

ORIGINAL ARTICLE

Indirect structural crosslinguistic influence in early Catalan–Spanish bilinguals in adulthood: Predicate selection in Catalan existential constructions

Silvia Perpiñán^{1*}  and Adriana Soto-Corominas² 

¹Universitat Pompeu Fabra, Department of Translation and Language Sciences, Barcelona, Spain and

²Universitat Internacional de Catalunya, Department of Applied Linguistics, Barcelona, Spain

*Corresponding author. Email: silvia.perpinan@upf.edu

(Received 11 June 2020; revised 18 June 2021; accepted 18 June 2021; first published online 06 September 2021)

Abstract

This study reports an oral production experiment investigating the expression of existentiality in the Catalan of adult Catalan–Spanish early bilinguals ($N = 58$) with comparable proficiencies but different language dominance. The results show qualitative differences among the bilinguals in existential predicate selection and in their supply of partitive pronouns, modulated by language dominance. Balanced Bilinguals as well as Spanish-dominant bilinguals significantly produced more *estar* (in detriment of *ser-hi* and *haver-hi*) not only in locative contexts, where Catalan already presents optionality regulated by semantic differences, but also in existential constructions, where this optionality does not exist. We argue for *indirect* crosslinguistic influence (CLI), when the bilingual *perceives* certain structural overlap within constructions, mediating the influence from one structure to another one and expanding the limits of CLI. The qualitative differences found among bilinguals challenge the idea of a bilingualism continuum in Catalan–Spanish bilingualism with an identical mental representation.

Keywords: crosslinguistic influence; Catalan; bilingualism; existential sentences; *ser*; *estar*; *harver-hi*; partitive clitics; language dominance

The interaction between the two linguistic systems of the bilingual has been the object of much research, both in terms of its psycholinguistic consequences for the speaker (e.g., Müller & Hulk, 2001; Serratrice, 2013; Blom et al., 2017) and in terms of its long-term impact to the given languages in the shape of language change (e.g., Silva-Corvalán, 2002; Thomason & Kaufman, 1988; Meisel et al., 2013). However, little research has investigated the long-term effects of crosslinguistic influence (CLI) in highly proficient early bilinguals in adulthood born and raised

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

in cases of sustained societal and official bilingualism, and how language dominance modulates CLI in these cases. CLI refers to “the influence of a person’s knowledge of one language on that person’s knowledge or use of another language” (Jarvis & Pavlenko, 2008, p. 1) at the level of the individual as a psycholinguistic phenomenon, and it is usually regulated by certain structural and domain constraints. This study investigates CLI in the production of Catalan existential constructions by adult Catalan–Spanish bilinguals born, raised, and living in Catalunya, an officially bilingual region in the northeast of Spain.

Classifying bilinguals

In order to investigate bilingualism, and more precisely, how two languages (or more) cohabit and interact in the mind of a speaker, we first need to describe the degree of bilingualism in a speaker, since this has a clear effect in the interplay between the two languages. Nowadays, there is little debate about the dynamic and gradient nature of bilingualism (Birdsong, 2014; De Houwer, 2018; Grosjean, 1989, 2008, 2010; Luk & Bialystok, 2013; Schmid & Köpke, 2017). Yet, in research studies, there is a tendency to classify bilinguals into subgroups for the convenience, reliability, and scientific advantages that systematic comparisons across and within bilingual populations bring to our understanding of bilingualism. The classification of bilinguals, as well as the setting of cut-off points to create subgroups, is not exempt of controversy, for *degree of bilingualism*, *language dominance*, or *bilingual experience* are complex and multidimensional constructs, very difficult to operationalize and measure (Gertken et al., 2014; Luk & Bialystok, 2013; Schmid & Yılmaz, 2018; Silva-Corvalán & Treffers-Daller, 2015; Treffers-Daller, 2019). Language dominance refers to the relative strength and accessibility of the two languages of a bilingual and encompasses a series of internal and external linguistic variables such as language aptitude, proficiency, fluency, ease of processing, current and past use, age of onset of acquisition (AOA), amount of input and output, quality of input, cultural identification, linguistic environment, education, place of birth and residence, among others. In this study, we examine language performance in terms of some of the constituent dimensions of bilingual experience, in particular in terms of language identity, place of residence, and current input and output.

A large amount of research on early bilingualism has traditionally classified bilinguals according to their AOA, using it as an initial agent of language dominance or even as a way of measuring quantity of input (Montrul, 2014). While it is generally agreed that AOA is a crucial factor that impacts language knowledge later in life (Bylund et al., 2021; Hyltenstam & Abrahamsson, 2003; Meisel, 2009, 2018; Montrul, 2008, 2016b; Nicoladis, 2018), in a bilingual environment, AOA effects may be mitigated by input quality, input quantity, and cumulative exposure (Unsworth, 2016). For instance, Bedore et al. (2012) found that the current use accounted better for the data than AOA in their study with 1029 Spanish–English bilingual children. Similarly, research on heritage language development has repeatedly shown that the development of the family language is usually compromised, even though it is the first language of the bilinguals (see Montrul, 2016a; Polinsky, 2018 for recent overviews). A growing body of research on the development of minority

languages in bilingual societies has similarly found that exposure-related factors may have a stronger effect than AOA. For example, investigating the development of Welsh plural morphology by young simultaneous Welsh–English bilinguals who differed in the language spoken at home, Thomas et al. (2014) found that the performance in Welsh by children of bilingual (Welsh–English) homes was comparable to that of English-speaking children exposed to Welsh at school after age 4 and both of these groups were significantly different from Welsh-only-homes’ children (see also Gathercole & Thomas, 2009 and Soto-Corominas, 2020 for similar results in early bilinguals). While home language may exert an influence on the development in the minority language, other studies have found that language use outside the home may also impact bilingual development and outcomes (Kupisch & Weijer, 2015). In short, these studies on minority languages force us to reconsider the use of AOA as a predictor for bilingual performance in bilingual societies. From a point of view in which bilingualism is perceived as a dynamic and active process (Schmidt & Köpke, 2017), AOA should not be the determining factor to describe bilinguals, as later experience can have a greater impact on bilingualism outcomes in adulthood. Following this reasoning, we decided not to include AOA in our overall measurement of bilingual dominance. Indeed, when AOA is compared to measurements of current use, as we do in this study, it is shown that AOA would not reliably predict a fine-grained classification of bilinguals.

Another way of typically classifying bilinguals is according to their proficiency (Luk & Bialystok, 2013; Treffers-Daller, 2019), that is, according to their knowledge or competence of the two languages. This is an external measurement of the performance of the individual in certain tests or language assessments. As Montrul (2015) discusses, language proficiency is one dimension of language dominance that relates to linguistic ability, and although proficiency and dominance often correlate, they cannot be equated, as we can find bilingual individuals dominant in one language that are not necessarily highly proficient in that language (Birdsong, 2006). Indeed, in a bilingual society such as Catalunya (Spain), in which virtually everybody under the age of 40 has been educated in and exposed to both languages, language proficiency may indicate the educational level of the speaker rather than the relative weight of the two languages in the mind of the bilingual. So much that we can actually find highly proficient bilinguals in Catalan as assessed by the *Direcció General de Política Lingüística* who are not dominant or completely at ease in that language, despite their high explicit knowledge about it. For these reasons, we purposefully excluded proficiency as a measurement in the overall calculation of bilingual dominance and we demonstrate that self-rated proficiency fails to reliably distinguish among significant types of Catalan–Spanish bilinguals.

In a nutshell, we argue for a dynamic view of both bilingualism and language dominance, measured in a way that the relevant predictors are chosen according to the reality of the context. Thus, instead of applying a generic language dominance questionnaire, we created an ad hoc measurement, the Language Dominance Questionnaire (LaDoQ), taking into account the particularities of our target population and the potentially changing bilingual reality of our speakers. Thus, we give prominence to factors such as current language use in and outside the home, language identity, as well as quality of input determined by place of origin of the main input providers, as well as place of residence. This is the way we operationalize

language dominance. In this study, we asked whether *language dominance* predicts CLI in adulthood and whether adult early bilinguals with different degrees of language dominance share the same mental representation of a specific linguistic phenomenon: Catalan existential constructions.

Language dominance and CLI at the endstate of early bilingualism

Whereas a burgeoning body of research has investigated the relationship between language dominance and CLI in bilingual children (Serratrice, 2013; Kupisch, 2018), less is known about whether CLI may persist in early bilinguals into adulthood (Kupisch, 2012). The limited number of studies that have investigated morphosyntactic outcomes of early bilingualism in bilingual societies have found instances of fossilized CLI, particularly in the minority language. CLI in these cases has been usually found to be modulated by language dominance. For instance, Perpiñán (2017) found that adult Catalan–Spanish bilinguals, both Balanced Bilinguals (mostly simultaneous bilinguals) and Spanish-dominant bilinguals (early sequential bilinguals), fell short in the production of non-personal Catalan clitics compared to Catalan-dominant bilinguals and they displayed insensitivity to ungrammatical constructions. These results indicate that even early bilinguals present differences with other more dominant Catalan bilinguals and, as such, downplay the role of AOA. CLI might also be linguistically selective. In a bidirectional study in which Catalan-dominant and Spanish-dominant speakers were tested, Puig-Mayenco et al. (2018) found selective bidirectional CLI depending on the linguistic phenomenon: negative concord did not display any type of influence from one language to the other, whereas differential object marking (DOM) did. Puig-Mayenco et al. (2018) concluded that CLI might be dependent on whether the phenomenon presents ‘optionality’ (negative concord) or ‘obligatoriness’ (DOM), as CLI, they argued, may be less likely when there is an optional construction that does not compromise the grammaticality of the structure. This conclusion is somewhat at odds with the general idea that CLI serves as a relief strategy precisely for ambiguous input (Müller, 1998), and the fact that optionality is a characteristic of bilingual grammars (Sorace, 2005).

Following Serratrice (2013), we consider cases of CLI when there is evidence for the effect of one language on the other. Typically, CLI is reported in cases in which language A has a syntactic construction allowing for multiple grammatical analyses and input from language B reinforces one of those analyses. The most classic approaches to CLI locate it in areas that involve the C-system, that is, between syntax and other cognitive systems, and in particular in the discourse-syntax interface (Hulk & Müller, 2000; Müller & Hulk, 2001; Silva-Corvalán, 1993; Sorace, 2011). Also, these assume certain unidirectionality determined by the structural possibilities of the languages with respect to a particular overlapping linguistic phenomenon: the language that only accepts one syntactic analysis would influence the language that allows for two or more analyses, irrespective of language dominance. However, bidirectionality in CLI effects has been shown (Jarvis & Pavlenko, 2008; Pavlenko & Jarvis, 2002; Soto-Corominas, 2020; Zhou et al., 2020), as well as effects outside the discourse-pragmatics interface (Sorace & Serratrice, 2009; Silva-Corvalán &

Montanari, 2008). Another relevant consideration in the study of bilingualism is whether the observed variation, optionality, and/or CLI suppose a quantitative or a qualitative distinction with respect to other varieties. According to Meisel, Elsig and Rinke (2013), “[quantitative changes] reflect an increase or decrease in frequency of occurrence of constructions, while [qualitative changes] consist of the emergence of new forms or constructions or the elimination of previously attested ones. Only qualitative changes can plausibly be hypothesized to reflect restructurings of grammatical systems (. . .)”, p. 2. From a synchronic point of view then, qualitative differences could be hypothesized to describe current divergent underlying grammars and therefore, potential ongoing change. This is the approach we adopt in this study.

Thus, drawing from proposals specific for early bilingualism, ultimate attainment in SLA, as well as bilingualism in language contact situations, the purpose of this study is to describe the interaction, if any, of the two linguistic systems, Catalan and Spanish, the nature of CLI between the two systems, how CLI may be modulated by language dominance, as well as its scope and limitations in the Catalan of Catalan–Spanish bilinguals. Most accounts that try to explain CLI in different bilingualism situations (heritage language acquisition, L1 attrition, asymmetrical societal bilingualism) point toward three key factors that favor CLI (Putnam & Sánchez, 2013; Schmid & Köpke, 2017; Hicks & Domínguez, 2020): 1 – frequency of coactivation; 2 – early age of onset of bilingualism; and 3 – similarity of the linguistic systems. All these factors are present in the context of Catalan–Spanish bilingualism.

Linguistic background: The locative paradigm

Existential and locative constructions are a set of constructions that denote spatial relations (Ramos, 1998). According to this author, the main differences between locative and existential constructions are within the scope of information structure, that is, of pragmatic/discourse nature. Due to conceptual, semantic, and frequently also lexical overlap, the two constructions are often considered together (Freeze, 1992; Rigau, 1997). In the following sections, existential and locative constructions are described in detail for Central Catalan and Spanish.

Existential constructions

Existential constructions introduce a participant or an element in the discourse (Zeitoun et al., 1999). These constructions present non-canonical morphosyntax in both Catalan and Spanish (McNally, 2011; 2016). Existential constructions in Catalan are expressed with the existential predicate *haver-hi* ‘there is/are’ (1), which comes from a medieval possessive verb followed by a locative pronoun (Bassaganyas, 2015; McNally, 2016). Catalan *haver-hi* has the locative clitic *HI* lexicalized in the verbal form, and the predicate *haver* (i.e., without *HI*) does not exist (Rigau, 1994).

- (1) *Hi ha molta gent.*
HI HAV-3SG many people
 ‘There are many people.’ (lit. There is many people.)

Ser-hi, which is a copulative verb followed by the same locative clitic *hi*, can only be used with new information in very limited cases such as with proper nouns or with personal pronouns (Ramos, 1998). As such, in first mention existential constructions (i.e., presentational uses), *haver-hi*, and not *ser-hi*, is used with common nouns. As opposed to *haver-hi*, the clitic *HI* in *ser-hi* is not considered lexicalized (Rigau, 1994). With respect to semantic restrictions, when the theme of the existential structure is a definite phrase, *ser* and *haver* are both available in Catalan with a difference between proper nouns (2a) and common nouns (2b) with respect to the availability of *ser*, examples based on Ramos (1998). That is, Catalan does not seem to obey the *definiteness effect* constraint (Milsark, 1977), which clearly applies in Spanish and regulates the selection of *haber/estar* in Spanish (6), and hence, the choice between an existential structure and a locative one.

This complementary distribution in word order is not present in restrictive relative clauses, a syntactic construction in which information structure and definiteness effects are neutralized in both, Spanish and Catalan. In Catalan restrictive relative clauses, *haver-hi* (2c) is more frequent than *ser* (2d), but both are possible (Ramos, 2002).

- (2) a. En aquella assemblea hi havia {la Marta/una dona}.
 in that meeting, HI HAV-3SG-PAST {the Marta/a woman}
 ‘In that meeting, {Marta was present/there was a woman}’
- b. En aquella assemblea hi era {el Miquel/*un home}.
 in that meeting, HI SER-3SG-PAST {the Miquel/un home}
 ‘In that meeting, {Miquel was present/there was a man}’
- c. Les claus que hi ha a casa són meves
 the keys that HI HAV-3SG at home are mine
- d. Les claus que són a casa són meves
 the keys that SER-3PL at home are mine
 ‘The keys that are at home are mine.’

Estar is not available as an existential predicate in Standard Catalan. The constituent order in the existential use of *ser-hi* is fixed in Catalan: pivot (THEME/subject) – copula (3), the expression of the location is optional.

- (3) (A la casa), el gat no hi {és/*està}.
 at the house the cat neg. HI SER/ EST-3SG
 ‘The cat isn’t at the house.’ (Lit. ‘At the house, the cat isn’t there.’)

Since *ser-hi*, when used without an explicit locative phrase, is the equivalent to ‘to be there/to exist’, its use to refer to something for the first time without a topicalized locative phrase is very restricted. In fact, it is only used either to refer to people (4a) or to negate the existence of something which was pre-supposed by the speakers (4b):

- (4) a. El Joan hi és.
 the Joan HI SER-3SG
 'Joan is here/there.'
- b. No hi és la clau, on l' has posat?
 neg. HI SER-3SG, the key where it have-2SG put
 'The key is not (here), where have you put it?'

In short, Spanish existential constructions present three main differences from Catalan existentials: 1) Spanish only has one existential verb, *haber*, (5). That is, there is no equivalent of *ser-hi*; 2) no clitic is lexicalized in *haber* (in fact, Spanish lacks locative clitic HI altogether); and 3) the definiteness effect strongly applies in Spanish (6). As a result of the definiteness effect, if the theme to be localized is definite, it needs to appear in a locative syntactic construction in Spanish (6a vs. 6b), but if the theme is indefinite, it would be localized with an existential construction, as in (5).

- (5) **Hay** mucha gente.
 HAV-3SG/PL many people
 'There are many people.' (lit. There is many people.)
- (6) a. ***Hay** las llaves.
 HAV-3SG/PL the keys
 'There are the keys.'
- b. Las llaves **están** ahí.
 the keys EST-3PL there
 'The keys are there.'

Locative constructions: *estar* and *ser*

Locative constructions establish the location of a given element or participant. While both *ser* and *estar* 'to be' are possible in locative constructions in Catalan (7a-b), their exact distribution depends on a complex interplay of syntax, semantics, animacy of theme, and aspect (Ramos, 2002). The Grammar of the Institut d'Estudis Catalans (GIEC, 2016, p. 872) considers *ser* the prototypical verb for locative constructions, whereas *estar* adds an aspectual contribution such as duration or endpoint. Therefore, the distribution of *estar/ser*, as in Spanish, is determined by semantic (aspectual) differences. Nonetheless, it has been described that copula *estar* in Catalan is gaining ground in the locative context to *ser*, particularly in the southern Catalan varieties (Sanz & González, 1995) and some areas of Barcelona (Solà, 1994), as locative *estar* is losing its aspectual connotation of duration or endpoint. According to the GIEC, this *estar* extension is due to 'foreign influences', that is, to Spanish. The precise description of the distribution of *ser* and *estar* in Catalan locative constructions is difficult due to its current variation in the language; for the purpose of this study, we will laxly assume that present-day Catalan accepts both copulas to locate objects, the etymological *ser* and the innovative *estar*.

- (7) La pilota {és / està} a casa.
 the ball SER-3SG EST-3SG at home
 ‘The ball is/remains at home.’

Spanish employs the copula *estar* for all locative predicates (8a), except to locate eventive subjects (8b), which employs *ser* (Perpiñán et al., 2020).

- (8) a. La pelota **está** en casa.
 the ball EST-3SG at home
 ‘The ball is at home.’
 b. La fiesta es en el jardí.
 the party SER-3SG in the garden
 ‘The party is in the garden’

One commonality of locative constructions in Catalan and Spanish is that the theme or pivot (i.e., what is being located) is usually a definite determiner phrase, and the order of constituents is fixed: theme – copula – location (Ramos, 2000, 2002). If the order of the constituents changes into location – copula – theme, an existential construction obtains, which is in complementary distribution with the locative construction, and presents different predicate selection and information structure.

Negative and partial existentials

Given the nature of our experimental corpus, we also need to describe how locative and existential constructions work when negated or referred back in the discourse to modify or deny its presupposition. In this sense, “second mention constructions” are closely related to negative existentials, which commonly work differently from positive existentials in syntactic and morphological terms (Veselinova, 2013). To our knowledge, second mention existentials, semi-existentials, and negative existentials have not been investigated in depth in Catalan (or Spanish) and have one formal description in semantic and syntactic terms (i.e.: the availability of a type/token reading), comparing Catalan and Spanish (Fàbregas, 2019). For this study, we will limit our presentation to descriptive data.

Ser-(hi)/Estar

As in locative first mention constructions, the use of *ser* (standard) or *estar* (common) can be used to refer to any antecedent in the discourse for a second time, as long as the location is specified (9a). If the location is not specified, then we are negating or modifying its existence; therefore, the existential *ser-hi* is required (9b), with the obligatory locative clitic *hi*. The use of *estar* in this case would render the sentence ungrammatical as *estar* is a locative, not an existential predicate, and it does not license the obligatory clitic *hi*:

- (9) a. Hi havia un cotxe al carrer però ara {és/està} a la plaça.
 HI HAV-3SG/PL-PAST a car in.the street but now SER/EST-3SG in the square.
 ‘There was a car in the street but it is now at the square.’

- b. Hi havia un cotxe al carrer però ara no {hi és/ (hi) *està}
 HI HAV-3SG/PL-PAST a car in.the street but now neg. HI SER/EST-3SG
 'There was a car in the street but it isn't there now.'

Haver-hi/Ser-hi

Given the nature of our task, we also found negative existentials in first mention, negating a pre-supposed element not introduced yet in the discourse, but known to the receiver thanks to the pictures. In this case, Catalan can use the existential *haver-hi* with definites and indefinites (10a), whereas Spanish requires *estar* with definite themes (11c).

- (10) a. No hi ha {un/l} esquiol.
 neg. HI HAV-3SG a/the squirrel
 'There isn't a /the squirrel'
- b. {*Un/ L}'esquiol no hi és.
 A/the squirrel neg. HI SER-3SG
 'The squirrel isn't there'

All these negative existentials and second mention locatives would be expressed with *estar* in Spanish (11a-b):

- (11) a. Había un coche en la calle pero ahora no {está/*hay).
 HAV-3SG/PL-PAST a car in the street but now neg. EST-3SG/HAV-3SG/PL
 'There was a car in the street but it isn't there now.'
- b. No hay {una/*la} ardilla.
 Neg. HAV-3SG a/the squirrel
 'There isn't a /the squirrel'
- c. {*Una/La} ardilla no está.
 A/the squirrel neg. EST-3SG
 'The squirrel isn't there'

In second mention, negative existentials with mass nouns or indefinites can use *haver-hi* as long as it is accompanied with the partitive clitic *en* (12a). Since *ser-hi* does not license the partitive clitic (12b), which in these cases is obligatory, it cannot be used in this context.

- (12) Hi havia farina a la cuina, però ara ja ...
 HI HAV-3SG/PL-PAST flour in the kitchen but now already
- a. no *(n)' hi ha.
 neg. EN HI HAV-3SG/PL
- b. *no n' hi és.
 neg. EN HI SER-3SG
 'There was flour in the kitchen but now ... there isn't'

Partitive clitics

The Catalan partitive clitic cannot refer back to the whole phrase. Instead, the partitive can only refer back to the head, or the head and its complements, but never to the specifier of the determiner phrase. Thus, if a theme is quantified, the partitive clitic cannot substitute the quantifier (13a, 13b). If the entirety of the referred entity is negated, though, the partitive is not necessary, and hence, the existential *ser-hi* can appear (13c), but not the locative *estar* (13d).

- (13) Hi havia {una botiga / cinc botigues} al carrer però ara . . .
 HI HAV-3SG/PL-PAST a shop / five shops in.the street but now
 ‘There was/were a store/ five stores in the street but now . . .’
- a. no *(n’) hi ha *(cap).
 neg. EN HI HAV-3SG/PL none
 ‘There isn’t any.’
- b. *(n’) hi ha *(quatre).
 EN HI HAV-3SG/PL four
 ‘There are four.’
- c. no hi és / no hi són.
 neg. HI SER-3SG neg. HI SER-3PL
 ‘There isn’t/There aren’t.’
- d. *no està / *no estan.
 neg. EST-3SG neg. EST-3PL
 ‘There isn’t/There aren’t.’

Spanish does not have a partitive clitic. For second mention existentials, Spanish uses *haber* as long as the quantifier appears (14a, b). If the quantifier is not present, then *estar* is required (14c).

- (14) Había una tienda/cinco tiendas en la calle pero ahora . . .
 HAV-3SG/PL-PAST a shop/five shops in the street but now
 ‘There was/were a store/ five stores in the street but now . . .’
- a. no hay *(ninguna).
 neg. HAV-3SG/PL none
 ‘There aren’t any/There isn’t one.’
- b. hay *(cuatro).
 HAV-3SG/PL four
 ‘There are four.’
- c. no está / no están.
 neg. EST-3-SG neg. EST-3-PL
 ‘There isn’t/There aren’t.’

To summarize, locative and existential constructions in Catalan and Spanish share general properties such as the predicate selection and its word order consequences in some contexts, as well as the fine line between locative predicates and existential sentences generally regulated by information structure, and some semantic restrictions and discourse mechanisms. However, they also present important differences

in the expression of the locative and existential paradigms in these very same domains. All of these appear are summarized in Table 1. In particular, the divergence is significant in the selection of existential/copular verbs (*estar* vs. *ser* vs. *haver*), since Catalan alternates between *haver-hi* and *ser-hi* in its existential (positive, negative, partial) constructions, whereas Spanish alternates between *haber* and *estar*. Furthermore, this predicate distinction has morphosyntactic consequences as Catalan requires the locative clitic *hi* in all its existential uses, and also the partitive clitic *en* in partial and negative existentials with *haver-hi*. The Catalan copulas, as the Spanish copulas and the Spanish *haber*, do not license the partitive clitic. In fact, Spanish does not have any type of overt non-personal clitic such as *HI* or *EN*. Thus, Catalan–Spanish bilinguals need to keep two different predicate systems regulated by the semantics of the theme (lexical-semantics), as well as information structure, besides the two clitic systems (functional morphology) in the mind. This, we predict, will be a locus of difficulties prone to CLI.

Previous studies

Despite the general agreement among Catalan grammarians regarding the expansion of *estar* in locative contexts, there are very few studies that have investigated the phenomenon in depth with quantitative data. Sanz and González (1995) investigated *ser* and *estar* in Tortosí Catalan, a southern variety of Catalan. Overall, the authors found extension of *estar* over *ser* with adjectives and locatives; the overextension was found to be larger in younger speakers. In locative contexts, they did not find a single instance of *ser* with animate or inanimate subjects, displaying a complete replacement of *ser* with *estar* in this context. They concluded that this language change is internally motivated, since it is also attested in Spanish and French. The language contact situation, as Silva-Corvalán (1986) argued, if anything, may have accelerated the process.

Another study that investigated the *ser/estar* uses in locative constructions is Labèrnia (2015). 248 bilingual university students completed a production task and employed *estar* 59.5% of the time as a locative copula, compared to 40.5% of use of *ser*. However, none of these studies considered the speakers' language dominance, or existential sentences in first or second mention, and negative existentials. The only study of the development of existential constructions in Catalan is by Gràcia et al. (2008), who studied L2 acquirers of Catalan whose L1 was Romanian, Arabic, or Chinese. On the other hand, Perpiñán (2015) investigated the acquisition of Spanish locative and existential constructions by Catalan native speakers. These very dominant Catalan speakers, very comparable to the participants that set our baseline in the current study, completed two tasks in Spanish: an oral production task (the equivalent to the one described in the current study but in Spanish), and an acceptability judgment task (AJT). Production results indicated an overall reduced use of *estar* and an increased use of clitics, as compared to monolingual Spanish native speakers. The AJT showed that native speakers of Catalan accepted ungrammatical sentences with *ser* to locate in Spanish, and ungrammatical definite nouns with *haber*, violating the definiteness effect constraint

Table 1. Predicate selection distribution in existential and locative constructions in Catalan and Spanish

Syntactic Construction	Semantics of THEME	Catalan		Spanish	
		Mention in the Discourse			
		1 st	2 nd	1 st	2 nd
<i>Existential</i> (<i>pro</i>) VERB- THEME- (LOC)	Definite nouns	Haver-hi / Ser-hi (person)	Ser-hi	Locative estar	Locative estar
	Indefinite nouns	Haver- hi	NA	Haber	NA
	Mass nouns / Bare plurals	Haver- hi	Haver- hi + en	Haber	Haber
	Quantified nouns	Haver- hi	Haver- hi + en + quantifier	Haber	Haber + quantifier
<i>Negative Existential</i>	Definite nouns	Haver-hi / Ser-hi	Ser-hi	Locative estar	Locative estar
	Indefinite nouns	Haver- hi	Ser-hi	Haber	Estar
<i>Locative</i> THEME-VERB-LOC	Definite nouns	Ser/Estar	Ser/Estar	Estar	Estar
	Indefinite nouns	NA	Ser-hi Estar + loc.	NA	Estar

Note: Differences across languages appear in bold.

and displaying significant differences in their grammars, with both functional convergence and language interference from Catalan into Spanish.

Clitics

The acquisition and production of Catalan clitics have received a fair amount of attention both in children (Gavarró, Mata, & Ribera, 2006; Gavarró et al., 2011; Soto-Corominas, 2020) and adults (Gavarró, 2018; Perpiñán, 2017; 2018). Gavarró et al. (2011) argued that the partitive clitic is fully acquired by the age of 5 in Catalan. However, if language dominance is taken into account as in Soto-Corominas (2018; 2020) or Perpiñán (2017), trajectories differ. These studies found that Catalan-dominant bilinguals produced locative clitic *HI* and partitive clitic *EN* significantly more than the Balanced and Spanish-dominant bilinguals in an oral production task and significantly more clitic omission in the non-Catalan-dominant bilingual groups. However, only Gràcia (2015) investigated the presence of non-personal clitics with existential verbs, though not in a systematic way. Therefore, a secondary objective of this study is to investigate the appearance of lexicalized and not lexicalized non-personal clitics with existential and locative constructions.

The present study

Our study investigates the expression of existentiality in Catalan by adult Catalan–Spanish early bilinguals with different degrees of language dominance. Considering the dearth of studies addressing the effects of bilingualism on these constructions, and the structural similarities and differences between the Catalan and Spanish universal locative paradigms, the general research questions that guide this study are the following:

- RQ1: Does language dominance predict the separation/interaction of the two systems?
- RQ2: If so, what is the scope of this interaction and how does it occur?
- RQ3: If we find differences among the grammars of bilinguals with different degrees of language dominance, are they quantitative or qualitative?

With respect to the first research question, and in line with what recent studies on Catalan–Spanish bilingualism have investigated (Perpiñán, 2017, 2018; Puig-Mayenco et al., 2018; Soto-Corominas, 2018; 2020), we predicted significant differences in the expression of existentiality in Catalan depending on the degree of language dominance of the bilinguals, so that the more Catalan-dominant bilinguals would present a more etymological and/or standard use of the language, whereas the more Spanish-dominant bilinguals would present the highest percentage of interference. If, on the other hand, coactivation of languages and early onset of bilingualism are the driving forces of CLI, then we expect Balanced Bilinguals to present the highest percentages of CLI.

Regarding RQ2, and focusing on the linguistic structure under investigation, we predict that, when input from one language reinforces one of the possible analyses available in the other language, as proposed by Silva-Corvalán (1993) and Müller

and Hulk (2001), then we could find interferences between the languages in the overlapping structures. For instance, some variants of Catalan have two possibilities to localize, i.e.: the patrimonial *ser* and the innovational *estar*, but Spanish only reinforces one of these analyses (*estar* for location). This could arguably favor the use and spread of *estar* in Catalan locatives, so that we would find CLI at the lexico-semantic interface of locative predicates. However, this overlap does not exist in existential constructions, the focus of our investigation. Therefore, we hypothesize that existential structures will be spared from CLI effects. Another important locus of crosslinguistic differences between Catalan and Spanish is the pronominal system, in the overt expression of non-personal clitics *en*, *hi*. We also predict CLI in these morphosyntactic contexts, but only in the non-lexicalized clitics constructions, that is, when the explicit location is omitted and then the locative clitic *hi* is needed, or with the partitive clitic *en*.

Finally, we hypothesize that the differences among the three depicted groups will be quantitative and not qualitative (RQ3), as they all belong to the same continuum of bilingualism (Perpiñán, 2017).

Method

Participants

Fifty-eight Catalan–Spanish bilingual speakers took part in this study. The mean age of all participants was 29.12 ($SD = 4.71$). According to participants' responses to the questionnaire, they were divided into three dominance groups: Catalan-dominant bilinguals (CatDom; $N = 21$), Balanced Bilinguals (BBil; $N = 15$), and Spanish-dominant bilinguals (SpaDom; $N = 22$). All Catalan-dominant participants had been born and raised in Central Catalunya (Manresa and surrounding towns), whereas the BBil and the SpaDom bilinguals are from the Barcelona Metropolitan area. They were recruited and tested in their place of origin and residence. While the two dialects (Central Catalan and its subvariant Barceloní) have small differences, especially in their phonology, there is no reason to believe that the two dialects have any variability with respect to existential constructions, at least not in the properties we are investigating¹.

Materials

Language Dominance Questionnaire

Participants responded to the LaDoQ online, a questionnaire targeting language dominance, made up of 47 questions. We decided to design a questionnaire for this study to determine language dominance in lieu of using existing tools since existing questionnaires have not been designed to be administered to bilinguals born and raised in a bilingual community, thus yielding limited information, particularly regarding the quality of input received. In addition to the common demographic questions, the LaDoQ asked participants to describe their current and past language use, their language preference, language identification, AOAs of each language, nativeness of input received, and self-rated proficiency in both languages with regard to different language domains. Participants were divided into three language dominance groups: Catalan-dominant (CatDom), Balanced Bilingual (BBil), and



Figure 1. Sample of 'spot-the-difference' task.

Spanish-dominant (SpaDom) according to their answers on 13 of the questions targeting language use in the home and outside, language identification, and origin of close relatives (nativeness of input). The specific questions used to determine language dominance appear highlighted in the LaDoQ in Appendix A. Most of these questions had closed answers (multiple choice: Catalan, Spanish, Both indistinctively), and they were transformed into a numerical value that created a scale of bilingualism, with Catalan on one end, and Spanish on the other. Responses that favored Catalan were coded with 1 (for instance, if the family is originally from Catalunya, if they live outside the metropolitan area of Barcelona, if they identified themselves with Catalan, etc.), and responses that favored Spanish were coded as 0; responses that favored both languages were coded as 0.5. This calculation resulted in a maximum score of 13, which served to classify the participants into the above-mentioned groups.

Oral production task

Participants completed a difference finding/telling task, as in Perpiñán (2014, 2015). In this activity, participants were shown five pairs of pictures (see Figure 1) that only differed in ten minor features. They were asked to find and explain the differences to the researcher, who was not looking at the pictures.

All OPT audio recordings were transcribed (and subsequently revised) by two trained native speakers of Catalan who coded for mention (first or second), verb used (*ser, estar, haver, tenir, portar, other, omission*), semantics of the theme ([± quantified], [± definite]), the use of the clitic (HI, EN, [± doubled], [± omitted]), the appearance of the explicit location, word order, agreement between subject and verb, and code-switching. Disagreements between transcribers were settled by discussion.

Results

Participant characteristics (LaDoQ)

On the one hand, the sum of the total points of the variables computed (a total of 13) classified the bilinguals in groups. If the score ranged from 0 to 4.5, the participant was classified as “Spanish-dominant” (SpaDom; $M = 1.55$, $SD = 1.18$); if the total score ranged from 5 to 9.5, the participants were classified as “Bilingual” (BBil;

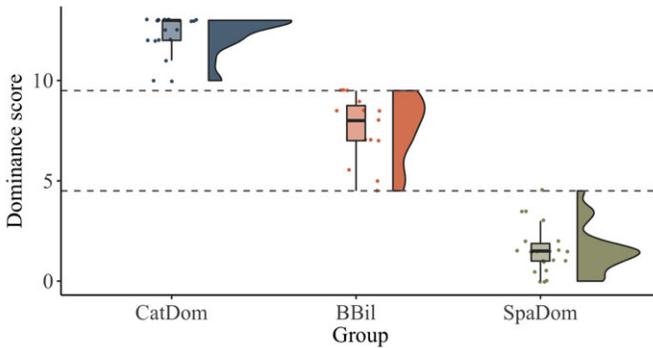


Figure 2. Participants' dominance scores.

Note: Each point indicates one participant. Split violin indicates density, so that peaks demonstrate where participants are concentrated.

$M = 7.93$, $SD = 2.03$), and if the score ranged from 10 to 13, the participants were classified as “Catalan-dominant” (CatDom; $M = 12.38$, $SD = 0.96$). Figure 2 shows how these participants constitute quite homogeneous groups and cluster together at different ranges of the language dominance continuum. Horizontal lines indicate the dominance score cut-offs employed to divide participants into dominance groups.

We then compared these dominance groups in terms of other relevant variables from the questionnaire that had not been used to compute the overall classification. Table 2 presents participant characteristics that illustrate between-group contrasts, together with the results of Kruskal–Wallis rank sum tests². Whenever a test was significant (i.e., $p < .05$), a post-hoc test with a Bonferroni adjustment was performed and results are reported in the text.

AOA was determined by asking participants when they started being exposed to Catalan and to Spanish in an uninterrupted manner. Participants had to choose one of the following ranges: 0-3, 3-6, 6-10, 10-14, or after the age of 15. These bins reflected main milestones in language development and education. We took the lower bound of each range (that is, 0, 3, 6, 10, and 15) to calculate the descriptive statistics shown in Table 2. The CatDom and BBil groups had a significantly younger Catalan AOA than the SpaDom group ($p < .001$). No differences were found between the CatDom and BBil groups. A significant difference was also found in terms of Spanish AOA. The CatDom group had a significantly older Spanish AOA than both the BBil and SpaDom groups ($p < .001$), but the BBil and SpaDom groups had a similar AOA.

The variable *Home language* was calculated by averaging language use with the mother, father, siblings, cousins, and uncles. Significant differences between the three groups were found (CatDom-BBil: $p = .001$; CatDom-SpaDom: $p < .001$; BBil-SpaDom: $p = .003$), indicating that CatDom bilinguals use more Catalan in the home than BBils and these, in turn, use more Catalan than SpaDom bilinguals. There were also differences between the three groups in terms of language use with friends (all $p < .001$). Importantly for the quality of input, 90.47% of CatDom participants had mothers born in Catalunya, compared to 80% of BBils and 14.64% of the SpaDom participants. The remaining mothers were all born in other parts of Spain.

Table 2. Participant characteristics, results from Linguistic Dominance Questionnaire (LaDoQ)

	CatDom (N = 21)	BBil (N = 15)	SpaDom (N = 22)	Kruskal–Wallis test (in all cases, $\chi^2(2)$)
	M (SD)	M (SD)	M (SD)	
Catalan AOA	0 (0)	0 (0)	3.55 (5.04)	23.415, $p < .001^{***}$
Spanish AOA	4.76 (3.30)	0.64 (1.28)	0.27 (1.28)	34.742, $p < .001^{***}$
Home language	0.96 (0.08)	0.68 (0.30)	0.05 (0.09)	45.158, $p < .001^{***}$
Language use with friends	0.98 (0.10)	0.43 (0.37)	0.18 (0.25)	38.956, $p < .001^{***}$
Years of education	14.86 (2.41)	14.40 (2.75)	14.73 (2.86)	0.36739, $p = .832$
Language identity	1 (0)	0.90 (0.21)	0.18 (0.25)	48.369, $p < .001^{***}$
Bilingual profile	0.98 (0.11)	0.60 (0.28)	0.11 (0.21)	44.538, $p < .001^{***}$
Catalan proficiency	4.30 (0.62)	4.22 (0.89)	3.96 (0.42)	6.2667, $p = .043^*$

Participants were asked to indicate their highest degree of completed education. No differences were found for this variable.

Participants were asked for their *Language identity*; whether they identified themselves more with Catalan (coded as 1), with both languages (0.5) or with Spanish (0). Differences were found between the CatDom/BBil groups and the SpaDom group ($p < .001$), but not between the CatDom and BBil groups ($p = .16$). Participants were also asked for their self-assessed *Bilingual profile*. That is, whether they considered themselves more Catalan-dominant (1), balanced bilingual (0.5), or more Spanish-dominant (0). Significant differences were found between all three groups ($p < .001$), showing that our independent classification into dominance groups largely matched participants' linguistic self-perceived profile.

Finally, participants were asked to rate their Catalan abilities in terms of spoken comprehension, reading comprehension, spoken production/pronunciation, written production, fluency, and perception of dialectal differences. Participants used a categorical scale which was subsequently turned into numerical: perfect (5), very good (4), good (3), sufficient (2), and weak (1). Even though a Kruskal–Wallis test found a significant between-groups difference, a post-hoc test did not find any significant contrasts between the three groups.

Oral production task (OPT)

The OPT proved to be a successful task to elicit locative and existential constructions, yielding 2605 predicates introducing an object in the discourse and/or locating it spatially or temporally. We will present these data by mention, whether the theme was introduced for the first time (first mention), or whether it was referred back to in the discourse (second mention).

Table 3. Counts and percentages of verb selection to express the locative paradigm in first mention

Group Verbs	<i>Haver-hi</i>	<i>Estar</i>	<i>Ser-(hi)</i>	<i>Tenir</i>	<i>Portar</i>	<i>Other</i>	<i>Null</i>	<i>Total</i>
CatDom	266	2	0	128	77	17	0	490
%	54.3	0.4	0	26.1	15.7	3.5	0	100
BBil	178	26	3	97	64	84	1	453
%	39.3	5.7	0.7	21.4	14.1	18.5	0.2	100
SpaDom	255	46	11	126	59	83	18	598
%	42.6	7.7	1.8	21.1	9.9	13.9	3.0	100

First mention

Assuming the universal locative paradigm (Freeze, 1992; Rigau, 1994), in which locative, existential, and possessive constructions are considered all ‘locative’, the OPT yielded 1541 locative or existential first mention constructions. Here, we initially counted sentences with verbs such as *portar* ‘to carry’ or *tenir* ‘to have’, which are in many respects in complementary distribution, in alternations such as *Hi ha aigua a la galleda* (‘There is water in the bucket’) vs. *La galleda porta aigua* (‘The bucket carries water’). However, in order to simplify the analysis, we will focus on the selection of existential *haver* and the copulas *ser/estar*, where the main differences between the two languages arise. The overall counts/percentages of verb selection in first mention constructions appear in Table 3.

For all bilingual groups, the verb used most frequently for first mentions was *haver-hi*, showing that this task overwhelmingly elicited existential constructions over locative predicates. In order to determine whether there were group differences in the predicate selection for first mentions, we fit a generalized linear mixed effects model (GLMM) with the Poisson distribution using the package *glmmTMB* (Brooks et al., 2017) in R (version 3.6.1). In this model, the number of instances produced for each verb, by participant, were set as the dependent variable; Group (CatDom, BBil, SpaDom), Verb (*estar*, *haver*), and their interaction were set as fixed factors and a random intercept was included for participant, with a correlated slope for verb. Due to the lack of use of *ser* across the three groups, we decided to exclude the counts for this copula in the analysis to avoid problems derived from the zero-inflated count data. Tables B1–B3 in Appendix B display the full output of this model and of the pairwise contrasts with a Bonferroni adjustment run using the *emmeans* package (version 1.4.7; Lenth, 2021). We present here the main conclusions of these analyses. All groups were significantly more likely to use *haver* than *estar*, but the CatDom group used *estar* significantly less than the BBil and SpaDom groups (both $p < .001$), with no differences between the other two groups. It should also be considered that *ser* was used to a limited extent by the BBil and SpaDom groups, while it was not used by the CatDom group at all.

In short, *haver-hi* and *estar* presented quantitative differences, and *ser* qualitative ones between the Catalan-dominant speakers on the one hand, and the two non-Catalan-dominant groups on the other hand. The utterances in (15) exemplify this phenomenon: whereas a Catalan-dominant speaker most often used *haver-hi* to introduce and localize a theme, *the little pig* in (15a), BBil and SpaDom speakers tended to do this with *estar*, as in (15b, c).

- (15) a. El porquet que **hi ha** a la dreta... CatDom#55
 the piggy that HI HAV-3SG/PL at the right
 ‘The piggy that is on the right...’
- b. En el dibuix de baix, el conillet que **està** més a la dreta... BBil#3
 in the picture of below the bunny that EST-3SG more to the right
 ‘In the picture below, the bunny that is more to the right...’
- c. El conill que **està** a la part més central... SpaDom#84
 the rabbit that EST-3SG at the part more central
 ‘The rabbit that is towards the centre...’

This was particularly the case when the theme introduced in the discourse was modified by a relative clause, a construction in which the antecedent can escape definiteness effects in both Catalan and Spanish (as explained in 2c, 2d). Forty-nine out of these 74 *estar* cases (66.2%) in first mention occurred inside a relative clause whose antecedent was the theme, as in (15b,c). The only 2 instances of *estar* produced by CatDom participants appeared inside a relative clause, whereas SpaDom and BBil participants also employed *estar* to introduce the novel theme without the modifying relative clause. Not involving a relative clause, non-Catalan-dominant bilinguals are using *estar* as a presentational verb in first mention (11/74, 14.9%), with the word order locative – copula – theme, which requires an existential predicate as in (16a); these cases were qualified as ungrammatical. The remaining uses of *estar* in first mention are arguably accepted extensions of the locative paradigm (14/74, 19%), with indefinite themes (only 2 instances), as in (16b) or with definite themes, as in (16c). All these first mention uses of *estar* constitute qualitative differences between CatDom bilinguals and non-Catalan-dominant bilinguals and are calculated by group in Table 5.

- (16) a. *A la fotografia d’abans **està** la malvada... SpaDom#73
 at the picture of before EST-3SG the witch
 ‘In the picture before there is the witch’
- b. ?Un altre ocellet a l’ arbre **està** allà dalt, a baix no. SpaDom#80
 An other bird in the tree EST-3SG there up, down not.
 ‘Another bird in the tree is up there, not down there’
- c. La noia **està** al costat d’ una columna de pedra... BBil#82
 The girl EST-3SG at.the side of a column of stone
 ‘The girl is next to a stone column’

Table 4. Counts and percentages of verb selection to express locative/existential meaning in second mention

Groups Verb	<i>Haver-hi</i>	<i>Estar</i>	<i>Ser-(hi)</i>	<i>Tenir</i>	<i>Portar</i>	<i>Other</i>	<i>Null</i>	<i>Total</i>
CatDom	146	17	66	96	44	9	20	398
%	36.7	4.3	16.6	24.1	11.0	2.3	5.0	100
BBil	75	15	69	44	18	27	51	299
%	25.1	5.0	23.1	14.7	6.0	9.0	17.1	100
SpaDom	110	44	57	70	15	23	48	367
%	30.0	12	15.5	19.1	4.1	6.3	13.1	100

Table 5. Total use of *estar* (tokens and percentage) in 1st and 2nd mention constructions

	1 st mention				2 nd mention		
	Tokens of <i>estar</i>	Locative inside Relative	Locative	Existential (ungram.)	Tokens of <i>estar</i>	Locative	Existential (ungram.)
CatDom (<i>n</i> = 21)	2	2 (100%)	0	0	17	17 (100%)	0
BBil (<i>n</i> = 15)	26	17 (65.4%)	5 (19.2%)	4 (15.4%)	15	12 (80%)	3 (20%)
SpaDom (<i>n</i> = 22)	48	30 (62.5%)	9 (18.75)	7 (14.6%)	44	30 (68.2%)	14 (31.8%)
Total	74	49 (66.2%)	14 (18.9%)	11 (14.9%)	76	59 (77.6%)	17 (22.4%)

The restricted uses of *ser* in first mention correspond to locative uses in topicalized constructions, and typically negated, as in (17). That is, *ser* was not really used as a true locative predicate in this OPT by any speakers, a finding to which we return in the Discussion section.

- (17) El cinturó del príncep a dalt no **hi és** i a baix, sí. SpaDom#81.
 the belt of the prince at top neg. HI SER-3SG and at down yes
 'The prince's belt, in the picture above it isn't there, and in the one below, it is.'

Second mention

The OPT yielded 1064 locative or existential second mention constructions. Verb choice is described in Table 4. This table includes the grammatical uses of *ser-hi* as second mention, usually negated as in (18a), and the arguably accepted locative extensions of *estar* (18b), as well as its existential uses (18c), which are not possible in the grammar of Catalan natives and described in detail in Table 5. The column

Null includes cases where the verb was omitted (19), which is a possibility in both languages.

- (18) a. Al dibuix de dalt, [...]hi ha un ocell que a dalt **no hi és**. BBil#3
 in.the drawing of top. HI HAV-3SG a bird that to top neg HI SER-3SG
 'In the picture above, there's a bird that isn't there in the picture above.'
- b. La rata que hi havia a la dreta del príncep.
 the rat that HI HAV-3SG-PAST at the right of.the prince
 ara **està** al costat de la princesa. CatDom#119
 now EST-3SG at.the side of the princess
 'The rat that was to the right of the prince is now next to the princess.'
- b. *Abans hi havia un gos i ara no **està**. SpaDom#58
 before HI HAV-3PL-PAST a dog and now neg. EST-3SG
 'There was a dog before and now it isn't there.'
- (19) A la galleda hi ha aigua i a la segona, no. BBil#1
 at the bucket HI HAV-3SG water and at the second, neg.
 'There is water in the bucket and in the second (picture) there isn't.'

For second mention verbs, we fit a GLMM as the one fit for first mention verbs. Tables B4–B6 in Appendix B include the full output of the GLMM model, together with the selected pairwise contrasts. The main findings of these analyses were as follows: the three groups displayed a preference for *haver* over *estar*, especially the CatDom group (CatDom and BBil: $p < .001$; SpaDom: $p = .001$); CatDom bilinguals and BBils showed a preference for *ser* over *estar* (CatDom: $p = .001$; BBil: $p = .006$), but not SpaDom bilinguals. CatDom bilinguals showed a preference for *haver* over *ser* ($p = .015$). SpaDom bilinguals showed a trend in the same direction ($p = .091$), but BBils did not.

Non-standard verbal uses

Some of the verbs above appeared in non-standard constructions, which we investigate in this section. Arguably, we considered 'standard' innovations all locative uses of *estar* (16c)³ but not its uses as an existential verb (16a), which are only produced by the non-dominant Catalan speakers and qualified as impossible by any Catalan native speaker. We acknowledge that cases such as that of (16b) can be considered both: locative because it is situating something in a particular space and existential because it is introducing it in the discourse for the first time. We only had two of such cases. The source of the ungrammaticality in first mention, explained in more detail above, is basically the replacement of the existential *haver-hi* with the locative *estar*. In second mention, the source of the ungrammaticality consists of the replacement of *ser-hi* with *estar*, which appears without an overt location and is thus used as an existential, usually a negative existential. An example of this ungrammatical structure is presented in (20). Counts and percentages of grammatical and ungrammatical uses of *estar* by group are summarized in Table 5.

- (20) *El bressol que hi ha en primer pla, el símbol BBil#72
 the crib that HI HAV-3PL in first plain the symbol
 en la d' abans, i en la d' ara no està.
 in the of before and in the of now neg. EST-3SG
 'The crib that appears to the fore, the symbol in the before picture and in the
 now picture it isn't there.'

The second verb that presents some deviant uses is *ser(-hi)*. When this predicate is used in non-standard constructions, it is used by BBil and SpaDom participants in first and second mention constructions, not by CatDom participants. The source of ungrammaticality with *ser-hi*, although not very frequent, is twofold. On the one hand, it is used as a first mention existential predicate ($n = 3$), where *haver-hi* should be used (21a). On the other hand, *ser-hi* is used without clitic HI, but we only found one instance of this latter case (21b). We also found several cases ($n = 8$) of ungrammatical licensing of the partitive clitic *en* with *ser-hi* (21c), in both BBil and SpaDom participants.

- (21) a. *La fauna hi és. SpaDom#67
 the fauna HI SER-3SG
 'The fauna is there.'
- b. *Hi ha un marc petit que és abans i ara, no. BBil#62
 HI SER-3SG a frame small that SER-3SG and now neg.
 'There is a small frame that is there before but not now.'
- c. *L' ocellet a dalt també n' hi és i a baix, a sota, no. SpaDom#80
 The birdie to top also EN-HI SER-3SG and to down, to below, neg.
 'The birdie (on the picture) above is also there and (in the picture) below, it isn't.'

Clitic uses

As described above, clitics HI and EN are required in certain constructions and this constitutes an important difference between the grammar of Catalan and that of Spanish. All obligatory contexts for HI or EN were coded with respect to whether the given clitic was produced in a target manner (22a), or whether it was ungrammatically omitted (22b), doubled (22c), or replaced (22d).

- (22) a. Hi ha un ocell a la segona imatge que abans no hi és. SpaDom#88
 HI HAV-3SG a bird at the second image that before neg. HI SER-3SG
 'There is a bird in the second picture that isn't there before.'
- b. *Els ocells que estan al costat del mico, BBil#70
 the birds that EST-3SG at the side of the monkey
 que o hi han més a baix.
 that HIHAV-3PL more at below
 'The birds that are next to the monkey, which there are more of in the picture below.'

- c. *En la primera imatge no **n_i**' hi ha aigua_i SpaDom#89
 in the first image neg. EN_i HI HAV-3SG/PL water_i
 'In the first picture there isn't any of water.'
- d. *?A dalt un dels monos té cua_i BBil#03
 at top one of.the monkeys has tail
 i en el dibuix de baix no **la_i** té.
 and in the picture of below neg. acc. have-3SG
 'In the picture above one of the monkeys has a tail and in the picture below,
 it doesn't have it.'

We first turn our attention to partitive clitic EN. There was a total of 448 contexts where clitic EN was obligatory (CatDom = 154, BBil = 110, SpaDom = 184). CatDom participants produced the clitic in a target-like manner in their productions 98.1% of the time ($n = 151$), compared to BBil participants' 60.9% ($n = 67$) and SpaDom participants' 37% ($n = 68$). BBil participants omitted target EN in 19.1% obligatory contexts ($n = 21$) and SpaDom participants, 34.8% ($n = 64$). The CatDom group only omitted 1.3% of obligatory EN ($n = 2$). Clitic EN was ungrammatically doubled in 19.1% of the cases by BBil participants ($n = 21$) and in 27.2% of the cases by SpaDom participants ($n = 50$), contrasting with the 0.6% rate by CatDom participants ($n = 1$). Replacements were highly infrequent in these two groups: BBil participants replaced *en* 0.9% of the cases ($n = 1$), and SpaDom participants, 1.1% ($n = 2$). CatDom participants did not replace this clitic. These results can be visualized on Figure 3.

In order to assert whether there were group differences in the expression of partitive EN, we fit a GLMM with the Binominal distribution using the package *lme4* (Bates et al., 2015) in R (version 3.6.1). The outcome of this model was whether clitic *en* was produced (1) or not produced in a target-like manner (0) in each obligatory context for the clitic. That is, omissions, clitic doublings, and replacements were all coded as 0. The only predictor was Group (CatDom, BBil, SpaDom), and a random intercept was included for participant. The complete output of the GLMM model, together with the pairwise contrasts with a Bonferroni correction, appears in Tables B7–B9 in Appendix B. Results showed that CatDom participants produced more target instances of the clitic than BBil and SpaDom participants (both contrasts, $p < .001$). On the other hand, the difference between the BBil and SpaDom groups was not found to be significant ($p = .18$).

Results for contexts where clitic HI was expected are visualized in Figure 4. Importantly, these results do not include cases of HI in the verb *haver-hi* since the clitic is grammaticalized as part of the verb and we did not find a single omission with *haver*. Nevertheless, these results include cases of HI in *ser-hi*, a decision to which we return below. A total of 290 contexts where clitic HI was required were produced (CatDom = 94, BBil = 112, SpaDom = 84). CatDom participants produced the highest percentage of target HI in obligatory contexts, 88.3% of the time ($n = 83$), followed by SpaDom participants with 81% of target productions ($n = 68$), and BBil participants with 75.9% ($n = 85$). Omissions accounted for 11.7% of the utterances by the CatDom group ($n = 11$), 19% by the SpaDom group ($n = 16$), and 22.3% by the BBil group ($n = 25$). Replacements and cases of doubling were

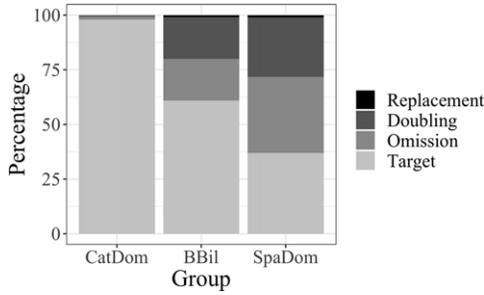


Figure 3. Productions with respect to target clitic EN, divided by group.

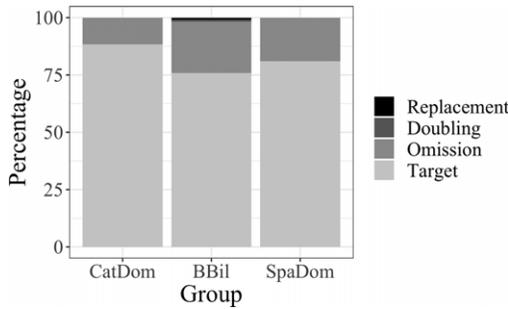


Figure 4. Productions with respect to target clitic HI, divided by group.

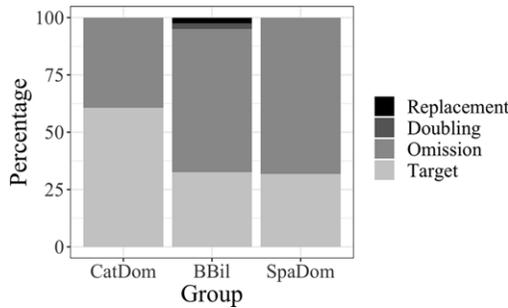


Figure 5. Productions with respect to target clitic HI, divided by group, excluding instances in verb *ser-hi*.

virtually absent. A logistic GLMM comparing group performance, such as the one used for partitive EN above, found that Group was not a significant predictor. The complete output of the GLMM model appears in Appendix B, Tables B10 and B11.

Considering the possibility that HI is grammaticalized in verb *ser-hi* as it is in *haber-hi*, results for HI were re-calculated eliminating all instances of *ser-hi*. This left a total of 90 contexts where HI was obligatory (CatDom = 28, BBil = 40, SpaDom = 22). Results can be visualized in Figure 5.

A binominal GLMM (Tables B12 and B13; Appendix B) found a significant difference between the CatDom and BBil groups ($p = .036$), indicating that the former group produced significantly more instances of target HI in obligatory contexts. It also showed a trend toward significance between the CatDom and SpaDom groups ($p = .052$) in the same direction. However, when pairwise contrasts were run with a Bonferroni correction, they did not show any significant differences between the three groups (see Table B14, Appendix B). These non-significant contrasts could likely be due to this specific analysis being underpowered due to the few contexts where *hi* was obligatory with verbs other than *ser-hi*.

Discussion and conclusions

This study set out to examine the effects of language dominance on the grammar of early Catalan–Spanish bilinguals in adulthood with respect to Catalan existential constructions. Specifically, we asked three questions regarding (1) the capacity of language dominance to predict the interaction of the two linguistic systems, (2) the extent of this interaction, and (3) the quantitative/qualitative nature of differences among bilinguals.

First of all, in order to operationalize language dominance and make comparisons possible, we employed a three-way division of bilingual speakers. As can be seen in Figure 2, the BBil group is the most heterogenous one, and also the most difficult to find in the real world (cfr. Idescat, 2018). Nevertheless, given the results shown in Table 2, we argue that our LaDoQ is an effective tool to categorically classify Catalan–Spanish bilinguals and that our three groups reliably represent not only three different sociolinguistic realities but also three different linguistic behaviors. Next, we compared the linguistic results among the three bilingual dominance groups and found significant differences in the way they expressed existentiality in Catalan. Our data indicate selective indirect CLI from Spanish into Catalan modulated by language dominance. Our results suggest that these differences cannot be accounted for by AOA in Catalan and/or Catalan proficiency alone. If these two variables alone were responsible for the observed CLI, we would not have expected any differences between the CatDom and BBil participants, who did not differ in either of these two dimensions. Since differences between the CatDom and BBil group were observed in the expression of existential constructions, we conclude that accurate measures of language dominance need to consider other sources of variability, such as language exposure and quality of input, in cases of sustained widespread bilingualism.

Furthermore, BBils, allegedly the speakers with more coactivation and an earlier onset of bilingualism, are not the participants who presented the highest percentage of CLI. In particular, we found SpaDom bilinguals, and to a lesser extent also BBil, used significantly more frequently *estar* as a presentational verb, to locate objects for the first time in the discourse, and as a negative or partial existential in second mention, uses that are not accepted in the Catalan grammar and critically not present in the grammar of CatDom bilinguals, the baseline group. Thus, for these non-dominant Catalan speakers, *estar* is not only the verb for locative predicates but also for some existential constructions. While there is scarce independent evidence that

estar has been extended in locative predicates in some varieties of Catalan (Sanz & González, 1995), to our knowledge, no study to date had investigated the effect of this semantic extension also in existential constructions in Catalan–Spanish bilingualism. Granted, the extension of the Catalan copula *estar* in detriment of the use of *ser* with adjectives and locatives, which this study does not investigate, might be due to internal factors (i.e., the natural development of the copulas in Catalan), accelerated by external factors (i.e., the effect of Spanish), as in the case of Spanish in contact with English (Silva-Corvalán, 1986; Solà, 1994). We argue that these *estar* extensions in locative contexts are possible and fairly accepted in Catalan because these linguistic contexts already presented optionality in Catalan, with subtle different aspectual interpretations. However, existential contexts do not present this optionality; therefore, the incipient use of *estar* in existential constructions observed in this study, which is highly marked and considered ungrammatical, cannot be the direct result of an internal extension of *estar* given that *estar* does not alternate with *haver-hi* or *ser-hi* in existential constructions.

In other words, the two extensions of *estar* are of different nature: on the one hand, the locative extension (not directly investigated here) stems from the fact that etymological Catalan can already employ *estar* for locative predicates, initially with a marked aspectual meaning of duration or permanence (see examples in (7)), in a semantic bleaching process. That is, there is already certain overlap of uses between *ser* and *estar* in the locative structure, and thus the perfect condition for CLI to occur: speakers are choosing to use *estar* over *ser* in this context, and the fact that Spanish only provides input for one of these options favors the extension of *estar* in locative constructions. However, this optionality between the copulas *ser/estar* does not apply to Catalan existential constructions, neither in first nor in second mention existentials (see Table 1 for easy reference); therefore, CLI is not expected here. But indeed, we are finding innovative uses of *estar* also in existential constructions, a conceptually related structure to the locative structure. Therefore, we argue that *indirect* CLI can also surface in structure X if the bilinguals *perceive* that a related structure Y from language A, in this case the locative construction, has multiple solutions, and input from language B in that structure Y reinforces only one of those solutions. In such a way that CLI applies comprehensively and culminates in the same outcome across the entire paradigm: the spread of *estar* in both, locative and existential constructions⁴. Our proposal of indirect CLI is schematized in Figure 6.

Notice that the use of *estar* in existential constructions technically is not alleviating any ambiguity or structural overlap in the input, as originally argued for CLI (Müller, 1998), since optionality between *ser/estar* was not available in this context. The only ambiguity we conceive is that between locative and existential constructions. As explained in the linguistic description, locative and existentials are two complementary lexico-syntactic ways to express spatial relationships with different semantic and discursive entailments, but in sentences such as (16b), repeated here for convenience, the fine line between the two is blurred.

- (16) b. ?Un altre ocellet a l'arbre **està** allà dalt, a baix no. SpaDom #80
 an other birdie in the tree est-3sg there up, down not.
 'Another birdie in the tree is up there, not down there'

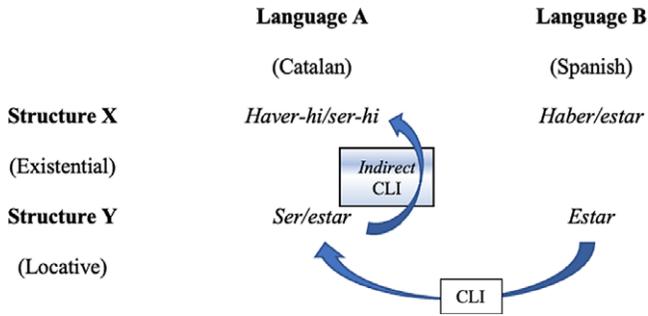


Figure 6. Theoretical Proposal: Indirect Crosslinguistic Influence.

We hypothesize that the source of the emergence of *estar* as an existential verb could be precisely the ambiguity and complex lexico-syntactic puzzle that the existential and locative constructions constitute, driven by semantic and discourse-pragmatics considerations. Indeed, the fact that Spanish is forced to alternate the locative structure with the existential one according to the semantics of the NP (6a vs. 6b) may have a non-trivial effect: Spanish speakers may perceive the two structures as equivalent and make this equivalence also in their non-dominant language, in this case, Catalan.

One could argue that the unacceptable appearance of *estar* in second mention existential contexts is a direct transfer from Spanish, as this language can employ *estar* with negative existentials and with token readings. We do not entertain this explanation here because we believe that the incorporation of new morphosyntactic elements in a language has its limits (Silva-Corvalán, 1993) and are only possible when the receptive system allows for them. It is more plausible to resort to and spread an already existing feature in the language, than to import an external one. Moreover, if this was a case of direct transfer, and not of mediated or indirect CLI, its appearance would be much abundant, similar to the spread of locative *estar*. Instead, we argue that the Catalan variety that these bilingual speakers are creating allows for this existential *estar* precisely because *estar* already exists in their system in a related structure, the locative one. That is, we predict that only speakers that productively use *estar* as a locative may extend it also to existential contexts. This statement remains to be proven with further data. Likewise, the replacement of *haver-hi* with *estar* in first mention existentials cannot be explained in any way with a direct transfer from Spanish, as Spanish does not allow *estar* in these contexts and requires *haber*. This constitutes further evidence for indirect CLI across related structures in the paradigm, and against a direct transfer explanation.

The other main linguistic result concerns the differential use of obligatory non-personal pronouns *en*, *hi* that bilinguals make; here, and particularly with the use of partitive *en*, we find striking differences between the CatDom bilinguals on the one hand, and the non-CatDom groups (BBil and SpaDom) on the other hand. With respect to the partitive clitic, Perpiñán (2017) already tested bilingual speakers with transitive sentences, in which acceptance rates of ungrammatical clitic omission were quite high, and omission in production was exclusive of SpaDom bilinguals and BBils. Similarly, omission of the obligatory partitive clitic in existential contexts

in this study, as well as its illicit combination with *ser-hi* occurs only in non-Catalan-dominant bilinguals. This supposes a qualitative difference among the groups. With respect to the clitic *hi*, we only found differences among the groups when *ser-hi* was not factored in, a result that we interpret as the complete lexicalization of this clitic with the verb *ser* in existential uses, since when measured productively (not with *ser-hi*), the non-Catalan-dominant groups presented large amount of (ungrammatical) omission. It needs to be said that also CatDom omit the clitic *hi* to a significant degree and that no qualitative differences, but only quantitative ones are found with regard to the locative clitic in productive contexts.

Consequently, our results indicate that the three-way dominance distinction we created is necessary to better reflect the linguistic reality of bilingual societies, in this particular case of Catalunya. This three-way distinction independently correlates quantitatively and qualitatively with the linguistic results, as we found selective differences among the three groups depending on the linguistic feature. Most frequently, CatDom bilinguals behaved differently than the non-Catalan-dominant groups, particularly in the use of *estar* in first mentions, and the partitive clitic *en*; BBils also patterned with the CatDom bilinguals in certain features, such as in the reduced use of *estar* in second mentions. Finally, the CatDom and SpaDom groups systematically differed in the crucial aspects investigated in this study, despite having the same high proficiency in Catalan. These systematic differences challenge the idea that these two groups share the exact same linguistic representations regarding existentiality and locatives. Likewise, these results could call into question whether these two discrete groups represent two points from the same continuum of bilingualism, as previously argued (Perpiñán, 2017), as these groups presented not only significant quantitative differences but also qualitative ones. Here, we propose a bilingualism continuum with the possibility of a divergent production; this divergent production could well be the reflection of a divergent mental representation, although other type of data (intuitional, comprehension, processing data) would be needed in order to confirm this approach.

To summarize and responding to our first research question regarding the interaction of the two systems depending on language dominance, the answer is yes. Our data showed that Catalan–Spanish bilingualism degree tightly correlates with linguistic knowledge in the minority language, to the point that non-Catalan-dominant speakers employ a series of differential structures in their Catalan, such as the frequent innovative uses of presentational *estar* in first mention structures, as well as ungrammatical existential uses of *estar* in second mention. That is, non-Catalan-dominant bilinguals have restructured their Catalan in a significant (i.e., qualitative) way, favoring the use of *estar* even in existential contexts in which standard Catalan does not present overlapping structures, unlike what could be the case for locative predicates. At this point, one can only wonder when these highly marked existential uses of *estar* would be so frequent in the language that we cannot longer consider them ungrammatical, as we venture might have happened with the locative uses of *estar*. Certainly, we cannot predict when language change would occur, or if it would occur at all, but we can be certain that bilingualism is playing a role in these restructurings. The fact the BBil group is already showing some of these distinctive features makes us think that this is a change in progress.

With respect to the scope of CLI (RQ2), we have attested CLI in the lexico-syntactic interface (predicate selection) mostly regulated by semantics and information structure, as well as in morphosyntax (clitics). Moreover, we found *indirect* CLI, when the influence is implemented by a related structure in the language in which there is no optionality, but bilinguals *perceive* certain overlap within the linguistic structures (locative and existential constructions). This is a novel finding that, to our knowledge, had not been described in the literature before. Finally, our study found qualitative differences (RQ3) among the different types of bilinguals, despite their commensurate levels of proficiency, a finding that forces us to reformulate the proposed continuum of bilingualism in Catalan–Spanish speakers into a continuum with divergent production, and potentially divergent mental representations, at least for the expression of existentiality.

Acknowledgements. We gratefully acknowledge the financial support of two University of Western Ontario SSHRC-Internal Grants. We would like to thank all the participants of the study, as well as the Ajuntament de Sant Adrià de Besòs and Mònica Tarrés for their help in the data collection. We thank the three anonymous reviewers for their constructive feedback, as well as the editor of the journal. All the remaining errors are our own.

Notes

1. One anonymous reviewer asked about possible dialectal differences between these two varieties. Ramos (1998) examines presentational sentences in six Catalan dialects and classifies them in four groups according to their behavior with respect to this construction: on the one hand, Central Catalan, -our object of study-, Northwestern Catalan, and Balearic work as described in our linguistic section; on the other hand, Valencian, Roussillonese, and Algerese have all distinct properties whose description goes beyond the scope of this study. There is a reported dialectal difference in existential constructions in Central Catalan, and it consists on the non-standard agreement of the auxiliary *haver* with the theme of the construction, such as *hi ha cinc llibres* vs. *hi han cinc llibres* ('There are five books'). Normative Catalan, as well as normative Spanish, recommends the use of the invariant singular *haver* in all cases, however, the agreement form is very common in Catalan, more than in Spanish. Oriental Catalan dialects tend to use the agreement form, whereas the occidental Catalan dialects use the invariant 3rd person singular *have* (Rigau, 1997; Solà, 1973). And even though the varieties we are dealing with (central Catalan and Barcelonès) are both oriental, so they both present substandard agreement of *haver*, we purposely decided not to investigate this issue here (number in *haver*) given the dialectal optionality this form may present in general Catalan. Thus, the production of *haver* we find in our participants might be agreeing or not agreeing with the theme.
2. We used the Kruskal-Wallis test (rather than a parametric one-way ANOVA) because all variables were collected as ordinal data, rather than as interval data.
3. Admittedly, this is a very progressive way to define a standard form since many native speakers of Catalan would already consider these locative uses of *estar* ungrammatical, even when they appeared with an overt location, as in (16c). Nevertheless, we coded them here as accepted and not non-standard given their advance in nowadays Catalan. Recall that our task does not directly elicit locative predicates, so we are not directly providing data of the extension of *estar* in nowadays Catalan (Sanz & González, 1995; Solà, 1994).
4. An anonymous reviewer questions the use of the verb *perceive* in this context, given that we do not have perception or awareness data in this study, and asks us to clarify that what we mean with *perceive* here is that the speaker conceptually relates the two structures. That is, that the bilingual may not make a categorical distinction between the two. We thank the reviewer for this observation.

References

- Bassaganyas-Bars, T. (2015). "Have" and the link between perfects and existentials in Old Catalan. *Proceedings of Sinn Und Bedeutung*, 19, 107–124.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., Bohman, T. M., & Gillam, R. B. (2012). The measure matters: Language dominance profiles across measures in Spanish–English bilingual children*. *Bilingualism: Language and Cognition*, 15(3), 616–629. <https://doi.org/10.1017/S1366728912000090>
- Birdsong, D. (2006). Age and Second Language Acquisition and Processing: A Selective Overview. *Language Learning*, 56(S1), 9–49.
- Birdsong, D. (2014). Dominance and Age in Bilingualism. *Applied Linguistics*, 35(4), 374–392. <https://doi.org/10.1093/applin/amu031>
- Blom, E., Cornips, L., & Schaeffer, J. (Eds.). (2017). *Cross-linguistic Influence in Bilingualism: In honor of Aafke Hulk*. Amsterdam: John Benjamins Publishing Company.
- Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., ... & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, 9(2), 378–400.
- Bylund, E., Hyltenstam, K., & Abrahamsson, N. (2021). Age of acquisition – not bilingualism – is the primary determinant of less than nativelike L2 ultimate attainment. *Bilingualism: Language and Cognition*, 24(1), 18–30. <https://doi.org/10.1017/S1366728920000188>
- De Houwer, A. (2018). The role of language input environments for language outcomes and language acquisition in young bilingual children. In D. Miller, F. Bayram, J. Rothman & L. Serratrice (Eds.), *Bilingual cognition and language: The state of the science across its subfields* (pp. 127–154). Philadelphia: John Benjamins Publishing Company.
- Fábregas, A. (2019). *The definiteness effect in Catalan and Spanish: same principle different parametrization*. Talk presented at Jornades de la Lingüística Catalana a Viena. Universitat Wien, November, 2019.
- Freeze, R. (1992). Existentials and Other Locatives. *Language*, 68(3), 553–595.
- Gathercole, V. C. M., & Thomas, E. M. (2009). Bilingual first-language development: Dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition*, 12(2), 213–237.
- Gavarró, A. (2018). Language acquisition and change: The acquisition of the Catalan partitive and locative clitics. In P. Guijarro-Fuentes & A. Cuza (Eds.), *Language Acquisition and Contact in the Iberian Peninsula* (pp. 11–36). De Gruyter Mouton. <https://www.degruyter.com/view/title/538913>
- Gavarró, A., Guasti, M. T., Turell, L., Prévost, P., Belletti, A., Cilibrasi, L., Delage, H., & Vernice, M. (2011). The acquisition of partitive clitics in Romance five-years-olds. *Iberia*, 3(2), 1–19.
- Gavarró, A., Mata, M., & Ribera, E. (2006). L'omissió dels clítics d'objecte i partitius en el català infantil: Dades espontànies. In C. D. Pusch (Ed.), *La gramàtica pronominal del català: Variació, evolució, funció = The grammar of catalan pronouns: Variation, evolution, function* (pp. 27–46). Shaker.
- 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.
- Gràcia, L. (2015). Catalan existential sentences produced by Chinese learners. *Italian Journal of Linguistics*, 27(2), 1–46.
- Gràcia, L., Crous, B., & Garganta, L. (2008). Typological Differences and Second Language Acquisition 2: Attributive, Locative, Existential, and Possessive Constructions. *Revista española de lingüística*, 38(2), 47–74.
- Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36(1), 3–15. [https://doi.org/10.1016/0093-934X\(89\)90048-5](https://doi.org/10.1016/0093-934X(89)90048-5)
- Grosjean, F. (2008). *Studying Bilinguals*. Oxford University Press. <http://ebookcentral.proquest.com/lib/upf-trial/detail.action?docID=415801>
- Grosjean, F. (2010). *Bilingual: Life and reality*. Harvard University Press.
- Hicks, G., & Domínguez, L. (2020). A model for L1 grammatical attrition. *Second Language Research*, 36(2), 143–165. <https://doi.org/10.1177/0267658319862011>
- Hulk, A., & Müller, S. (2000). Bilingual First Language Acquisition at the Interface between Syntax and Pragmatics. *Bilingualism: Language and Cognition*, 3(3), 227–244.

- Hyltenstam, K., & Abrahamsson, N. (2003). Age of onset and ultimate attainment in near-native speakers of Swedish as a second language. In K. Hyltenstam & K. Fraurud (Eds.), *Multilingualism in Global and Local Perspective* (pp. 319–340). Stockholm University.
- Idescat (2018). EULP, Enquesta d'usos lingüístics de la població 2018. Institut d'Estadística de Catalunya. <https://www.idescat.cat/>
- Institut d'Estudis Catalans. (2016). *Gramàtica de la llengua catalana (Primera edició)*. Institut d'Estudis Catalans.
- Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. Routledge.
- Kupisch, T. (2012). Specific and generic subjects in the Italian of German-Italian simultaneous bilinguals and L2 learners. *Bilingualism: Language and Cognition*, 15(4), 736–756. <https://doi.org/10.1017/S1366728911000691>
- Kupisch, T. (2018). Introduction: Recent developments in early bilingualism. *Bilingualism: Language and Cognition*, 21(4), 653–655. <https://doi.org/10.1017/S1366728918000548>
- Kupisch, T., & Weijer, J. van de. (2015). The role of the childhood environment for language dominance: A study of adult simultaneous bilingual speakers of German and French. In C. Silva-Corvalán & J. Treffers-Daller (Eds.), *Language Dominance in Bilinguals: Issues of Measurement and Operationalization* (pp. 174–194). Cambridge University Press. <https://doi.org/10.1017/CBO9781107375345.009>
- Labèrnia, A. (2015). *La incidència de certs factors sociolingüístics en la implantació de setze variants sintàctics en català*. Ph.D. Dissertation, Universitat Pompeu Fabra.
- Lenth, R. V. (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.6.0. <https://CRAN.R-project.org/package=emmeans>
- Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology (Hove, England)*, 25(5), 605–621. <https://doi.org/10.1080/20445911.2013.795574>
- McNally, L. (2011). Existential Sentences. In C. Maienborn, K. V. Heusinger, & P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning* (pp. 1829–1848). Walter de Gruyter.
- McNally, L. (2016). Existential Sentences Crosslinguistically: Variations in Form and Meaning. *Annual Review of Linguistics*, 2(1), 211–231. <https://doi.org/10.1146/annurev-linguistics-011415-040837>
- Meisel, J. M. (2009). Second Language Acquisition in Early Childhood. *Zeitschrift Fur Sprachwissenschaft*, 28(1), 5–34.
- Meisel, J. M. (2018). Early child second language acquisition: French gender in German children. *Bilingualism: Language and Cognition*, 21(4), 656–673. <https://doi.org/10.1017/S1366728916000237>
- Meisel, J. M., Elsig, M., & Rinke, E. (2013). *Language acquisition and change: A morphosyntactic perspective*. Edinburgh University Press.
- Milsark, G. L. (1977). Toward an Explanation of Certain Peculiarities of the Existential Construction in English. *Linguistic Analysis*, 3(1), 1–29.
- Montrul, S. (2008). *Incomplete Acquisition in Bilingualism: Re-examining the Age Factor*. John Benjamins Publishing Company.
- Montrul, S. (2014). Structural changes in Spanish in the United States: Differential object marking in Spanish heritage speakers across generations. *Lingua*, 151, 177–196. <https://doi.org/10.1016/j.lingua.2014.05.007>
- Montrul, S. (2015). Dominance and proficiency in early and late bilingualism. In C. Silva-Corvalán & J. Treffers-Daller (Eds.), *Language Dominance in Bilinguals: Issues of Measurement and Operationalization* (pp. 15–35). Cambridge University Press. <https://doi.org/10.1017/CBO9781107375345.002>
- Montrul, S. (2016a). *The Acquisition of Heritage Languages*. Cambridge University Press.
- Montrul, S. (2016b). Age of Onset of Bilingualism Effects and Availability of Input in First Language Attrition. In E. Nicoladis & S. Montanari (Eds.), *Bilingualism across the Lifespan: Factors Moderating Language Proficiency* (p. 141).
- Müller, N. (1998). Transfer in bilingual first language acquisition. *Bilingualism, Language and Cognition*, 1(3), 151–171.
- Müller, N., & Hulk, A. (2001). Crosslinguistic Influence in Bilingual Language Acquisition: Italian and French as Recipient Languages. *Bilingualism: Language and Cognition*, 4(1), 1–53.
- Nicoladis, E. (2018). Simultaneous Child Bilingualism. In D. Miller, F. Bayram, J. Rothman, & L. Serratrice (Eds.), *Bilingual cognition and language: The state of the science across its subfields* (pp. 81–101). Philadelphia : John Benjamins Publishing Company.

- Pavlenko, A., & Jarvis, S.** (2002). Bidirectional Transfer. *Applied Linguistics*, 23(2), 190–214.
- Perpiñán, S.** (2014). Locatives and existentials in L2 Spanish: The acquisition of the semantic contrasts among *ser*, *estar* and *haber*. *Second Language Research*, 30(4), 485–513. <https://doi.org/10.1177/0267658313518215>
- Perpiñán, S.** (2015). The locative paradigm in the L2 Spanish of Catalan native speakers. In T. Judy & S. Perpiñán (Eds.), *The Acquisition of Spanish in Understudied Language Pairings* (pp. 105–132). John Benjamins.
- Perpiñán, S.** (2017). Catalan-Spanish bilingualism continuum: The expression of non-personal Catalan clitics in the adult grammar of early bilinguals. *Linguistic Approaches to Bilingualism*, 7(5), 477–513. <https://doi.org/10.1075/lab.15004.per>
- Perpiñán, S.** (2018). On Convergence, Ongoing Language Change, and Crosslinguistic Influence in Direct Object Expression in Catalan–Spanish Bilingualism. *Languages*, 3(2), 14. <https://doi.org/10.3390/languages3020014>
- Perpiñán, S., Marín, R., & Villamar, I. M.** (2020). The role of aspect in the acquisition of *ser* and *estar* in locative contexts by English-speaking learners of Spanish. *Language Acquisition*, 27(1), 35–67. <https://doi.org/10.1080/10489223.2019.1610408>
- Polinsky, M.** (2018). *Heritage languages and their speakers*. Cambridge University Press.
- Puig-Mayenco, E., Cunnings, I., Bayram, F., Miller, D., Tubau, S., & Rothman, J.** (2018). Language Dominance Affects Bilingual Performance and Processing Outcomes in Adulthood. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.01199>
- Putnam, M. T., & Sánchez, L.** (2013). What's so incomplete about incomplete acquisition?: A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism*, 3(4), 478–508. <https://doi.org/10.1075/lab.3.4.04put>
- Ramos, J. R.** (1998). Presentational Sentences in Catalan Dialects. *Catalan Working Papers in Linguistics*, 6, 41–58.
- Ramos, J. R.** (2000). *Ésser, estar i haver-hi en català antic: Estudi sintàctic i contrastiu*. Institut Interuniversitari de Filologia Valenciana.
- Ramos, J. R.** (2002). El SV, II: la predicació no verbal obligatòria. In J. Solà, M.-R. Lloret, M. Joan, & M. Pérez Saldanya (Eds.), *Gramàtica del català contemporani I, II i III* (pp. 1951–2044). Editorial Empúries.
- Rigau, G.** (1994). Catalan Presentational Sentences and the Properties of Agr Nodes. In G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi, & R. Zanuttini (Eds.), *Paths towards Universal Grammar: Studies in Honor of Richard S. Kayne* (pp. 343–359). Georgetown University Press.
- Rigau, G.** (1997). Sintaxi i variació dialectal: Les construccions locatives. In M. R. Lloret, E. Boix, M. Llorente, L. Payrató, & M. P. Perea (Eds.), *Anàlisi de la variació lingüística: Actes de la 2a Jornada sobre la Variació Lingüística i del 3r Colloqui Lingüístic de la Universitat de Barcelona (Club-3)* (pp. 179–202). PPU, SA.
- Sanz, C., & González, M. J.** (1995). *Ser* and *Estar* in Tortosí Catalan: Language Contact, Language Variation, and Language Change. *Sintagma*, 7, 5–25.
- Schmid, M., & Yilmaz, G.** (2018). Predictors of Language Dominance: An Integrated Analysis of First Language Attrition and Second Language Acquisition in Late Bilinguals. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.01306>
- Schmid, M. S., & Köpke, B.** (2017). The relevance of first language attrition to theories of bilingual development. *Linguistic Approaches to Bilingualism*, 7(6), 637–667. <https://doi.org/10.1075/lab.17058.sch>
- Serratrice, L.** (2013). Cross-linguistic influence in bilingual development: Determinants and mechanisms. *Linguistic Approaches to Bilingualism*, 3(1), 3–25. <https://doi.org/10.1075/lab.3.1.01ser>
- Silva-Corvalán, C.** (1986). Bilingualism and Language Change: The Extension of *estar* in Los Angeles Spanish. *Language*, 62(3), 587–608.
- Silva-Corvalán, C.** (1993). On the permeability of grammars. In W. J. Ashby, M. Mithun, & G. Perissinotto (Eds.), *Linguistic Perspectives on Romance Languages: Selected Papers from the XXI Linguistic Symposium on Romance Languages, Santa Barbara, February 21–24, 1991* (pp. 19–42). John Benjamins. <https://benjamins.com/catalog/cilt.103.08sil>
- Silva-Corvalán, C.** (1993). On the permeability of grammars. In W. J. Ashby, M. Mithun, & G. Perissinotto (Eds.), *Linguistic Perspectives on Romance Languages: Selected Papers from the XXI Linguistic Symposium on Romance Languages, Santa Barbara, February 21–24, 1991* (pp. 19–42). John Benjamins. <https://benjamins.com/catalog/cilt.103.08sil>

- Silva-Corvalán, C.** (2002). *Language Contact and Change: Spanish in Los Angeles*. Clarendon Press ; Oxford University Press.
- Silva-Corvalán, C., & Montanari, S.** (2008). The acquisition of *ser*, *estar* (and *be*) by a Spanish–English bilingual child: The early stages*. *Bilingualism: Language and Cognition*, *11*(3), 341–360. <https://doi.org/10.1017/S136672890800357X>
- Silva-Corvalán, C., & Treffers-Daller, J.** (2015). Digging into dominance: A closer look at language dominance in bilinguals. In *Language Dominance in Bilinguals: Issues of Measurement and Operationalization* (pp. 1–14). Cambridge University Press. <https://doi.org/10.1017/CBO9781107375345.001>
- Solà, J.** (1972/1973). *Estudis de sintaxi catalana* (1a ed). Edicions 62.
- Solà, J.** (1994). *Sintaxi normativa: Estat de la qüestió* (1. ed). Editorial Empúries.
- Sorace, A.** (2005). Selective optionality in language development. In L. E. A. Cornips & K. P. Corrigan (Eds.), *Syntax and Variation: Reconciling the Biological and the Social* (pp. 55–80). John Benjamins.
- Sorace, A.** (2011). Pinning down the concept of interface in bilingualism. *Linguistic Approaches to Bilingualism*, *1*(1), 1–33. <https://doi.org/10.1075/lab.1.1.01sor>
- Sorace, A., & Serratrice, L.** (2009). Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism*, *13*(2), 195–210.
- Soto-Corominas, A.** (2018). Acquisition of Catalan and Spanish Morphosyntax in the Catalan-Spanish Bilingual Context. *Electronic Thesis and Dissertation Repository*. <https://ir.lib.uwo.ca/etd/5441>
- Soto-Corominas, A.** (2020). Acquisition of Quantified Partitivity in Catalan-Spanish Bilingualism. Influence from Child-Level and Language-Level Factors. *Linguistic Approaches to Bilingualism*. <https://doi.org/10.1075/lab.19016.sot>
- Thomas, E. M., Williams, N., Jones, L. A., Davies, S., & Binks, H.** (2014). Acquiring complex structures under minority language conditions: Bilingual acquisition of plural morphology in Welsh. *Bilingualism: Language and Cognition*, *17*(3), 478–494.
- Thomason, S. G., & Kaufman, T.** (1988). *Language contact, creolization, and genetic linguistics*. University of California Press.
- Treffers-Daller, J.** (2019). What Defines Language Dominance in Bilinguals? *Annual Review of Linguistics*, *5*(1), 375–393. <https://doi.org/10.1146/annurev-linguistics-011817-045554>
- Unsworth, S.** (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition*, *16*(01), 86–110. <https://doi.org/10.1017/S1366728912000284>
- Unsworth, S.** (2016). Quantity and Quality of Language Input in Bilingual Language Development. In E. Nicoladis (Ed.), *Bilingualism Across the Lifespan* (pp. 103–122). De Gruyter Mouton. <https://doi.org/10.1515/9783110341249-008>
- Veselinova, L.** (2013). Negative existentials: A cross-linguistic study. *Rivista Di Linguistica*, *25*(1), 107–145.
- Zeitoun, E., Huang, L. M., Yeh, M. M., & Chang, A. H.** (1999). Existential, Possessive, and Locative Constructions in Formosan Languages. *Oceanic Linguistics*, *38*(1), 1–42. JSTOR. <https://doi.org/10.2307/3623391>
- Zhou, J., Mai, Z., & Yip, V.** (2020). Bidirectional cross-linguistic influence in object realization in Cantonese–English bilingual children. *Bilingualism: Language and Cognition*, 1–15. <https://doi.org/10.1017/S1366728920000231>

Appendix 1. Language Dominance Questionnaire (LaDoQ)

The original questionnaire was delivered online and in Catalan. For the original version, contact the first author. Questions used to determine participants' language dominance are preceded by an asterisk ().*

1. Participant number: _____
2. How old are you? _____
3. *Where do you live? _____
4. *Where was your mother born? _____
5. *Where was your father born? _____
6. *Where were your maternal grandparents born? _____
7. *Where were your paternal grandparents born? _____
8. *What language do you speak with your mother?
 Catalan Spanish Both Catalan and Spanish Other: _____
9. *What language do you speak with your father?
 Catalan Spanish Both Catalan and Spanish Other: _____
10. *What language do you speak with your siblings?
 Catalan Spanish Both Catalan and Spanish Other: _____
11. What language do you speak with your maternal grandparents?
 Catalan Spanish Both Catalan and Spanish Other: _____
12. What language do you speak with your paternal grandparents?
 Catalan Spanish Both Catalan and Spanish Other: _____
13. *What language do you speak with your close friends?
 Catalan Spanish Both Catalan and Spanish Other: _____
14. What language do you speak with your cousins?
 Catalan Spanish Both Catalan and Spanish Other: _____
15. *What language do you speak with your partner?
 Catalan Spanish Both Catalan and Spanish NA/Other: _____
16. *How old were you when you started being exposed to Catalan in a continuous manner?
 0–3 years 3–6 years 6–10 years 10–14 years After 15 years
17. How old were you when you started being exposed to Spanish in a continuous manner?
 0–3 years 3–6 years 6–10 years 10–14 years After 15 years
18. If you used to speak a different language with your siblings than you do now, how old were you when you switched languages?
 0–3 years 3–6 years 6–10 years 10–14 years After 15 years No change
19. *Do you consider yourself...
 More Catalan-dominant A balanced bilingual More Spanish-dominant
 Other: _____
20. *What language do you identify yourself with most?
 Catalan Spanish Both Catalan and Spanish Other: _____
21. What is your highest level of education attained?
 Primary education Secondary education University education

22. How would you rate your Catalan skills?

	Perfect	Very good	Good	Sufficient	Weak
Spoken comprehension					
Reading comprehension					
Spoken production/pronunciation					
Written production					
Fluency					
Perception of dialectal differences					

23. How would you rate your Spanish skills?

	Perfect	Very good	Good	Sufficient	Weak
Spoken comprehension					
Reading comprehension					
Spoken production/pronunciation					
Written production					
Fluency					
Perception of dialectal differences					

24. Indicate your language use for each period of time in your life, depending on the situation:

	Only Spanish	Mainly Spanish	More Spanish than Catalan	More Catalan than Spanish	Mainly Catalan	Only Catalan
As a child, before the onset of schooling						
During primary education, at school						
During primary education, at home						
During primary education, in the streets						
During primary education, in other places						
During secondary education, at school						

(Continued)

(Continued)

	Only Spanish	Mainly Spanish	More Spanish than Catalan	More Catalan than Spanish	Mainly Catalan	Only Catalan
During secondary education, at home						
During secondary education, in the streets						
During secondary education, in other places						
As an adult, in university/at work						
As an adult, at home						
As an adult, in the streets						
As an adult, in other places						

Appendix B. Full output models

Table B1. Output for Poisson GLMM for first mention verbs

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-2.535	0.732	-3.462	.001
GroupBBil	2.913	0.767	3.795	<.001***
GroupSpaDom	3.088	0.750	4.117	<.001***
Verbhaver	5.017	0.740	6.778	<.001***
GroupBBil:Verbhaver	-3.015	0.786	-3.836	<.001***
GroupSpaDom:Verbhaver	-3.213	0.766	-4.196	<.001***

Note: the reference level for Group was *CatDom*, and for *Verb* it was *estar*.

Table B2. Random effects for Poisson GLMM for first mention verbs

Groups	Name	Variance	Std. Dev.	Corr
Participant	(Intercept)	0.378	0.615	
	Verbhaver	0.542	0.736	-0.83

Table B3. Selected pairwise contrasts with Bonferroni adjustment for Group \times Verb interaction for first mention verbs

term	contrastfield	insidelevel	contrast	p.value	estimate
Group:Verb	Group	estar	BBil > CatDom	<.001***	-2.913
Group:Verb	Group	estar	SpaDom > CatDom	<.001***	-3.088
Group:Verb	Group	estar	SpaDom = BBil	1.000	-0.176
Group:Verb	Group	haver	CatDom = BBil	1.000	0.103
Group:Verb	Group	haver	CatDom = SpaDom	1.000	0.125
Group:Verb	Group	haver	BBil = SpaDom	1.000	0.022
Group:Verb	Verb	CatDom	haver > estar	<.001***	-5.017
Group:Verb	Verb	BBil	haver > estar	<.001***	-2.002
Group:Verb	Verb	SpaDom	haver > estar	<.001***	-1.804

Note: selected contrasts drawn from a total of 15 comparisons, at an adjusted alpha level of .003.

Table B4. Random effects for Poisson GLMM for second mention verbs

Groups	Name	Variance	Std. Dev.	Corr	
Participant	(Intercept)	0.465	0.682		
	Verbhaver	0.365	0.606	-0.72	
	Verbser	0.714	0.845	-0.47	0.12

Table B5. Output for Poisson GLMM for second mention verbs

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.400	0.303	-1.318	.187
GroupBBil	0.169	0.434	0.388	.698
GroupSpaDom	0.823	0.364	2.260	.024*
Verbhaver	2.249	0.308	7.306	<.001***
Verbser	1.281	0.354	3.620	<.001***
GroupBBil:Verbhaver	-0.546	0.445	-1.226	.220
GroupSpaDom:Verbhaver	-1.191	0.374	-3.185	.001**
GroupBBil:Verbser	0.177	0.505	0.350	.726
GroupSpaDom:Verbser	-1.104	0.445	-2.478	.013*

Note: the reference level for Group was *CatDom*, and for Verb it was *estar*.

Table B6. Selected pairwise contrasts with Bonferroni adjustment for Group × Verb interaction for second mention verbs

term	contrastfield	insidelevel	contrast	p.value	estimate
Group:Verb	Group	estar	BBil = CatDom	1.000	-0.169
Group:Verb	Group	estar	SpaDom = CatDom	.874	-0.818
Group:Verb	Group	estar	SpaDom = BBil	1.000	-0.657
Group:Verb	Group	haver	CatDom = BBil	1.000	0.377
Group:Verb	Group	haver	CatDom = SpaDom	1.000	0.369
Group:Verb	Group	haver	SpaDom = BBil	1.000	-0.008
Group:Verb	Group	ser	BBil = CatDom	1.000	-0.344
Group:Verb	Group	ser	CatDom = SpaDom	1.000	0.278
Group:Verb	Group	ser	BBil = SpaDom	1.000	0.622
Group:Verb	Verb	CatDom	haver > estar	<.001***	-2.242
Group:Verb	Verb	CatDom	ser > estar	.001**	-1.274
Group:Verb	Verb	CatDom	haver > ser	.015*	0.968

(Continued)

Table B6. (Continued)

term	contrastfield	insidelevel	contrast	p.value	estimate
Group:Verb	Verb	BBil	haver > estar	<.001***	-1.703
Group:Verb	Verb	BBil	ser > estar	.006**	-1.457
Group:Verb	Verb	BBil	haver = ser	1.000	0.247
Group:Verb	Verb	SpaDom	haver > estar	.001**	-1.054
Group:Verb	Verb	SpaDom	ser = estar	1.000	-0.177
Group:Verb	Verb	SpaDom	haver > ser	.091.	0.877

Note: selected contrasts drawn from a total of 36 comparisons, at an adjusted alpha level of .001.

Table B7. Output for Binomial GLMM for clitic EN

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	4.682	0.771	6.072	<.001***
GroupBBil	-4.469	0.905	-4.937	<.001***
GroupSpaDom	-5.599	0.902	-6.206	<.001***

Table B8. Random effects for Binomial GLMM for clitic EN

Groups	Name	Variance	Std. Dev.
Participant	(Intercept)	2.032	1.426

Table B9. Pairwise contrasts with Bonferroni adjustment for Group for clitic EN.

term	contrastfield	contrast	p.value	estimate
Group	Group	CatDom > BBil	< .001***	-4.470
Group	Group	CatDom > SpaDom	< .001***	-5.601
Group	Group	BBil = SpaDom	.181	1.130

Note: 3 contrasts in total at an adjusted alpha level of .0167.

Table B10. Output for Binomial GLMM for clitic HI

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	2.173	0.399	5.440	<.001***
GroupBBil	-0.910	0.495	-1.838	.066.
GroupSpaDom	-0.651	0.510	-1.277	.202

Table B11. *Random effects for Binomial GLMM for clitic HI*

Groups	Name	Variance	Std. Dev.
Participant	(Intercept)	0.478	0.692

Table B12. *Output for Binomial GLMM for clitic HI without instances of ser-hi*

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.442	0.403	1.098	.272
GroupBBil	-1.201	0.572	-2.102	.036*
GroupSpaDom	-1.208	0.623	-1.940	.052.

Table B13. *Random effects for Binomial GLMM for clitic HI without instances of ser-hi*

Groups	Name	Variance	Std. Dev.
Participant	(Intercept)	0.093	0.306

Table B14. *Pairwise contrasts with Bonferroni adjustment for Group for clitic HI without instances of ser-hi*

term	contrastfield	contrast	p.value	estimate
Group	Group	CatDom = BBil	.107	1.201
Group	Group	CatDom = SpaDom	.157	1.208
Group	Group	BBil = SpaDom	.181	0.007

Note: 3 contrasts in total at an adjusted alpha level of .0167.

Cite this article: Perpiñán, S. and Soto-Corominas, A. (2021). *Indirect structural crosslinguistic influence in early Catalan-Spanish bilinguals in adulthood: Predicate selection in Catalan existential constructions.* *Applied Psycholinguistics* 42, 1463–1502. <https://doi.org/10.1017/S0142716421000308>