

EMPIRICAL STUDY

Language Dominance Shapes Protracted Development of Nonpersonal Clitics (Even) in Cases of Widespread Bilingualism

Adriana Soto-Corominas ^a, Sílvia Perpiñán ^b,
and Scott James Perry ^b

^aUniversitat Autònoma de Barcelona ^bUniversitat Pompeu Fabra

Abstract: Little research has explored how language dominance may affect the development and ultimate attainment of morphosyntax in a situation of widespread and social bilingualism, where exposure to both languages starts early on and can be sustained over time. This study investigates the development of three nonpersonal clitics that have received little attention by empirical research cross-linguistically, namely the partitive and two different oblique clitics. We investigated whether and how the development and ultimate attainment of these clitics in Catalan were modulated by language dom-

CRediT author statement – **Adriana Soto-Corominas:** conceptualization; funding acquisition; methodology; investigation; data curation and analysis; writing – original draft preparation, review and editing. **Sílvia Perpiñán:** conceptualization; funding acquisition; methodology; investigation; data curation; writing – original draft preparation, review, and editing. **Scott James Perry:** data analysis; writing – original draft, review, and editing.

A one-page Accessible Summary of this article in nontechnical language is freely available in the Supporting Information online and at <https://oasis-database.org>.

This research project was supported by four main grants: Ontario Graduate Scholarship (PI: Soto-Corominas), I+D+i Project awarded by the Science and Innovation Ministry of Spain (#PID2020-114276GB-I00, PI: Perpiñán), Western University SSHRC Internal Research Grant, and Western Social Science and Humanities Research Board (SSHRB) Bridge Grant 36789 (PI: Perpiñán).

We kindly thank all participants, and, in the case of child participants, their families and their schools for making this project possible.

Correspondence concerning this article should be addressed to Adriana Soto-Corominas, Universitat Autònoma de Barcelona, Adriana.Soto@uab.cat.

The handling editor for this article was Aine Ito.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

inance. To this end, we tested child (ages 4–9) and adult Catalan–Spanish bilinguals with different degrees of language dominance. Our findings show robust differences among language dominance groups in the development and attainment of Catalan non-personal clitics. Specifically, despite continuous and widespread bilingualism, clitics emerge later in non-Catalan-dominant bilinguals, and later development does not lead to convergence with Catalan-dominant bilinguals, even in adulthood.

Keywords bilingualism; clitics; Catalan; language dominance

Introduction

Different linguistic properties emerge and are mastered at different times. Varying timelines have been accounted for in various ways: frequency of a construction/property in the input (Ambridge et al., 2015), amount of unambiguous evidence in the input supporting one possible analysis of that property (Yang, 2002), transparency of the form–function mapping (Gathercole, 2007), the linguistic domain that the target property belongs to (Tsimpli, 2014), or maturational constraints on the learner's part (Guasti, 2002). Likewise, when a category becomes productive in children's speech, not all of its exponents are equally productive or target-like. A case in point are clitics. Children acquiring languages with both subject and object clitics, such as French, produce and master subject clitics before object clitics (Hamann et al., 1996), and when object clitics first appear, they do not do so consistently, often coexisting with optional, nontarget clitic omission (Pérez-Leroux et al., 2018). Still, most studies have reported that object clitics are acquired by age 5, with optional omission up to this age (see, e.g., Gavarró et al., 2010, and Wexler et al., 2004, for Catalan; Jakubowicz & Rigaut, 2000, for French). However, very few of these studies investigate clitics beyond third-person accusative singular clitics, except for some studies on datives and partitives (Babyonyshev & Marin, 2006; Cardinaletti et al., 2021; Gavarró et al., 2011; Gavarró & Mosella, 2009; Soto-Corominas, 2018, 2021; Tarrés & Bel, 2015). This study investigates the development of the partitive and oblique clitics in Catalan bilingual children together with Catalan bilingual adults. The purpose of this research is twofold: On the one hand, we want to document the development of these less-studied clitics in Catalan; on the other hand, we want to investigate how language dominance affects the development of this morphosyntactic property through childhood and adulthood.

Researchers have tried to account for why bilinguals take longer to master certain properties of grammar than others (Cornips & Hulk, 2008; Gathercole & Thomas, 2009; Tsimpli, 2014). Several studies have shown that considering

language dominance is key to understanding the timing and outcomes of bilingual development (Cornips & Hulk, 2008; Gathercole & Thomas, 2009; Paradis, 2010). We take language dominance to denote the relative weight of the two (or more) languages in the mind of the bilingual. Although dominance has been defined and accounted for in a variety of ways (see Treffers-Daller, 2019, for a review), authors generally agree that it is a bidimensional construct made up of proficiency/competence and use/exposure (Montrul, 2015). Though it has long been argued that the notion of balanced bilingualism is unrealistic (Romaine, 1989), cases of societal bilingualism provide unparalleled opportunities for balanced bilingualism to emerge. In cases where bilingualism not only is widespread but also receives official support, the majority of the population has access to both languages in education and most domains and contexts. This is the case of Catalan and Spanish in Catalonia, a region in northeast Spain, where the present study takes place.

Thanks in large part to an obligatory Catalan-immersion educational system, all people who have completed obligatory schooling in Catalonia are fluent in both Catalan and Spanish. Bilinguals in Catalunya have access to input in the two languages from a diversity of sources, both oral (e.g., other people, TV, radio) and written (e.g., in the linguistic landscape, newspapers, literature). Given this context, we ask whether development is homogenous across bilinguals or whether language dominance determines the timing of acquisition and the ultimate attainment of nonpersonal clitics, a morphosyntactic phenomenon where Catalan and Spanish differ significantly. Specifically, we investigate three Catalan clitics (in bold) that are not present in Spanish, as shown in the examples (1–3).

- (1)

Tinc

tres coses.

En tinc

tres

(de coses).

Catalan

Tengo

tres cosas.

∅

Tengo

tres

(cosas).

Spanish

“have.1sg three things

have.1sg three

(of things).”

“I have three things. I have three (of them).”

(2)

Parlo

de tres coses.

En parlo

molt

(d’elles).

Catalan

Hablo

de tres cosas.

∅

Hablo

mucho

(de ellas).

Spanish

“talk.1sg about three things

talk.1sg lots

(about them).”

“I talk about three things. I talk lots (about them).”

(3)

Penso

en tres coses.

Hi penso

molt

(en elles).

Catalan

Pienso

en tres cosas.

∅

Pienso

mucho

(en ellas).

Spanish

“think.1sg about three things

think.1sg lots

(about them).”

“I think about three things. I think lots (about them).”

3

Language Learning 00:0, April 2026, pp. 1–35

1479722, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/lang.12026 by Spanish Cochran National Provision (Ministerio de Sanidad), Wiley Online Library on [16/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Dominance Effects on the Acquisition of Morphosyntax

Studies that have investigated the development of specific morphosyntactic properties in bilingual children growing up in areas with widespread and official bilingualism have noted that language dominance predicts not only accuracy but also timing of development, particularly in the minority language of the society. For example, for properties such as Welsh plural or gender morphology in Welsh–English bilinguals (Binks & Thomas, 2019; Thomas et al., 2014) or ergative case marking in Basque by Basque–Spanish bilinguals (Austin, 2007; Ezeizabarrena, 2012), bilingual children with less exposure to the language were found to be less accurate and lag behind children who were more dominant in the minority language.

Assuming a critical role of input in the development of morphosyntax (e.g., Gathercole, 2007; Gathercole & Thomas, 2009; Oller & Eilers, 2002; Thomas et al., 2014), language dominance is expected to affect such development insofar as it entails differences in patterns of exposure to each language, resulting in the nondominant language in simultaneous bilingual children (temporarily) lagging behind the dominant one (Schlyter, 1994; Yip & Matthews, 2000). Language dominance may also affect morphosyntactic development through its connection to crosslinguistic influence (CLI). The generally accepted idea that the bilingual's two linguistic systems are connected and in interaction predicts that some morphosyntactic structures may be over- or underused in one of the languages. Many studies have found that children are more likely to experience CLI in their nondominant language, especially when the nondominant language is not the societal language (Van Dijk et al., 2022).

Fewer studies have investigated if the initial differences attributable to language dominance disappear as development continues throughout the lifespan. A small body of research indicates that dominance can indeed affect bilinguals' end-state grammar. For instance, Thomas et al. (2014) investigated the development of plural marking in Welsh by Welsh–English bilingual adults. These participants were divided into three groups according to their home language: L1-Welsh speakers had been raised in households where only Welsh was used, and the simultaneous bilingual (2L1) group had been exposed to both languages in the home. The L2-Welsh group had attended Welsh-medium schooling but had used English only in the home. Results from an oral production task found that only the L1-Welsh group was at ceiling, whereas the 2L1 and the L2-Welsh groups averaged lower accuracy. Likewise, Perpiñán and Soto-Corominas (2021) investigated the expression of locative and existential constructions in Catalan in three groups of adult Catalan–Spanish early bilinguals with similar proficiency but different language dominance profiles. They

found qualitative differences among the three groups in the use of predicate selection regulated by semantics and information structure, suggesting that these three types of bilinguals might well have different linguistic representations. These studies suggest, first, that language dominance not only may affect the early stages of bilingual acquisition, but may result in different outcomes in morphosyntactic development, even after many years of continued exposure to both languages. Second, these divergent outcomes may be present in situations of widespread societal bilingualism, where bilinguals have consistent exposure to both languages.

In this study, we investigate the development and ultimate attainment of bilinguals born and raised in a situation of official bilingualism where exposure to the minority language occurs both in informal and formal (school) contexts and where use of and exposure to the two languages remains available throughout a bilingual's lifespan. By expanding on and comparing the datasets partially presented in two previous studies, Perpiñán's (2017) and Soto-Corominas's (2021), we attempt to give an account of the development across ages of Catalan clitics in children and adults, keeping constant comparable language dominance variables. In doing so, we investigate the development and ultimate attainment of bilinguals born and raised in a situation of official bilingualism.

Clitic Development in Children

The development of object clitics has received ample attention in research on Romance languages, though most studies have focused on accusative clitics only. Several findings from this body of work appear to hold cross-linguistically. First, from the early stages of clitic production, children place clitics in the expected (pre or postverbal) position following language-specific requirements (e.g., Guasti, 2002; Wexler et al., 2004). Second, in the initial stages of clitic production, it is not uncommon to find clitic omission and, less frequently, agreement errors, by which a mismatch in gender and/or number may occur between the clitic and its referent (Montrul, 2004; Prévost, 2009). Regarding omission specifically, studies have found that by age 4–5, children have overcome accusative and dative clitic omission in most Romance languages (Babyonyshev & Marin, 2006; Cardinaletti et al., 2021; Gavarró et al., 2011; Gavarró & Mosella, 2009; Varlokosta et al., 2016). Omission of the partitive clitic, although more persistent in earlier ages than that of accusative and dative clitics (Gavarró et al., 2006), is similarly overcome by age 5 in French, Italian, and Catalan monolingual children (Gavarró et al., 2011), seemingly reaching adult-like production by this age.

Importantly, in Portuguese, children have been shown to take longer to overcome the ungrammatical direct object omission stage (Varlokosta et al., 2016; see also Costa & Lobo, 2007). As opposed to other Romance languages, Portuguese allows for direct objects to be null as well as to be represented with accusative clitics in main clauses (Raposo, 1986). Since ungrammatical direct object omission lasts longer in European Portuguese than in other languages, it seems that the inconsistency of the production of accusative clitics in the input may delay the acquisition of this property (see also Pérez-Leroux et al., 2018).

Studies that have investigated object production and omission in bilingual children specifically have found that bilinguals tend to take longer to overcome the null object stage than their monolingual peers (Müller & Hulk, 2001; Yip & Matthews, 2000) and that this difference cannot be attributed solely to CLI (Pérez-Leroux et al., 2009). For this reason, a “bilingual effect” has been proposed for the development of objects: Bilingual children omit objects more often than their monolingual peers and show more protracted development (Pirvulescu et al., 2014).

It is important to reiterate that most of the studies reviewed above have focused on accusative clitics, and it is unclear to what extent findings are generalizable to other clitics, including oblique clitics, which we investigate in the present study. A comprehensive theory of language acquisition should account not only for the development of core grammatical elements like accusatives and datives, but also for less frequently studied forms such as partitive and oblique clitics. We discuss these forms in detail in the section below.

The Linguistic Phenomenon: Clitics in Catalan

The pronominal system of Catalan includes accusative, partitive, dative, and oblique clitics. These clitics can combine and fuse, and their surface forms vary depending on their position with regards to the verb, their phonetic context, and/or the specific clitic combination.

In this study, we focus on the partitive /n/, the oblique /n/, and the oblique /i/ due to our limited understanding of how these forms are acquired. These clitics in Standard Catalan are usually spelled as *en* and *hi*, respectively, and do not have a pronominal equivalent in present-day Spanish (i.e., they are equivalent to a null phonetic form in Spanish, as shown in examples (1)–(3) above).¹ Although analogous forms (*ne*, *ci/y*) exist and are productive in Italian and French, they have similarly been overlooked in research on clitic acquisition in Romance languages, with the exception of the partitive, which has been studied in monolingual 5-year-old children acquiring Catalan, Italian, and French (Gavarró et al., 2011).

Just as accusative and dative pronouns do, partitive and oblique pronominal clitics refer back to an antecedent in the discourse. When they refer to argumental complements, these pronouns are both syntactically and pragmatically obligatory. The partitive pronoun *en* can replace nonspecific subjects of unaccusative verbs, complements of the noun, and bare or quantified direct objects of transitive verbs as in example (4), but crucially, not definite direct objects, for which the accusative form is required. The partitive can refer back to a part of the phrase or its totality. In this study, we will focus on quantified direct objects. As indicated by the asterisk, the omission of the partitive clitic in these contexts results in an ungrammatical construction in Catalan.

(4) *Veig tres persones_i. No *(en)_i veig dues_i.*

“see.1sg three persons neg. EN_{part} see.1sg two”

“I see three people. I don’t see two.”

Oblique pronouns refer back to a prepositional phrase (PP). Catalan has two oblique pronouns, *en* and *hi*. Oblique *en* refers back to a PP introduced by the preposition *de* in example (5), whereas oblique *hi* refers back to a PP introduced by a preposition other than *de* in example (6). These two oblique clitics refer to prepositional arguments of the sentence, so the omission of these clitics changes the meaning of the sentence, as it implies a reduction in the thematic valence of the verb. For instance, example (5) without the clitic *en* does not violate any grammatical rule in isolation, but it can mean only that the subjects practice the activity of speaking very much. In a linguistic system with oblique clitics, the omission of the clitic cannot allow for the recovery or reference to the internal argument of talking, in this case, of money. Likewise, in example (6), the lack of the clitic *hi* could indicate only that parents think very much, but we would have no recovery of the topic of this thinking, in this case, of their children. For simplicity, we use asterisks throughout the study to indicate either a grammaticality violation or a pragmatically anomalous use/omission of the clitic.

(5) *Només parlen [de diners]_i. *(En)_i parlen sempre.*

“only speak.3pl of money EN_{obl} speak.3pl always”

“They speak only about money. They speak about it always.”

(6) *Els pares pensen [en els seus fills]_i. *(Hi)_i pensen constantment.*

“the parents think.3pl of the their children HI_{obl} think.3pl constantly”

“Parents think about their children. They think about them constantly.”

There are some key differences between these three clitics and the well-studied accusatives that could affect acquisition. First, oblique objects, though argumental, are mediated by a preposition and therefore maintain a more distant transitivity relationship with the verb compared to direct objects, which directly map onto core verbal arguments. According to Bosson (1985) and Moreno Cabrera (1987), this structural distinction implies a more distant semantic dependency on the part of oblique objects compared to direct objects. This difference means that the syntactic and semantic integration of oblique clitics could require additional processing steps due to their greater distance from the verb, potentially increasing cognitive load during parsing and production, thus affecting development, particularly in the early stages of language acquisition. In contrast, the partitive clitic refers to a quantified, indefinite, or bare direct object that is no more distant from the verb than an accusative clitic, which typically refers to a definite direct object. Thus, the primary distinction between the partitive and accusative clitics is in their definiteness rather than their syntactic position (Perpiñán, 2018; Roca, 1992). Consequently, no major differences in acquisition paths would be expected between these two direct object clitics, whereas developmental divergences may emerge between oblique clitics on the one hand and accusative and partitive clitics on the other.

If, however, CLI and frequency effects are the key variables, all three clitics examined here may pose acquisition difficulties. Although studies on clitic frequency in Catalan are lacking, it is plausible that oblique and partitive clitics occur less frequently than accusative arguments in child-directed speech. Furthermore, from a cross-linguistic perspective, Spanish lacks overt counterparts for the clitics under investigation here, but it does have equivalents for accusative and dative clitics. Given the potential for CLI, this asymmetry could influence developmental trajectories. Taken together, these considerations suggest that oblique and partitive clitics may present greater challenges for acquisition than their accusative counterparts.

The Present Study

This study aims to explore the associations between language dominance and the development of nonpersonal clitics in Catalan–Spanish bilinguals. To capture both development and outcomes, we consider children (ages 4–9) and adults born and raised in Catalunya. We asked three questions:

1. To what extent are the partitive and oblique clitics acquired by age 5 as typically described for other verbal clitics?

2. To what extent does language dominance in Catalan–Spanish bilinguals affect the timing of acquisition (emergence and development) for non-personal clitics in Catalan?
3. If we find language dominance effects, to what extent are these modulated by age? That is, do differences related to language dominance diminish or widen as children become adults?

Since dominance entails differences in exposure, we expected differences whereby less exposure would mean slower development. On the other hand, we expected differences to be less prominent in adults than in children given the fact that adults would have had more exposure to the target clitics over time. As such, it should be expected that adults should have attained full command of the given structures.

Data for this study were collected in two partly parallel projects in Catalunya: One investigating ultimate attainment of Catalan by Catalan–Spanish bilingual adults and one investigating development of Catalan and Spanish by Catalan–Spanish school-age children. As such, this study presents the data from 103 adult and 338 child Catalan–Spanish bilinguals. Although all participants had some knowledge of other languages (mostly English), none of them reported using a language other than Catalan and/or Spanish in the home at any point in their lives.

We present the results for this study as two distinct experiments. We present first the children's data, given that it makes more chronological sense. However, the experiment with adults was conducted earlier. Even though an effort was made to make the elicitation procedure for children resemble the adult one, adaptations were made in the design to make the test suitable for children. The most important adaptation was the elicitation procedure: Adults completed a question-answer format test, and children had to provide a correction to an utterance. This is explained in detail below. The switch from a question-answer format to an utterance-correction format was due to the fact that, during the piloting stage, children showed a clear preference for answering questions with felicitous and economical responses that did not require the use of the clitic.

Experiment 1: Children

Part of the data from this experiment was previously published in Soto-Corominas's (2021) study. Specifically, this earlier study focused only on children's production of partitive arguments in Catalan and in Spanish.

Method

Participants

Child participants ($N = 338$) had a mean age of 6;5 years ($SD = 1;6$), with a range: 4;0–9;10. Only for the purposes of visualization, participants were binned into age groups of 12 months: age 4 (including children between the ages of 4;0 and 4;11; $N = 46$), age 5 ($N = 58$), age 6 ($N = 72$), age 7 ($N = 67$), age 8 ($N = 55$), and age 9 ($N = 40$). All participants were recruited from schools in Central Catalunya.

Instruments

Questionnaire. Caregivers filled in a background questionnaire of 62 questions, available in Soto-Corominas's (2018) study. This questionnaire allowed us to determine participants' language dominance, classifying them into three groups: Catalan-dominant, balanced bilingual, or Spanish-dominant. Assignment to the dominance group was decided on three dimensions: age of onset of acquisition (AOA), language preference, and input quantity. For each of the three dimensions, participants were given a score between 0 and 1 point: 0 favored Spanish, .5 favored both languages, and 1 favored Catalan.

For AOA, parents were asked to indicate how old their child was when stable exposure to each language began. Zero points were assigned when the AOA of Spanish was before age 3 and the AOA of Catalan was after 3, .5 points were assigned when the AOA of both languages was before age 3 (i.e., simultaneous bilinguals), and 1 point was assigned when the AOA of Catalan was before age 3 and the AOA of Spanish was after 3.

For the second dimension, language preference, parents had to indicate their child's preferred language when expressing themselves, arguing, or playing. Parents could choose Spanish (which would assign 0 points), both Catalan and Spanish (.5 points), or Catalan (1 point).

The third dimension, input quantity, yielded a percent score. Parents were asked to indicate the language of use in spoken interaction between the child and 13 interlocutors (including mother, father, and siblings) and the number of hours watching TV and reading in Catalan. A higher percentage score would thus indicate more exposure to Catalan. The point assignment for input quantity was as follows: When participants scored 0–34% (suggesting little exposure to Catalan) they were given 0 points, when they scored 35–65% they were given .5 points, and participants with 65% or above were given 1 point.

The three dimensions had a maximum of 1 point each, thus 3 being the highest possible number of points. In order to divide participants into the three dominance groups, those with less than 1 point were classified as



Figure 1 Elicitation material example for the child version of the elicited production task.

Spanish-dominant, those with more than 2 points were classified as Catalan-dominant, and those in between were classified as balanced bilinguals.

Elicited Production Task. After the adult elicited production task was piloted on children (see below), materials required certain adaptations. First, the vocabulary items had to be adapted and materials had to be reconceptualized to be more child-friendly, so they were hand-drawn and contained no writing (Figure 1). The length of the test had to be reduced, so we included only 4 items per clitic. Besides the 12 experimental items, there were 20 distractors (which targeted other clitics), and 6 fillers. The procedure of the elicited production task was reframed to be that of an utterance-correction. Participants were presented with a laminated image, such as the one in Figure 1, that was visible only to participants and not the experimenter. The experimenter, playing the role of the mother of the child, provided an utterance that needed correction, such as *Jo no sé molt de cuina però el pare...* “I don’t know much about cooking but, dad....” The participant, who played the role of the child, had to correct the statement by looking at the picture. Children were familiarized with the characters in the images during a storytelling activity that preceded this elicited production task.

The selected verbs were chosen after consultation with two primary school teachers, avoiding those that would form a clitic cluster. These were the four predicates used to elicit each clitic:

Partitive *en*: *comprar* “to buy,” *menjar* “to eat,” *tenir* “to have,” *voler* “to want”

Table 1 Response type, together with a description and an example of each type of response

Response type	Description	Example
Target	The target clitic is employed.	<i>en sap.</i>
Omission	The context for the clitic to be used is given, but the clitic is omitted.	* <i>Ø sap.</i>
Overt	The NP or PP is repeated overtly, yielding an infelicitous response.	? <i>sap de cuina.</i>
Doubling	Both the clitic and the co-referential NP/PP appear.	* <i>en sap de cuina.</i>
Replacement	A clitic is used, but not the target one.	* <i>ho sap.</i>
Short answer	Yes/No response (statement is not corrected).	sí. “Yes.”
Other	A different construction is used.	<i>fa cuinetes.</i> “does cooking.”

Note. NP = noun phrase; PP = prepositional phrase.

Oblique *en*: *dubtar de* “to doubt,” *parlar de* “to talk about,” *saber de* “to know of,” *presumir de* “to brag about”

Oblique *hi*: *creure en* “to believe in,” *somiar en* “to dream of,” *pensar en* “to think about,” *jugar a* “to play”

Participant responses were coded according to the classification in Table 1 below. The examples in Table 1 illustrate the different response types for the Item in Figure 1, whose expected response was (7):

- (7) *sí que en sap.*
- “yes that EN_{obl} know.3sg”
- “Does know (about cooking).”

Procedures

Recruitment. After obtaining ethics approval from the Research Ethics Board of the University of Western Ontario, nine schools in Central Catalunya were recruited out of the 20 that had been contacted. Only those participants for whom signed parental consent and a filled questionnaire had been obtained were tested.

Testing. Participants were tested during school hours in a quiet space in their school. Prior to the elicitation task, participants were familiarized with the

main characters of the task and were asked to complete an elicited narration. In total, sessions took less than 30 minutes including breaks, and they were recorded for posterior transcription and scoring.

Data Analysis. All visualizations and analyses were performed in R (R Core Team, 2024). Visualizations were created using the *marginalEffects* (Arel-Bundock et al., 2024) and *ggplot2* packages (Wickham, 2016). All statistical models were Bayesian hierarchical models fit using the package *brms* (Bürkner, 2017), with interactions analyzed with the aid of the *marginalEffects* package and some additional comparisons of slopes calculated using the package *emmeans* (Lenth, 2024).

To model the probability of target-like clitic production in children, we fit logistic regression models (Bernoulli likelihood with logit link function). The population-level parameters estimated the effect of language dominance, clitic, age at testing, and education of the main caregiver. Language dominance was a three-level variable (Catalan-dominant, balanced bilingual, Spanish-dominant) as was clitic (partitive *en*, oblique *en*, oblique *hi*). Both variables were entered into the model as sum-coded variables. Age at testing was a continuous variable representing the age of the child in months at the time of testing, which was standardized before it was entered into our models. Last, we included education of the main caregiver as a control variable, and proxy for socioeconomic status, as it varied across dominance groups. This predictor was also standardized before it was entered into the regression model. The model also included the three-way interaction between age at testing, clitic, and language dominance. For group-level parameters, we included varying intercepts for item and participant, as well as varying slopes for clitic by participant and varying slopes for language dominance, age at testing, and education of the main caregiver by item. All priors were weakly informative, and their assumptions were verified through prior predictive simulation. We fit a second model with wider priors to evaluate the sensitivity of our posterior distributions to the prior distribution. All data and code used to fit the models and interpret the results are included in the OSF project page (<https://osf.io/wzn3d>). These materials include an additional analysis requested by an anonymous reviewer of the child data where our categorical variable of language dominance was substituted for one of three continuous measures: AOA, input quantity, and input quality. None of these models were estimated to have better out-of-sample predictive performance than the original model, and the interpretation of the effects is largely similar to the results below.

Table 2 Number of child participants per age and dominance group

Age	Catalan-dominant (<i>N</i> = 175)	Balanced bilingual (<i>N</i> = 68)	Spanish-dominant (<i>N</i> = 95)
4	27	7	12
5	28	13	17
6	42	12	18
7	30	15	22
8	30	9	16
9	18	12	10

Results

After classifying child participants into the three dominance groups, un-balanced groups obtained, probably as a result of the linguistic reality of Catalonia;² still, all dominance groups and ages were represented. The division of participants is given in Table 2.

Participant Characteristics

In Table 3, we present the characteristics for the child participants divided by dominance group, with participants of all ages amalgamated into each dominance group. Clear group differences were observed on all variables, even those variables that were not considered to classify participants.

All participants classified in the Catalan-dominant group were reported to prefer Catalan when expressing themselves. In the balanced bilingual group, 47 participants were rated as preferring both languages equally, 10 were reported to prefer Spanish, and 11 were reported to prefer Catalan. In the Spanish-dominant group, only two participants were reported to prefer both languages equally, and 93 were reported to prefer Spanish.

In terms of Catalan AOA, 171 participants of the Catalan-dominant group were exposed to Catalan from birth, and the remaining four were exposed to the language before age 2. In the Balanced bilingual group, 58 participants were exposed to Catalan from birth, eight were first exposed to it between ages 1 and 2, and the remaining two participants were exposed to it at age 3. This contrasts with the Spanish-dominant group, where 41 children were exposed to Catalan before age 2. Most of this group, 54 participants, were exposed to Catalan at or after age 3.

In terms of AOA of Spanish, all participants in the balanced and Spanish-dominant groups were exposed to Spanish from birth. In the Catalan-dominant

Table 3 Child participant characteristics, divided by dominance group

Variable	Catalan-dominant (<i>N</i> = 175)		Balanced bilingual (<i>N</i> = 68)		Spanish-dominant (<i>N</i> = 95)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Language preference	1	0	0.51	0.28	0.01	0.07
Catalan AOA (months)	0.41	2.85	3.00	8.13	26.46	16.15
Spanish AOA (months)	22.83	20.97	0	0	0	0
Input quantity in Catalan	85.43%	8.16%	56.60%	12.53%	22.43%	7.28%
Input diversity in Catalan	94.92%	8.47%	76.54%	14.79%	43.26%	20.28%
Parental education (years)	14.69	2.06	13.03	2.64	12.36	3.15

Note Language preference is presented as a continuous variable after converting categorical responses into numerical: 0 = Spanish preferred, .5 = both languages preferred equally, 1 = Catalan preferred. AOA = age of onset of acquisition.

group, 63 participants were exposed to Spanish from birth and 36 before age 3. The remaining participants (82) were exposed to Spanish at or after age 3.

On the dimension of input quantity, groups also differed. Whereas participants in the Catalan-dominant group were mostly exposed only to Catalan in their daily lives, the frequency of exposure to Catalan was much smaller in the other two groups.

Input diversity was determined by calculating the number of different sources (including interlocutors, TV, or books) from whom participants received input in Catalan. Diversity was measured by means of a percentage score, where 100 would indicate that all the possible sources in the questionnaire provided input in the given language, even if it was just some of the time. On average, Catalan-dominant participants heard at least some Catalan from the vast majority of the possible sources. This was also true, to a lesser extent, for balanced bilinguals. On the other hand, most Spanish-dominant participants received Catalan input from less than half of the possible sources.

As part of the questionnaire, the main caregiver was asked about their highest level of education attained. This was presented as categories that were later transformed into a number of years (e.g., primary education = 6 years, secondary education = 10 years, batxillerat (preuniversity preparation) = 12 years, university = 16 years in total). A total of 115 participants (66%) in the Catalan-dominant group had a main caregiver with a university degree, whereas this was true for only 23 balanced bilinguals (34%) and 27 Spanish-dominant participants (28%). On the other hand, three Catalan-dominant participants (1.71%) and three balanced bilinguals (4.41%) had parents with only primary education, compared to 14 (15%) of the participants in the Spanish-dominant group. These differences were controlled for in the statistical analyses.

Analysis of Child Data

Figure 2 below shows the children's overall percentage of target and nontarget productions for each of the three clitics. The analysis of the Bayesian models fit to analyze the present data is interpreted in the following sections with the aid of model-based predictions made using the package *marginalEffects*. This makes predictions on the back-transformed probability scale between 0 and 1. The presentation of these interpretations is narrative in structure and nonexhaustive. We focus on interpreting whether or not the relevant 95% interval includes 0 for the benefit of the reader (i.e., the 95% credible interval not including 0 means that there is at least a 95% probability that there is an effect). Further details regarding the models and their interpretation, where we look at

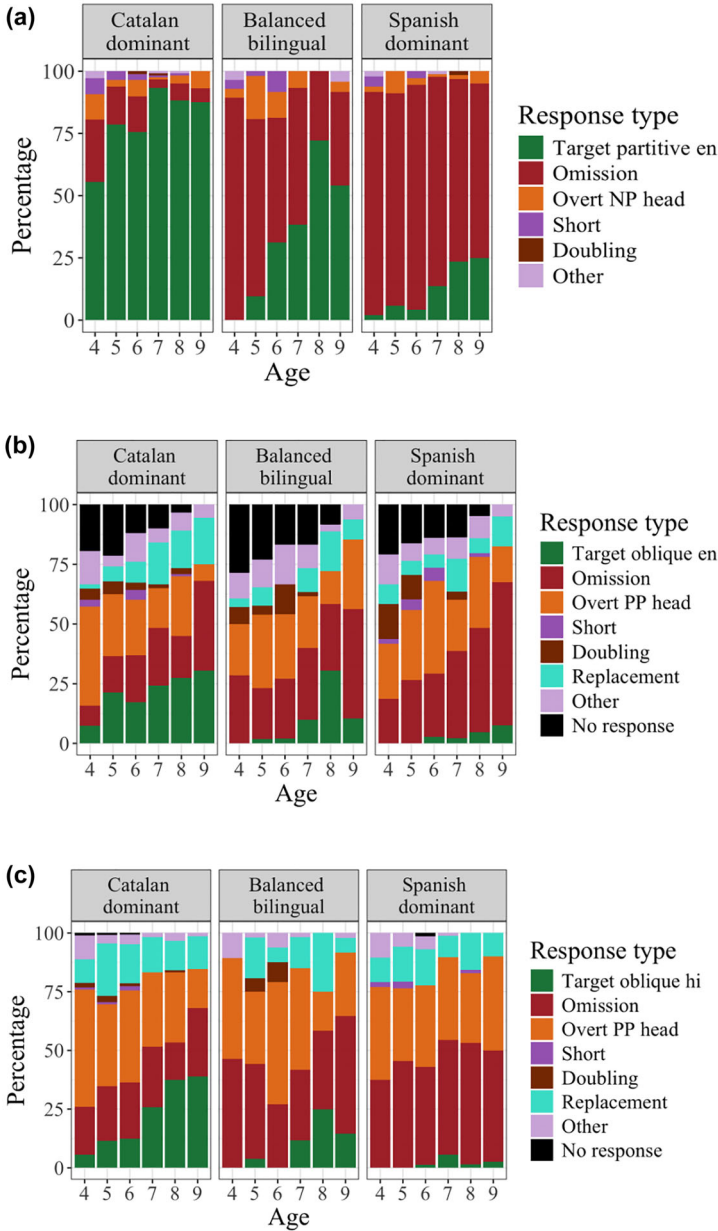


Figure 2 Percentage of each response type (y-axis) according to child participants' language dominance for partitive *en* (A), oblique *en* (B), and oblique *hi* (C). NP = noun phrase; PP = prepositional phrase.

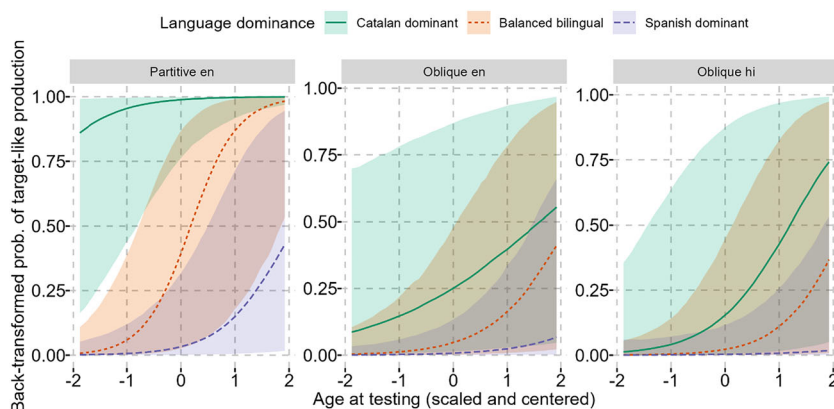


Figure 3 Model-based predictions of target-like clitic production according to age at testing, language dominance, and clitic, for an average participant and item. A scaled value of -2 would approximately correspond to a child that is about to turn 4 years of age and a scaled value of 2 would be a child that has just turned 10.

our level of uncertainty in a more gradient fashion, are given in the additional materials on the OSF project page (<https://osf.io/wzn3d>).

Many of the posterior distributions in our model of child data were quite wide, yielding uncertainty regarding the target-like production of some clitics for certain dominance groups (see Figure 3). However, some patterns emerged. For the partitive *en*, we are confident that Catalan-dominant children outperformed the other two dominance groups at the youngest ages tested. At the oldest ages tested, we are more than 95% confident of a three-way distinction between the dominance groups in their production of the partitive *en*. In addition, older children were more likely to produce target partitive *en* compared to younger children for all dominance groups and showed no strong evidence that the slope of age at testing is different for any of the groups.

Moving to oblique *en*, at the youngest ages in our sample, Catalan-dominant children were estimated to produce more target-like clitics than the other two groups, but the differences were much smaller compared to the partitive *en*. In the oldest children in our sample, we are more than 95% confident that Catalan-dominant children produced more target-like productions than Spanish-dominant children but are less certain that balanced bilinguals performed differently than either of the other groups. As with the partitive *en*, we are more than 95% confident that older children produced more target-like clitics than younger children and are uncertain of any differences in this improvement based on language dominance.

Lastly, for oblique *hi*, the youngest children in the sample performed similarly regardless of language dominance, and we are uncertain of any differences. At the oldest ages, both Catalan-dominant and balanced bilinguals produced more target clitics than Spanish-dominant children, with the difference between Catalan-dominant and balanced bilinguals being uncertain. We are not confident that older Spanish-dominant children produced more target-like clitics than younger children, in contrast to the other two groups. We are also uncertain that the slope of the age effect differs according to dominance.

Experiment 2: Adults

A portion of the current dataset was previously presented in Perpiñán's (2017) study. That earlier study, which also reported data from a grammaticality judgment task, included only a subset of the participants analyzed here (54 participants out of 102).

Method

Participants

Adult participants ($N = 102$) were sampled from two locations: Central Catalunya ($N = 48$) and the Metropolitan area of Barcelona ($N = 54$). Participants were between the ages of 18 and 45 ($M = 27.96$, $SD = 5.62$). We excluded all participants who did not complete secondary education to avoid any effects due to lack of formal instruction. We collected data from relatively young adults as we wanted to make sure that they had been schooled in Catalan (given that Catalan immersion in school was first implemented in the academic year of 1983–1984).

Instruments

Questionnaire. Participants responded to the Language Dominance Questionnaire (LaDoQ; Perpiñán & Soto-Corominas, 2021) using an online survey platform. This questionnaire was made up of 47 questions that yielded information on the current and past language experiences of participants, language identification, AOA of each language, quality of the input received, and self-rated proficiency in Catalan.

Like the children, adult participants were assigned to one of three dominance groups: Catalan-dominant, balanced bilinguals, and Spanish-dominant. The group categorization was made based on participants' responses targeting language use with family and friends, origin of close relatives, language identification and language preference, and place of residence (since some areas of Catalonia favor more or less the use of Catalan).

Qui parla de futbol sempre?



Figure 4 Experimental item for adult participants eliciting oblique *en*.

Elicited Oral Production Task. The elicited production task was presented to participants using a PowerPoint presentation. Participants were asked to answer a question written on a slide by including, minimally, the words provided to them in a textbox (the verb in the infinitive form and the subject of the sentence). The question was also read out loud by the experimenter. One example slide with the question “Who talks about soccer all the time?,” which elicited oblique *en*, is provided in Figure 4. Participants provided their answers orally and were recorded for posterior transcription and analysis.

Each target clitic was elicited by five different items, for a total of 15 items. Distractors targeting other clitics ($k = 15$) and fillers targeting other structures ($k = 25$) were also included, for a total of 55 items. The verbs that were used to elicit the three clitics were:

Partitive *en*: *tenir* “to have” ($\times 2$), *fer* “to do,” *haver* “there is/are,” *comprar* “to buy”

Oblique *en*: *parlar de* “to talk about,” *dubtar de* “doubt,” *dependre de* “to depend,” *presumir de* “boast about,” *saber de* “to know about”

Oblique *hi*: *contribuir a* “to contribute,” *somiar en* “to dream of,” *accedir a* “to access”
renunciar a “resign from,” *pensar en* “to think about”

Participants' responses were transcribed and later coded for the response type into one of the categories shown in Table 1 above.

Procedures

Recruitment. After obtaining ethics approval from the Research Ethics Board of the University of Western Ontario, adult participants were recruited via word-of-mouth and received 10€ as compensation.

Testing. Adult participants were tested in a quiet space of their choosing, most often in public library rooms. Testing sessions lasted approximately 60 minutes since participants completed other tasks unrelated to this one.

Data Visualization and Analysis. We employed the same data visualization and analysis protocols as for the child data above.

Results

Participant Characteristics

Table 4 presents the characteristics for the adult participants divided into the three dominance groups. AOA was determined by asking participants to select the age range when they were first exposed to each language. The possible ranges were 0–3, 3–6, 6–10, 10–15, and 15 and older, coinciding with educational periods. The means in Table 2 were calculated by considering the lowest age of the range. In terms of Catalan AOA, all Catalan-dominant participants declared having been exposed to this language between birth and age 3. This was also true for all but seven of the balanced bilingual participants. In the Spanish-dominant group, only 19 participants said they had been exposed to Catalan before age 3. The majority of this group, 25 participants, were exposed to Catalan between the ages of 3–6. In terms of Spanish AOA, there were only two Catalan-dominant participants who were exposed to Spanish before age 3, nine Catalan-dominant children were exposed to Spanish between 3–6, and 10 were exposed to it at age 6 or later, in primary school. In the balanced bilingual group, all participants but three were exposed to Spanish before age 3, whereas the rest were exposed to it at or after age 3. In the Spanish-dominant group, only one participant was exposed to Spanish between the ages of 3 and 6, and the remaining were exposed to it before age 3.

The variable input quantity was calculated by averaging language use with the mother, father, siblings, cousins, friends, and partner (when applicable). The mean was calculated after converting categorical responses into numerical: Spanish = 0; both languages = .5; Catalan = 1. The variable input diversity was created by considering how many of the six possible sources provided some input in Catalan.

Table 4 Participant characteristics for the adult participants

Variable	Catalan-dominant (<i>N</i> = 21)		Balanced bilinguals (<i>N</i> = 32)		Spanish-dominant (<i>N</i> = 49)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Catalan AOA (months)	0	0	0.77	1.54	2.22	2.17
Spanish AOA (months)	4.76	3.30	0.29	0.90	0.06	0.43
Input quantity in Catalan	0.98	0.04	0.52	0.25	0.11	0.11
Input diversity in Catalan	94.44%	10.97%	55.73%	26.30%	18.71%	17.56%
Self-assessed bilingual profile ^a	0.98	0.11	0.62	0.28	0.15	0.23
Language identity	1	0	0.75	0.34	0.16	0.26

Note. Presented as a continuous variable after converting categorical responses into numerical: Spanish-dominant = 0; balanced = .5; Catalan-dominant = 1. AOA = age of onset of acquisition.

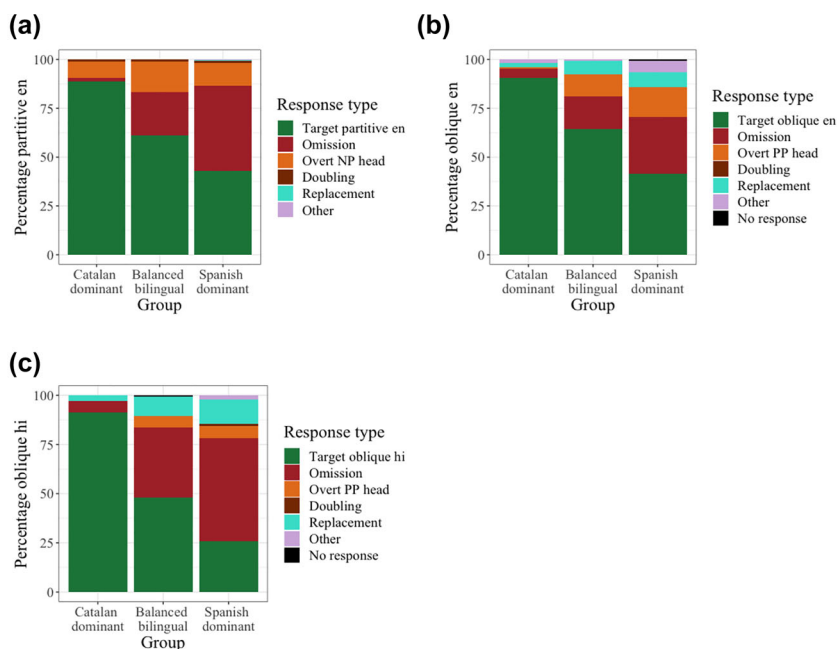


Figure 5 Percentage of each response type (y-axis) according to adult participants' language dominance for partitive *en* (A), oblique *en* (B), and oblique *hi* (C). NP = noun phrase; PP = prepositional phrase.

We asked about participants' self-assessed bilingual profile by asking them whether they considered themselves more Catalan-speaking (1 point), balanced (.5 points), or more Spanish-speaking (0 points). Language identity, on the other hand, was determined by asking participants whether they identified with Catalan (1 point), with Spanish (0 points), or with both languages equally (.5 points).

Analysis of Adult Data

Figure 5 shows the adults' overall percentage of target and nontarget productions for each of the three clitics. Given our third research question, which asked about whether age modulates dominance-related differences, our last model compared the oldest children (age 9) with the adult data in a model with a three-way interaction between group (9-year-old children vs. adults), clitic, and language dominance. All variables were sum-coded. The group-level predictors again included varying intercepts by item and participant, as well

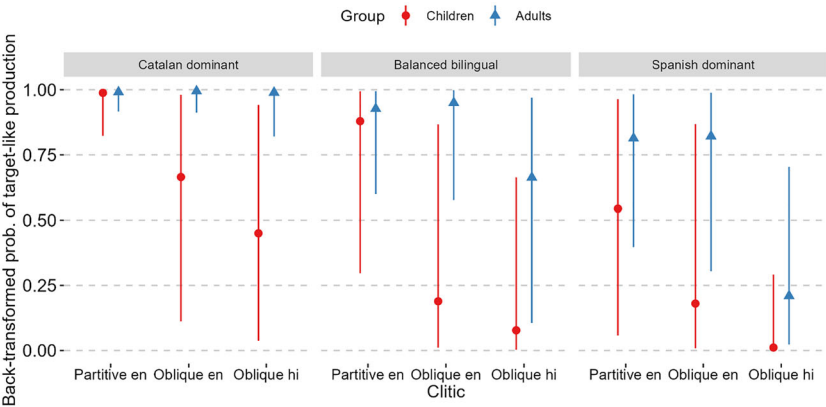


Figure 6 Back-transformed predicted probabilities of target-like production for an average item and participant. Predictions are made for each combination of clitic, dominance group, and age group.

as varying slopes for clitic by participant and varying slopes for language dominance by item. The goal of this analysis is to determine whether differences associated with language dominance diminish or widen as children become adults.

The predicted probability of a target-like clitic production in this model is presented in Figure 6. Overall, Catalan-dominant adults performed at ceiling for all clitics. We are less certain about the performance of balanced bilinguals, who show increased variability for all three clitics, although it is likely that we could interpret their performance at ceiling for partitive and oblique *en* (but not for oblique *hi*). For Spanish-dominant adults, we see more nontarget-like productions in all three clitics, especially of the oblique *hi*.

In comparing the two age groups (9-year-olds vs. adults), starting with the partitive *en*, we are unsure that adults produce more target-like clitics compared to children for any dominance groups. For the oblique *en*, we are more than 95% confident that adults produced more target-like clitics than children for all dominance groups. For the oblique *hi*, we are more than 95% confident that adults produce more clitics as compared to children for Catalan-dominant and balanced bilinguals, but we are less certain of this improvement for Spanish dominant speakers.

Discussion

This study explored the acquisition, development, and ultimate attainment of Catalan nonpersonal clitics in a bilingual context and whether these

14679722, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/lang.12036 by Spanish Cochran National Provision (Ministerio de Sanidad), Wiley Online Library on [16/06/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

processes were modulated by language dominance. To this extent, we tested child and adult Catalan–Spanish bilinguals with different degrees of language dominance.

Age 5 as the Age of Acquisition?

Our first research question asked about the age of acquisition of these clitics, given the previous research showing that (accusative and dative) object clitics are mastered by typically developing children by age 5 (e.g., Cardinaletti et al., 2021; Gavarró & Mosella, 2009; Varlokosta et al., 2016). Our findings demonstrate that the acquisition of clitic pronouns is not a uniform process, and that (a) some pronominal clitics are acquired before and more robustly than others, and (b) that any discussion of ages of acquisition in bilingual settings has to consider language dominance. Given the importance of language dominance, we discuss specific dominance-related timelines of acquisition in the next section. We focus here on the overall timing differences between the clitics.

The differences between the acquisition timing of partitive and accusative clitics (acquired relatively early) compared to oblique clitics (acquired very late) cannot be explained by an increased pragmatic load in certain pronominal uses, or a higher involvement of interface phenomena, as these clitics require similar discourse anchorage. In this sense, they make a similar pragmatic and discourse contribution. Even though omission is long-lasting in non-Catalan-dominant bilinguals (see discussion below), we hypothesize that this is not the result of a general bilingualism effect, but an overall structural consideration together with input effects. Concerning structural considerations, some clitics seem to constitute a more core property of the language than others, and this has consequences for their acquisition timeline, as proposed by Tsimpli (2014). That is, linguistic structure has a significant impact on the timing of acquisition. Accusative and partitive clitics are direct objects of transitive verbs, closely related to the verbal action, and constitute core properties of the language. Consequently, the results for partitive *en*, displayed in Figure 2A, show children either producing the clitic or omitting it from age 4, with very few alternative types of responses. Oblique objects, on the other hand, despite also being verbal arguments, are in a preposition-mediated, more distant transitivity relationship with the verb. This different degree in transitivity between accusative/partitive arguments, on the one hand, and oblique ones, on the other hand, may yield different degrees of distance processing and structural and semantic complexity, with consequences for their acquisition.

Figures 2B and 2C show that, especially at the youngest ages tested but throughout childhood, children produced a very sizeable percentage of overt

PPs. We interpret this result as indicating that children understand the argument structure of the verbs but have not yet acquired the conditions for cliticization of the PP object and thus unfelicitously repeat the overt PPs as an avoidance strategy. The contrast between the partitive and oblique clitics is seen most clearly when comparing partitive and oblique *en*, which, despite their homophony, are acquired and mastered at different times. These structural differences and their consequences for timing in acquisition are magnified by language dominance effects (see section below).

An alternative, and not mutually exclusive, explanation for why the oblique clitics develop later than accusatives and partitives in children is that oblique clitics may have a lower frequency in child-directed speech than the other clitics. Given that the frequency of a certain structure has been shown to affect its acquisition (Diessel & Hilpert, 2016), it is possible that frequency effects are connected to the results observed here. Frequency effects remain as a topic of future research, as there are currently no observational studies on the frequency of these clitics in Catalan.

Language Dominance Effects

Our second research question sought to determine whether language dominance would be associated with differential timelines of acquisition for the three clitics under study. Our results clearly suggest that both the emergence of the clitics and their ultimate attainment show differences according to dominance, whereas the later development of the clitics (i.e., the changes in target productions between 9-year-olds and adults) seems to be less contingent on language dominance and more dependent on the target clitic. That is, no groups showed evidence of improvement with respect to partitive *en* between age 9 and adulthood. However, that yielded different end states in the adult groups, given that Spanish-dominant bilinguals remained in a state of inconsistent production (i.e., optionality). For oblique *en*, all three dominance groups made significant gains between age 9 and adulthood, but again the Spanish-dominant adult group fell short of producing the target clitic consistently. The only dominance-related difference in growth is observed with respect to oblique *hi*, since the Spanish-dominant group did not demonstrate evidence of growth between age 9 and adulthood, unlike the other two groups. This differential growth, as well as the different baseline group performance at age 9, again yielded differences between the three groups in adulthood.

Our third, related, research question asked about the relationship between language dominance effects and age, that is, whether attested initial differences mainly due to disparities in input availability would disappear as continued

exposure to the minority language was maintained. Whereas we would need longitudinal data to properly assess individual development, we can conclude that initial differences in performance between the three profiles do not disappear as non-Catalan dominant speakers seem to have increased variability in target-like production in adulthood, with oblique *hi* showing the most optionality in its use.

Focusing on the appearance and development of clitics across ages in childhood, we observed that partitive *en* was acquired by age 5 by Catalan-dominant speakers, in line with what had been reported for the acquisition in “five-year-old children, *native* speakers of Catalan, French, and Italian” (Gavarró et al., 2011, p. 7; emphasis ours). This is also comparable to the results for accusative clitics cross-linguistically (Varlokosta et al., 2016). Gavarró et al. (2006), who compared the development of accusative and partitive clitics in Catalan children at earlier ages (1;2–4;1), concluded that children go through an optional omission stage of the partitive clitic, similar to the one described for accusative clitics, but that lasted longer. Indeed, we found what appears to be developmental omission in Catalan-dominant bilinguals until age 6, with 14% of omissions. In addition, our data captured the first appearance of the partitive clitic in balanced (around age 6) and Spanish-dominant bilinguals (around age 7), who started to use the partitive *en* markedly later than Catalan-dominant participants. This protracted acquisition and development in non-Catalan-dominant bilinguals result in an extended omission period that remains in adulthood, yielding “stable” optionality (Sorace, 2003), which should be distinguished from the developmental optionality previously described for Catalan-dominant children. For instance, in the balanced bilingual children, who are also *native* Catalan speakers (emphasis ours), partitive *en* starts to appear at age 6 and is not produced more or less consistently until age 8, with prolonged omission rates that stretch into adulthood. Spanish-dominant bilinguals, on the other hand, seem to never master the partitive clitic, not even in adulthood, after having been schooled in Catalan. Thus, our data not only show that not all Catalan speakers born and raised in a situation of widespread Catalan–Spanish bilingualism master this clitic at the landmark age of 5, but that protracted development, omission, and stable optionality are long-lasting characteristics of bilingual Catalan.

An important question is whether this late appearance and persistent omission rates in the non-Catalan dominant speakers can be considered the result of maturational constraints, as argued in earlier stages for accusative clitics (Gavarró et al., 2010), or whether it can be attributed to bilingualism effects. Our data show that this optional stage for partitive clitics not only lasts until

older childhood ages in balanced bilingual children, but that it is also present, to some extent, in adult balanced bilingual speakers. Given the late ages at which partitive clitics appear in these bilinguals, we believe that maturational constraints are not at the root of these different timelines and that omission and optionality have become defining characteristics of the Catalan grammar of these bilingual speakers and probably also of their input providers.

An Integrated Bilingual Grammar

Considering the results observed in Figure 5, it is apparent that bilinguals are exposed to variable input with respect to the three target clitics. Undoubtedly, there would be no group that is categorically shielded from variable input (i.e., even Catalan-dominant speakers are exposed to Catalan produced by balanced or Spanish-dominant speakers throughout their lives). Inconsistent input alone can be considered to make the task of acquiring these three clitics a more difficult task than for accusatives, for which Catalan speakers receive consistent input (see Costa & Lobo, 2007, for the acquisition of Portuguese accusatives, which shows variability in some syntactic contexts; Yang, 2002). However, we still need to account for why only Catalan-dominant speakers seem to arrive at the target representation for the three clitics.

An increasing number of studies support the holistic view of bilingualism that the grammars of a bilingual are not fully separable and that this complex integrated bilingual grammar allows for the interaction of the languages in the shape of CLI (Grosjean, 2024; Van Dijk et al., 2022). We believe that our results point in the same direction. Even in highly fluent bilingual adults, their Catalan and Spanish grammars do not seem to be fully independent.

Spanish does not have a phonetically overt counterpart to the Catalan clitics under study. As such, it is at least plausible to consider the use of the grammatical Spanish-like omission in Catalan as a case of CLI. As attested in our results, ungrammatical omission in Catalan was not rare. In adult and fluent bilinguals, omission of the clitic was not only frequent but, in some cases, the preferred response over the target clitic. In our sample, the degree of omission was associated with the degree of dominance. If we consider that reliance on omission is the result of CLI, then the degree of permeability to CLI is proportional to the degree of Spanish dominance. There are a few reasons why dominance would predict CLI. One of these reasons is exposure. Balanced bilinguals and Spanish-dominant bilinguals activate Spanish more than Catalan-dominant bilinguals on a day-to-day basis. It may be the case that the equivalent Spanish clitic-less construction primes omission in Catalan (Fernández et al., 2017). When faced with two options (clitic vs. omission) in Catalan,

some bilinguals thus choose omission, which produces the greater structural similarity between Spanish and Catalan. In previous research, bilinguals have been shown to be vulnerable to CLI especially when their nondominant language is also not the societal language. Given the minority status of Catalan, it is unsurprising that Spanish-dominant bilinguals are thus the most susceptible to CLI in their nondominant language. In addition, even though all groups are exposed to variable input with regards to the clitics of interest, it may be the case that Catalan-dominant bilinguals are exposed to input that is less variable (with a higher proportion of input provided by Catalan-dominant speakers themselves, such as their caregivers). Whether the acquisition by Catalan-dominant bilinguals is facilitated by a lack of CLI into their stronger language, by way of being exposed to input that is less variable with regards to the target structure, or by both variables at the same time remains a topic for further investigation.

Finally, we would like to add that the group differences have consequences for a potential language change in progress of Catalan. As of now, the clitic-less innovations produced by the balanced and Spanish-dominant bilingual groups currently coexist with the clitic productions of the Catalan-dominant group and, to a lesser extent, of the other two groups, leading to optionality in the construction. Given the fact that demographically, Catalan-dominant bilinguals are the minority (EULP, IDESCAT, 2023), this could lead to omission being the preferred option over time, eventually leading to language change (Fernández et al., 2017; Meisel et al., 2013)

Limitations and Conclusions

The present study is not without limitations. This was a cross-sectional study, where participants in different age groups were not the same. The ideal design of this study would be longitudinal. A longitudinal design would also allow us to track potential changes in language dominance during a bilingual's lifespan. In addition, child and adult participants were not tested using the exact same protocol or items, given that this was not a possibility due to the need to limit the testing complexity in the youngest participants.

Although it is risky to infer growth from cross-sectional data, it is clear that the dominance-related initial delays in clitic development observed in this study are not fully overcome later on with increased exposure, at least not during childhood or the school years we investigated (until age 9). Instead, certain structures improve in adolescence or adulthood. Regardless, differences are not fully neutralized between the different language dominance groups. These outcomes indicate that continued exposure also in adulthood

is crucial for performance in a minority language such as Catalan and suggest that complex morphosyntax may not stabilize until later in life, even in dominant (*native*, emphasis ours) speakers. These cases of protracted development and fossilization of certain language-specific morphosyntactic properties are in line with findings from other minority linguistic communities (Binks & Thomas, 2019; Ezeizabarrena, 2012; Gathercole, 2007; Gathercole & Thomas, 2009).

In summary, our results have shown that, given (or despite) the sociolinguistic reality of Catalonia, it is crucial to factor in language dominance when investigating the acquisition of Catalan morphosyntax. Our findings show robust differences among language dominance groups in the development and attainment of Catalan nonpersonal clitics. These differences are not fully mitigated with increasing age. This indicates that school support of a minority language is a necessary but not sufficient condition for all the properties of the minority language to be fully mastered. In terms of linguistic structures, this study suggests that partitive and oblique clitics are not mastered by age 5, when the acquisition of other object clitics appears to have been completed (e.g., Varlokosta et al., 2016). In addition, this study shows that even homophonous clitics such as *en* follow distinct trajectories. Partitive *en*, as a direct object of the verb, is a core linguistic property that develops earlier and more robustly than oblique *en* (and oblique *hi*). Oblique clitics, on the other hand, are vulnerable morphosyntactic properties that are acquired very late, if at all.

Final revised version accepted 13 February 2026

Open Research Badges



This article has earned an Open Data and Open Materials badge for making publicly available the components of the research methods needed to reproduce the reported procedure. All data and materials that the authors have used and have the right to share are available at <https://osf.io/wzn3d>. All proprietary data and materials have been precisely identified in the manuscript.

Notes

- 1 Campos (1986) proposed that Spanish indefinite direct objects project an empty category in the form of a trace of an operator. To our knowledge, no similar proposal has been made for oblique arguments, leaving open the question of

whether they involve an empty projection of the relevant category in Spanish or, alternatively, no syntactic projection at all.

- 2 According to the *Enquesta d'Usos Lingüístics 2023* (IDESCAT, 2023), less than 6% of the Catalan population over 15 years old reported having both Spanish and Catalan as their initial languages.

References

- Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language*, 42(2), 239–273. <https://doi.org/10.1017/S030500091400049X>
- Arel-Bundock, V., Greifer, N., & Heiss, A. (2024). How to interpret statistical models using marginaeffects for R and Python. *Journal of Statistical Software*, 111, 1–32. <https://doi.org/10.18637/jss.v111.i09>
- Austin, J. (2007). Grammatical Interference and the acquisition of ergative case in bilingual children learning Basque and Spanish. *Bilingualism: Language and Cognition*, 10(3), 315–331. <https://doi.org/10.1017/S1366728907003094>
- Babyonyshev, M., & Marin, S. (2006). Acquisition of pronominal clitics in Romanian. *Catalan Journal of Linguistics*, 5, 17–44. <https://doi.org/10.5565/rev/catjl.71>
- Binks, H. L., & Thomas, E. M. (2019). Long-term outcomes for bilinguals in minority language contexts: Welsh–English teenagers' performance on measures of grammatical gender and plural morphology in Welsh. *Applied Psycholinguistics*, 40(4), 1019–1049. <https://doi.org/10.1017/S0142716419000110>
- Bossong, G. (1985). *Empirische Universalienforschung: Differentielle Objektmarkierung in den neuiranischen Sprachen*. Narr. <https://doi.org/10.5281/zenodo.4697660>
- Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software*, 80, 1–28. <https://doi.org/10.18637/jss.v080.i01>
- Campos, H. (1986). Indefinite object drop. *Linguistic Inquiry*, 17(2), 354–359. <https://www.jstor.org/stable/4178492>
- Cardinaletti, A., Cerutti, S., & Volpato, F. (2021). On the acquisition of third person dative clitic pronouns in Italian. *Lingue e Linguaggio*, 2. <https://doi.org/10.1418/102817>
- Cornips, L., & Hulk, A. (2008). Factors of success and failure in the acquisition of grammatical gender in Dutch. *Second Language Research*, 24(3), 267–295. <https://doi.org/10.1177/0267658308090182>
- Costa, J., & Lobo, M. (2007). Clitic omission, null objects or both in the acquisition of European Portuguese? In S. Baauw, F. Drijkoningen, & M. Pinto (Eds.), *Romance Languages and Linguistic Theory 2005: Selected papers from 'Going Romance', Utrecht, 8–10 December 2005* (pp. 59–72). John Benjamins Publishing Company. <https://doi.org/10.1075/cilt.291.06cos>

- Diessel, H., & Hilpert, M. (2016). Frequency effects in grammar. In *Oxford Research Encyclopedia of Linguistics*.
<https://doi.org/10.1093/acrefore/9780199384655.013.120>
- Ezeizabarrena, M.-J. (2012). The (in)consistent ergative marking in early Basque: L1 vs. child L2. *Lingua*, 122(3), 303–317. <https://doi.org/10.1016/j.lingua.2011.11.009>
- Fernández, E. M., De Souza, R. A., & Carando, A. (2017). Bilingual innovations: Experimental evidence offers clues regarding the psycholinguistics of language change. *Bilingualism: Language and Cognition*, 20(2), 251–268.
<https://doi.org/10.1017/S1366728916000924>
- Gathercole, V. C. M. (2007). Miami and North Wales, so far and yet so near: A constructivist account of morphosyntactic development in bilingual children. *International Journal of Bilingual Education and Bilingualism*, 10(3), 224–247.
<https://doi.org/10.2167/beb442.0>
- 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.
<https://doi.org/10.1017/S1366728909004015>
- 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* (pp. 27–46). Shaker.
- Gavarró, A., & Mosella, M. (2009). Testing syntactic and pragmatic accounts of clitic omission. In J. Crawford, K. Otaki, & M. Takahashi (Eds.), *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA 2008)* (pp. 25–35). Cascadilla Proceedings Project.
<http://www.lingref.com/cpp/galana/3/abstract2303.html>
- Gavarró, A., Torrens, V., & Wexler, K. (2010). Object clitic omission: Two language types. *Language Acquisition*, 17(4), 192–219.
<https://doi.org/10.1080/10489223.2010.509262>
- Grosjean, F. (2024). *On bilinguals and bilingualism*. Cambridge University Press.
- Guasti, M. T. (2002). *Language acquisition: The growth of grammar*. The MIT Press.
- Hamann, C., Rizzi, L., & Frauenfelder, U. H. (1996). On the acquisition of subject and object clitics in French. In H. Clahsen (Ed.), *Generative perspectives on language acquisition: Empirical findings, theoretical considerations and crosslinguistic comparisons* (pp. 309–334). John Benjamins. <https://doi.org/10.1075/lald.14.13ham>
- IDESCAT. (2023). *EULP, Enquesta d'usos lingüístics de la població*.
https://llengua.gencat.cat/ca/serveis/dades_i_estudis/dades/enquesta-eulp/2023/index.html

- Jakubowicz, C., & Rigaut, C. (2000). The acquisition of nominative and object clitics in French. *The Canadian Journal of Linguistics/La Revue canadienne de Linguistique*, 45(1–2), 119–157.
- Lenth, R. (2024). *_emmeans: Estimated marginal means, aka least-squares means_*. [Computer software] (Version 1.10.2) <
<https://CRAN.R-project.org/package=emmeans>.
- Meisel, J. M., Elsig, M., & Rinke, E. (2013). *Language acquisition and change: A morphosyntactic perspective*. Edinburgh University Press.
- Montrul, S. (2004). *The acquisition of Spanish*. John Benjamins.
- 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>
- Moreno Cabrera, J. C. (1987). *Fundamentos de sintaxis general*. Síntesis.
- Müller, N., & Hulk, A. (2001). Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition*, 4(1), 1–21. <https://doi.org/10.1017/S1366728901000116>
- Oller, D. K., & Eilers, R. E. (Eds.). (2002). *Language and literacy in bilingual children*. Multilingual Matters Ltd.
- Paradis, J. (2010). Bilingual children's acquisition of English verb morphology: Effects of language exposure, structure complexity, and task type. *Language Learning*, 60(3), 651–680. <https://doi.org/10.1111/j.1467-9922.2010.00567.x>
- Pérez-Leroux, A. T., Pirvulescu, M., & Roberge, Y. (2009). Bilingualism as a window into the language faculty: The acquisition of objects in French-speaking children in bilingual and monolingual contexts. *Bilingualism: Language and Cognition*, 12(1), 97–112. <https://doi.org/10.1017/S136672890800391X>
- Pérez-Leroux, A. T., Pirvulescu, M., & Roberge, Y. (2018). *Direct objects and language acquisition*. Cambridge University Press.
- 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., & Soto-Corominas, A. (2021). Indirect structural crosslinguistic influence in early Catalan–Spanish bilinguals in adulthood: Predicate selection in Catalan existential constructions. *Applied Psycholinguistics*, 42(6), 1463–1502. <https://doi.org/10.1017/S0142716421000308>
- Pirvulescu, M., Pérez-Leroux, A. T., Roberge, Y., Strik, N., & Thomas, D. (2014). Bilingual effects: Exploring object omission in pronominal languages. *Bilingualism: Language and Cognition*, 17(3), 495–510. <https://doi.org/10.1017/S1366728913000631>

- Prévost, P. (2009). *The acquisition of French*. John Benjamins.
- Raposo, E. (1986). On the null object construction in European Portuguese. In O. Jaeggli & C. Silva-Corvalán (Eds.), *Studies in Romance linguistics*, 373–390. Foris.
- R Core Team. (2024). *R: A language and environment for statistical computing*. [Computer software]. R Foundation for Statistical Computing. <http://www.R-project.org/>
- Roca, F. (1992). Object clitics in Spanish and Catalan. In *Catalan Working Papers in Linguistics*, 245–280.
- Romaine, S. (1989). *Bilingualism*. Oxford University Press.
- Schlyter, S. (1994). The weaker language in bilingual Swedish–French children. In A. Viberg & K. Hyltenstam (Eds.), *Progression and regression in language: Sociocultural, neuropsychological and linguistic perspectives* (pp. 289–308). Cambridge University Press.
- Sorace, A. (2003). Near-nativeness. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition*. Blackwell Pub.
- Soto-Corominas, A. (2018). Acquisition of Catalan and Spanish morphosyntax in the Catalan–Spanish bilingual context. [Doctoral dissertation, University of Western Ontario].
- Soto-Corominas, A. (2021). Acquisition of quantified partitivity in Catalan–Spanish bilingualism. Influence from child-level and language-level factors. *Linguistic Approaches to Bilingualism*, 11(6), 873–908. <https://doi.org/10.1075/lab.19016.sot>
- Tarrés, M., & Bel, A. (2015). L’adquisició del català en infants portuguesos: Un estudi sobre l’ús i el coneixement dels pronoms clítics. *Zeitschrift Für Katalanistik*, 28, 261–281. <https://doi.org/10.46586/ZfK.2015.261-281>
- 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. <https://doi.org/10.1017/S1366728913000497>
- 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>
- Tsimpli, I. M. (2014). Early, late or very late?: Timing acquisition and bilingualism. *Linguistic Approaches to Bilingualism*, 4(3), 283–313. <https://doi.org/10.1075/lab.4.3.01tsi>
- Van Dijk, C., Van Wonderen, E., Koutamanis, E., Kootstra, G. J., Dijkstra, T., & Unsworth, S. (2022). Cross-linguistic influence in simultaneous and early sequential bilingual children: A meta-analysis. *Journal of Child Language*, 49(5), 897–929. <https://doi.org/10.1017/S0305000921000337>
- Varlokosta, S., Belletti, A., Costa, J., Friedmann, N., Gavarró, A., Grohmann, K. K., Gustai, M. T., Tuller, L., Lobo, M., Andelkovic, D., Argemi, N., Avram, L., Berends, S., Brunetto, V., Delage, H., Ezeizabarrena, M., Fattal, I., Haman, E., van Hout, A., de Lopez, K. J. ... & Yatsushiro, K. (2016). A cross-linguistic study of the

- acquisition of clitic and pronoun production. *Language Acquisition*, 23(1), 1–26.
<https://doi.org/10.1080/10489223.2015.1028628>
- Wexler, K., Gavarró, A., & Torrens, V. (2004). Feature checking and object clitic omission in child Catalan and Spanish. In R. Bok-Bennema, B. Hollebrandse, B. Kampers-Manhe, & P. Sleeman (Eds.), *Romance languages and linguistic theory* (pp. 253–269). John Benjamins. <https://doi.org/10.1075/cilt.256.15wex>
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer.
<https://doi.org/10.1007/978-0-387-98141-3>
- Yang, C. D. (2002). *Knowledge and learning in natural language*. Oxford University Press.
- Yip, V., & Matthews, S. (2000). Syntactic transfer in a Cantonese–English bilingual child. *Bilingualism: Language and Cognition*, 3(3), 193–208.
<https://doi.org/10.1017/S136672890000033X>

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Accessible Summary