un regalo.*

    She´s going *comprar un regalo.*

47. I´m going to *celebrar mi cumpleaños.*

    I´m going *celebrar mi cumpleaños.*

48. We´re going to *pintar la casa.*

    We´re going *pintar la casa.*

Spanish directionality

49. *Voy a* drink juice.

    Voy drink juice.

50. *Vamos a* visit my parents.

Vamos visit my parents.

51. *Va a* send the letter.

    Va send the letter.

52. *Voy a* make the phone call.

    Voy make the phone call.

53. *Vamos a* buy some food.

    Vamos buy some food.

54. *Va a* study for the exam.

    Va study for the exam.

55. *Voy a* balance my budget.

    Voy balance my budget.

56. *Vamos a* win the contest.

    Vamos win the contest.

Distractors

57. *Los* elephants are big.

    Elephants are big.

58. *El* love is mysterious.

    Love is mysterious.

59. *Los* peppers are green.

    Peppers are green.

60. *El* water is transparent.

    Water is transparent.

61. *La* truth is essential.

Truth is essential.

62. *Las* poppies are red.

    Poppies are red.

63. *Las* butterflies are light.

    Butterflies are light.

64. *La* music is beautiful.

    Music is beautiful.

Appendix D. Summary of descriptive data

Table 34

Summary of Descriptive Data — Grammaticality Judgment Task

Type of Sentence	Grammaticality	Group	N	Mean	SD
What	Grammatical	Control	20	92.50	23.44
		Heritage English	10	93.75	13.50
		Heritage Spanish	10	100.00	0.00
		L1Eng–L2Esp Immigrants	10	92.50	19.72
		L1Esp-L2Eng Immigrants	14	98.21	6.68
		L1Eng–L2Esp Students	10	100.00	0.00
		L1Esp-L2Eng Students	22	95.45	13.11
	Ungrammatical	Control	20	60.62	38.96
		Heritage English	10	91.25	19.59
		Heritage Spanish	10	76.25	33.05
		L1Eng–L2Esp Immigrants	10	55.00	34.96
		L1Esp-L2Eng Immigrants	14	72.32	37.07
		L1Eng–L2Esp Students	10	80.00	30.16
		L1Esp-L2Eng Students	22	79.54	25.46
Demonstrative	Grammatical	Control	20	87.50	25.00
		Heritage English	10	68.75	37.85
		Heritage Spanish	10	78.75	27.03
		L1Eng–L2Esp Immigrants	10	80.00	30.73
		L1Esp-L2Eng Immigrants	14	78.57	29.99
		L1Eng–L2Esp Students	10	73.75	37.48
		L1Esp-L2Eng Students	22	86.93	24.84

Type of Sentence	Grammaticality	Group	N	Mean	SD
	Ungrammatical	Control	20	61.87	36.15
		Heritage English	10	78.75	32.83
		Heritage Spanish	10	61.25	35.08
		L1Eng–L2Esp Immigrants	10	75.00	31.73
		L1Esp-L2Eng Immigrants	14	57.36	42.92
		L1Eng–L2Esp Students	10	71.25	33.88
		L1Esp-L2Eng Students	22	57.95	33.75
That	Grammatical	Control	20	94.37	22.39
		Heritage English	10	97.50	5.27
		Heritage Spanish	10	96.25	8.44
		L1Eng–L2Esp Immigrants	10	96.25	8.44
		L1Esp-L2Eng Immigrants	14	98.21	4.54
		L1Eng–L2Esp Students	10	92.50	15.81
		L1Esp-L2Eng Students	22	96.59	8.78
	Ungrammatical	Control	20	99.37	2.79
		Heritage English	10	100.00	0.00
		Heritage Spanish	10	87.50	25.00
		L1Eng–L2Esp Immigrants	10	81.25	32.41
		L1Esp-L2Eng Immigrants	14	98.21	4.54
		L1Eng–L2Esp Students	10	91.25	11.86
		L1Esp-L2Eng Students	22	97.16	7.65

Note 1. N = Number of participants; SD = Standard Deviation

Table 35

Summary of Descriptive Data — Two-alternative Forced-Choice Task

Type of Construction	Directionality	Group	N	Mean	SD
PeriMod	English	Heritage English	10	0.00	0.00
		Heritage Spanish	10	3.75	11.86
		L1Eng–L2Esp Immigrants	10	1.25	3.95
		L1Esp-L2Eng Immigrants	14	14.29	30.17
		L1Eng–L2Esp Students	10	0.00	0.00
		L1Esp-L2Eng Students	22	6.82	19.94
	Spanish	Heritage English	10	1.25	3.95
		Heritage Spanish	10	7.50	15.81
		L1Eng–L2Esp Immigrants	10	15.00	23.42
		L1Esp-L2Eng Immigrants	14	30.36	31.67
		L1Eng–L2Esp Students	10	11.25	20.79
		L1Esp-L2Eng Students	22	7.95	18.72
NonPeriMod	English	Heritage English	10	85.00	21.89
		Heritage Spanish	10	70.00	30.73
		L1Eng–L2Esp Immigrants	10	76.25	31.98
		L1Esp-L2Eng Immigrants	14	61.61	33.78
		L1Eng–L2Esp Students	10	90.00	15.37
		L1Esp-L2Eng Students	22	67.04	28.49
	Spanish	Heritage English	10	48.75	29.14
		Heritage Spanish	10	85.00	21.08
		L1Eng–L2Esp Immigrants	10	50.00	40.82

Type of Construction	Directionality	Group	N	Mean	SD
		L1Esp-L2Eng Immigrants	14	66.96	29.26
		L1Eng–L2Esp Students	10	62.50	39.53
		L1Esp-L2Eng Students	22	69.89	29.54
PeriAsp	English	Heritage English	10	98.75	3.95
		Heritage Spanish	10	96.25	11.86
		L1Eng–L2Esp Immigrants	10	96.25	8.44
		L1Esp-L2Eng Immigrants	14	99.11	3.34
		L1Eng–L2Esp Students	10	100.00	0.00
		L1Esp-L2Eng Students	22	94.32	21.73
	Spanish	Heritage English	10	100.00	0.00
		Heritage Spanish	10	91.25	11.86
		L1Eng–L2Esp Immigrants	10	95.00	15.81
		L1Esp-L2Eng Immigrants	14	89.29	16.88
		L1Eng–L2Esp Students	10	98.75	3.95
		L1Esp-L2Eng Students	22	91.48	23.27

Note 1. N = number of participants; SD = Standard deviation