




Structured variation, language experience, and crosslinguistic influence shape child heritage speakers' Spanish direct objects

Naomi Shin^a , Alejandro Cuza^b  and Liliana Sánchez^c ^aThe University of New Mexico, Albuquerque, New Mexico, USA; ^bPurdue University, West Lafayette, Indiana, USA and ^cUniversity of Illinois Chicago, Chicago, Illinois, USA

Research Article

Cite this article: Shin N, Cuza A, Sánchez L (2023). Structured variation, language experience, and crosslinguistic influence shape child heritage speakers' Spanish direct objects. *Bilingualism: Language and Cognition* 26, 317–329. <https://doi.org/10.1017/S1366728922000694>

Received: 28 October 2021
Revised: 13 September 2022
Accepted: 18 September 2022
First published online: 17 November 2022

Keywords:

child heritage speakers; direct object expression; direct object clitics; structured variation; animacy effects in morphosyntax; language experience; children's direct objects; Spanish direct objects

Address for correspondence:

Naomi Shin, Dept. of Linguistics, MSC03 2130, 1 University of New Mexico, Albuquerque, NM, 81731. USA. E-mail: naomishin@unm.edu

Abstract

This study investigates child heritage speakers' Spanish direct objects. A task designed to elicit direct objects was completed in Spanish and English by 40 child heritage speakers of Spanish in the U.S., and in Spanish by 24 monolingual children in Mexico. Both participant groups varied their direct object forms, following the same ranking: clitics>lexical NPs>omission>doubling. Animate referents promoted clitics; inanimate referents promoted lexical NPs. Among the heritage speakers, more Spanish experience and Spanish lexical proficiency predicted more clitic use (less omission and lexical NP use). We also argue that the child heritage speakers' production of strong pronouns, more lexical NPs, and masculine clitic *lo* with inanimate feminine referents suggest English influence. The study underscores the importance of examining structured variation, which revealed both similarities and differences between heritage and monolingual speakers.

1. Introduction

The burgeoning field of child heritage language acquisition has made major strides towards understanding how language development unfolds among bilingual children learning a minority language at home. Scholars have emphasized minority language input and crosslinguistic influence from the majority language as factors that shape heritage speakers' developing grammars (Cuza, 2016; Cuza & Pérez-Tattam, 2016; Flores, Santos, Jesus & Marques, 2017; Jia & Paradis, 2015; Montrul & Sánchez-Walker, 2013; Putnam & Sánchez, 2013). Yet, there remains a pressing need to better understand variation among speakers due to different language experiences as well as language-internal STRUCTURED VARIATION. Structured variation is defined as the interchange of linguistic forms where the choice to use one form over the other is probabilistically conditioned by linguistic and social factors (Labov, 1994). Traditionally, child language research has focused on more categorical aspects of language; however, the growing body of research on acquisition of structured variation suggests that monolingual and bilingual children alike learn probabilistic patterns by attending to these patterns in the input (Shin & Miller, 2022).

Investigating structured variation helps us move beyond an approach that sometimes inadvertently deems heritage grammars as deficient. As Flores and Rinke (2020, p. 25) write, “variation in heritage language grammars cannot be equated with deviance.” Indeed, researchers may characterize heritage speakers' variable use of forms as inaccurate (as compared to a baseline) only to find systematicity upon further inspection. Investigating structured variation may also reveal similarities between heritage and other speakers that otherwise go unnoticed. Polinsky and Scontras (2020) note that research has focused on how heritage speakers differ from other speakers. They write: “Most of heritage research looks at areas of vulnerability and deviation from the baseline. ... Without much-needed systematic exploration, most findings on resilient domains are accidental discoveries, which means that the empirical picture is likely fragmented” (p. 8). Requena (2022) provides a poignant example. Ticio (2015) had found that young Spanish–English bilingual children omitted Spanish differential object marker *a* more often than monolingual children. Requena reanalyzed Ticio's data, but took into account language-internal structured variation related to animacy. His re-analysis showed that the bilingual children, including U.S. child heritage speakers, expressed *a* as often as age-matched monolingual children.

In addition to examining structured variation, the current study investigates the impact of language experience and crosslinguistic influence on child heritage speakers' developing grammars. Previous research indicates that amount of heritage language experience predicts morphosyntactic development, with restricted input prolonging development (Flores et al., 2017; Gathercole & Thomas, 2009; Rodina & Westergaard, 2017; Shin, Rodríguez, Armijo & Perera-Lunde, 2019; Silva-Corvalán, 2014; Thordardottir, 2015; Unsworth, 2019). Crosslinguistic influence can manifest in various ways. Bilingual children sometimes transfer structures from one language into the other,

© The Author(s), 2022. 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.

such as preposition stranding in Spanish spoken by English–Spanish bilingual children (Silva-Corvalán, 2014). Crosslinguistic influence may also manifest in more subtle ways: for example, when two grammatical options are available in the heritage language (e.g., subject pronoun expression/omission in Spanish), one option may be reinforced by the child's other language (e.g., pronoun expression in English – Silva-Corvalán, 2014).

Taking into consideration the importance of structured variation as well as the impact of language experience and crosslinguistic influence on child heritage speakers' developing grammars, the current study investigates child heritage speakers' Spanish direct objects. A task designed to elicit direct objects was completed in Spanish and English by 40 child heritage speakers of Spanish in the U.S., and in Spanish by 24 monolingual children in Mexico. Structured variation is investigated by focusing on animacy. We hypothesize that the children will produce a variety of direct object forms and that this variation will be constrained by animacy. As discussed in Section 2, animacy plays a role in the forms that direct objects take in Spanish and English. In addition, given that animate entities are more accessible and salient than inanimate entities (Branigan, Pickering & Tanaka, 2008), they are likelier to be pronominalized (Ariel, 1990; Gundel, Hedberg & Zacharski, 1993). As such, we predict that animate direct object referents will be pronominalized more often than inanimate ones.

Further analyses test the hypothesis that child heritage speakers' direct object instantiation in Spanish will differ depending on amount of Spanish language experience. We predict that more Spanish experience results in higher rates of direct object clitic expression. In addition, since previous research has shown that lexical proficiency and age correlate with increased direct object expression in bilingual children (Pirvulescu, Pérez-Leroux, Roberge, Strik & Thomas, 2014; Shin, 2022), we predict that higher levels of Spanish lexical proficiency will result in higher rates of direct object expression, and that older children will express direct objects more often than younger children. Finally, we investigate possible crosslinguistic influence from English by analyzing direct object types and clitic gender. We argue that the following phenomena suggest English influence: the production of Spanish direct object strong pronouns, an increase in Spanish direct object lexical noun phrases (NPs), and the use of masculine clitic *lo* to refer to inanimate feminine referents.

Our approach aims to understand the linguistic systematicities of child heritage grammars in their own right, rather than assuming, a priori, a 'correct response' and deeming other responses erroneous. In this approach, comparisons with monolingual speakers remain useful for understanding which linguistic features are shared among all speakers and which are unique to child heritage speakers, thereby paving the way for an explanation for both types of features.

2. Direct objects in English and Spanish

English lexical and pronominal direct objects occur in postverbal position (1). Spanish direct object lexical NPs generally occur after the verb (2a), whereas Spanish direct object clitics typically occur before the verb (2b). Clitics can also occur after the verb in periphrastic constructions (2c), and must occur after the verb in imperative constructions (2d).¹ Direct object pronominalization

¹Clitics also appear in clitic left dislocations and clitic right dislocations (Zagona, 2000).

manifests as strong pronouns in English, but generally as clitic pronouns in Spanish (Pescarini, 2021). Spanish strong pronouns like *ella/él* 'she/he' are sometimes used as direct objects when a clitic is also present, i.e., in doubling constructions (2e) (Ormazabal & Romero, 2013; Zdrojewski & Sánchez, 2014). Bautista-Maldonado and Montrul (2019) found that Spanish-speaking adults in Mexico sometimes produced constructions with strong personal pronouns without the accompanying clitic (2f); however the pronoun was always preceded by differential object marker *a*.²

(1) The girl draws **the flower/it**.

(2a) La niña dibuja **la flor**.
The-F girl-F draws ACC.F.3SG flower
'The girl draws the flower'.

(2b) La niña **la** dibuja.
The-F girl-F it.ACC.F.3SG draws.
'The girl draws it'.

(2c) La niña está dibujando = **la**.
The-F girl-F is drawing = ACC.F.3.SG
'The girl is drawing it'.

(2d) ¡Dibújate = **la!**
Draw = ACC.F.3.SG
'Draw it!'

(2e) **La** vimos a **ella**.
ACC.F.3SG saw DOM her.F.3.SG
'We saw her'

(2f) Vimos a **ella**.
Saw DOM her.F.3.SG
'We saw her'

In Spanish, direct object clitics, articles, and adjectives agree in gender with the noun they modify or reference. Spanish direct object clitics are marked for gender for animate (2e–f) and inanimate (2b–d) referents.³ *La* is feminine (2b–e); *lo* is masculine. In English, singular pronouns are marked for gender for animate (2e–f), but not inanimate referents (2b–d). While many varieties of Spanish distinguish between direct and indirect object clitics (*lo/la* versus *le*), *le* is sometimes used for animate direct objects (Gómez Seibane, 2013; Ormazabal & Romero, 2013; Schwenter, 2006) and, though rarer, sometimes for animate and inanimate direct objects (Mayer & Sánchez, 2017). Given the importance of animacy and its contribution to direct object clitic forms in Spanish and pronominal forms in English, its impact on direct object usage is explored in this study.

2.1 Monolingual adults' and children's direct object types

Much of the literature on children's direct objects has focused on omission (3), which is a common cross-linguistic phenomenon that dissipates with age and lexical development (Pérez-Leroux, Pirvulescu & Roberge, 2018).

(3) What is Maria doing with the window? She's **opening**.

²Callen and Miller (2022) found that monolingual children in Mexico produced some post-verbal direct object pronouns without differential object marker *a*; however, these were demonstratives or non-specific pronouns like *uno* 'one', not personal pronouns (Callen, personal communication, October 20, 2021).

³In examples 2b–d, we are assuming clitic *la* refers to *la flor*, which is inanimate.

With respect to direct object types, both adults and young children modify their usage according to referent accessibility. Less accessible referents prompt more informative forms like lexical NPs, whereas more accessible referents prompt less informative forms like pronouns (Allen, Skarabela & Hughes, 2008). Even when the discourse context is held constant, however, speakers vary between lexical and pronominal forms. Pérez-Leroux, Pirvulescu, and Roberge (2008) investigated English- and French-speaking adults' and 2–5-year-olds' responses to prompts like (4).

- (4) Participant sees a picture of a girl drawing a flower and hears "Please tell Croco what the little girl is doing with the flower."
 Expected response with pronoun: "She's drawing **it**."
 Lexical NP response: "She's drawing **the flower**."

The English-speaking adults produced object pronouns and lexical NPs (64%, 36%, respectively). French-speaking adults' responses were mostly direct object clitics, but included some lexical NPs (82%, 18%, respectively). The French-speaking 4–5-year-olds produced more clitics than lexical NPs, whereas the English-speaking 4–5-year-olds produced more lexical NPs than pronouns.

Turning to Spanish, in Castilla and Pérez-Leroux (2010), 103 3–5-year-old children and 10 adults in Colombia were asked questions like *¿Qué le hace la mamá a la niña?* 'What is the mom doing to the girl?' The adults' responses consisted primarily of direct object clitics (96%), as in *la peina* 'she's combing her'; 4% were object omissions. The children's clitic production increased with age. Lexical NP direct objects were rare among the 3- and 4-year-olds and nonexistent among the 5-year-olds and adults.

In the studies discussed above, children omitted direct objects, but adults rarely did so. English-speaking adults produced direct object pronouns and lexical NPs, whereas Spanish-speaking adults always produced clitics, and French-speaking adults mostly did, too. English- and French-speaking children produced lexical NP direct objects, although these were more frequent among the former than the latter. Spanish-speaking children rarely produced lexical NPs. Together, these studies suggest that repeating lexical NPs across objects is more common in English than in Spanish or French.

There is some evidence that animacy impacts direct object forms. In a study of Polish- and Ukrainian-speaking adults and children, Mykhaylyk and Sopata (2015) found that the adults pronominalized animate direct objects more often than inanimate ones, and, conversely, their rates of lexical NPs were higher for inanimate direct objects than for animate ones. Animacy affected the children's direct object types at age five, at which point they pronominalized animate referents more often than inanimate ones. These animacy effects are likely related to accessibility and saliency. Animate entities are more accessible and salient than inanimate entities (Branigan et al., 2008), which makes them likelier to be pronominalized (Ariel, 1990; Gundel et al., 1993).

2.2 Spanish–English bilingual children's direct object types

Previous research has demonstrated that bilingual children omit direct objects more than monolingual children do (Castilla-Earls, Restrepo, Pérez-Leroux, Gray, Holmes, Gail & Chen, 2016; Castilla-Earls, Pérez-Leroux, Martínez-Nieto, Restrepo & Barr, 2020; Pirvulescu et al., 2014; Sánchez, 2003). This bilingual effect has been explained as the result of restricted input and lexical development (Pérez-Leroux et al., 2018;

Pirvulescu et al., 2014; Shin, 2022). The higher the child's receptive vocabulary score in their heritage language, the more direct objects they express in that same language (Shin, 2022; Shin et al., 2019). With respect to direct object types, in Castilla-Earls et al. (2016), 16 U.S. Spanish–English bilingual children, ages 4;7–8;4, either omitted objects or produced direct object clitics, but they produced almost no direct object lexical NPs. This study did not, however, control for animacy.

2.3 Direct object clitic gender in Spanish

Whereas school-age monolingual Spanish-speaking children generally match the gender of the direct object clitic to the gender of the noun the clitic references, as in (5a), U.S. child heritage speakers sometimes produce mismatches, as in (5b) (Castilla-Earls et al., 2020; Martínez-Nieto & Restrepo, 2022).

- (5) Prompt: *¿Qué hizo Susana con la comida?*
 What did Susana do with the food?
- a) Gender match: **La** congeló.
 ACC.F.3.SG froze
 'She froze **it**'
- b) Gender mismatch: **Lo** congeló
 ACC.M.3.SG froze
 'She froze **it**'

There is reason to suspect that English plays a role in Spanish gender mismatching. In a study of 37 U.S. child heritage speakers, Shin et al. (2019) found that more English experience and higher English vocabulary scores correlated with 'lo extension', i.e., increased use of masculine *lo* to refer to masculine and feminine referents alike. They found no relationship between Spanish vocabulary scores and *lo* extension. Goebel-Mahrle and Shin (2020) also found *lo* extension in a corpus study of Spanish spoken by Spanish–English bilingual children. Nevertheless, it is unclear whether *lo* extension applies to both animate and inanimate referents. Given the importance of animacy in the Spanish clitic system, and the fact that English only encodes gender in pronouns with animate referents, animacy may constrain *lo* extension.

3. Research questions and predictions

This study focuses on three research questions:

- 1) Structured variation: What direct object forms do child heritage speakers use in Spanish and how does animacy condition the use of those forms?
- 2) Spanish language experience: Does child heritage speakers' Spanish direct object instantiation differ according to amount of Spanish language experience, Spanish lexical proficiency, and age?
- 3) Crosslinguistic influence: Does English influence child heritage speakers' Spanish direct object instantiation and their direct object clitic gender?

With respect to structured variation (RQ1), we hypothesize that child heritage speakers will produce a variety of direct object forms, and this variation will be constrained by animacy. More specifically, we predict that, like monolinguals, child heritage speakers will produce mostly direct object clitics, but also lexical NPs (P1a). Further, since animate entities are more accessible and salient, and thus likelier to be pronominalized, we predict

that animate referents will promote the use of direct object clitics, whereas inanimate referents will promote the use of lexical NPs (P1b).

With respect to language experience (RQ2), we hypothesize that child heritage speakers' Spanish direct object types will differ depending on Spanish language experience, Spanish lexical proficiency, and age. We anticipate that increased Spanish experience and Spanish lexical proficiency will lead to increased rates of clitics and, concomitantly, lower rates of lexical NPs and omission (P2a). Since direct object omission has been shown to decrease with age, we anticipate that the younger heritage speakers will omit more objects than older ones (P2b).

With respect to crosslinguistic influence (RQ3), we hypothesize that child heritage speakers' direct objects will be influenced by English. English influence may manifest as strong pronouns used as direct objects in Spanish, even in non-doubled constructions and without differential object marker *a* (P3a). English influence may also manifest as an increase in post-verbal direct object lexical NPs in Spanish (P3b). With respect to clitic gender, we anticipate *lo* extension (*lo* referring to feminine referents) (P3c). Given that English marks gender in animate singular direct object pronouns ('him/her'), we predict that English influence will manifest as an extension of *lo* to inanimate feminine referents in particular (P3d).

4. Methods

4.1 Participants

Participants included 40 U.S. child heritage speakers of Spanish (ages 5;3-11;9, $M = 8;9$), 37 of whom were born in the U.S., 2 in Mexico, and 1 in Puerto Rico.⁴ Of the 40 children, 37 were exposed to Mexican Spanish at home and 1 to Puerto Rican Spanish. For the remaining two children, the variety of Spanish spoken at home was not reported. The study also included 24 monolingual children in Mexico (ages 5;5-11;0, $M = 8;2$).

4.2 Materials and procedure

Following similar methodology in previous studies (Cuza, Pérez-Leroux & Sánchez, 2013; Shin, in 2022), all participants completed a 24-item task in which the experimenter (the second author) described a picture and asked a question designed to elicit a direct object clitic. For example, participants saw a picture of children following a woman and heard *Los niños están con su tía en el parque. ¿Qué le están haciendo los niños a la tía?* 'The children are with their aunt in the park. What are the children doing to the aunt?' The experimenter provided the verb to be used by the participant, in this case *seguir* 'follow'. The expected response was *la están siguiendo* or *están siguiéndola* 'they are following her' (see Supplementary Materials, Appendix A, Figure S1 and Figure S2).

Different from previous studies of Spanish, the items were balanced for animacy: half the prompts included animate referents in the prepositional phrase, as in *a la tía* 'to the aunt', while the other half included inanimate referents. The items were also balanced for gender. Half the animate items included

⁴The children born in Mexico arrived in the U.S. at around 1 and 2 years old. The child born in Puerto Rico (CB15) arrived at around age 7 and had been in the U.S. for 1 year and 3 months at the time of testing. However, she spoke English fluently and often, had a high English vocabulary score (94), and her responses to the experimental task included 5 clitics, 15 lexical NPs, 3 omissions, and 1 strong pronoun, placing her squarely within the heritage speaker group.

Table 1. Direct object types produced by monolingual and child heritage speakers

Direct object type	Monolingual	Heritage
	N (%)	N (%)
Clitics in non-doubled constructions	436 (77.9)	453 (49.0)
Lexical NPs	73 (13.0)	296 (32.0)
Omission	43 (7.7)	134 (14.5)
Doubling constructions (clitic + lexical NP)	8 (1.4)	15 (1.6)
Demonstratives	—	2 (.2)
Strong pronouns	—	25 (2.7)
Total	560 (100)	925 (100)

feminine referents, e.g., *la tía*; the other half included masculine referents, e.g., *su hijo* 'her son'. So too, half the inanimate items were feminine, e.g., *la película*, and half were masculine, e.g., *el teléfono*. Thus, there were six animate-feminine items, six animate-masculine items, six inanimate-feminine items, and six inanimate-masculine items. All inanimate noun referents had canonical gender; they were either masculine and ended in *-o*, or feminine and ended in *-a*. Animate referents either included common names or words that followed canonical gender.

4.3 Data compilation and coding

All participant responses were transcribed by a research assistant and then checked by the second author. To focus on direct objects, only responses with transitive verbs were included. Responses with intransitive verbs like *caminando* 'walking' were excluded. Also excluded were cases in which the children did not respond at all or produced the verb in English (waking *la hija* 'waking the daughter').⁵ The process of isolating transitive constructions yielded 1,485 responses in Spanish. The child heritage speakers also completed the task in English. Exclusions from the English version included constructions with copula verbs, as in 'he is happy', and intransitive verbs, such as 'walking', leaving a total of 926 English responses.

Spanish responses were coded for the following:

1. Direct object type: clitic; lexical NP; omission; doubling construction with both a clitic and a lexical NP, as in *la asustó a su mamá* 'he scared his mom'; strong (personal) pronoun *él/ella* ('he/she'), or demonstrative.⁶
2. Animacy: Animate or inanimate depending on the direct object referent in the participant's response.
3. Noun gender: Masculine or feminine depending on the gender of the direct object referent in the participant's response.
4. Gender match vs. mismatch: Direct object clitics *lo/la* were coded as either matching the gender of the noun referent (5a), or not matching ("mismatch") (5b).

English responses were coded for Direct object type (pronoun, lexical NP, omission) and Animacy. Coding was completed by

⁵Three periphrastic verbs that included elements from Spanish and English were retained: *está blowando la sopa* 'is blowing the soup', *throwando la pelota* 'throwing the ball', and *lo está hugging* 'he is hugging him'.

⁶Three cases of *algo* 'something' were grouped with lexical NPs.

Table 2. Mixed effects multinomial regression analysis predicting Spanish direct objects, Monolingual and Heritage children. Random factor = Individual child

	N Tokens	Lexical NPs				Omissions			
		% Lex. NPs	β	SE	p	% omission	β	SE	p
Monolingual	552	13.2				7.8			
Heritage	883	33.5	1.44	.40	<.0001	15.2	1.14	.46	.01
Animate	674	17.1	—	—	—	12.3	—	—	—
Inanimate	761	33.4	.89	.14	<.0001	12.4	.24	.17	.15

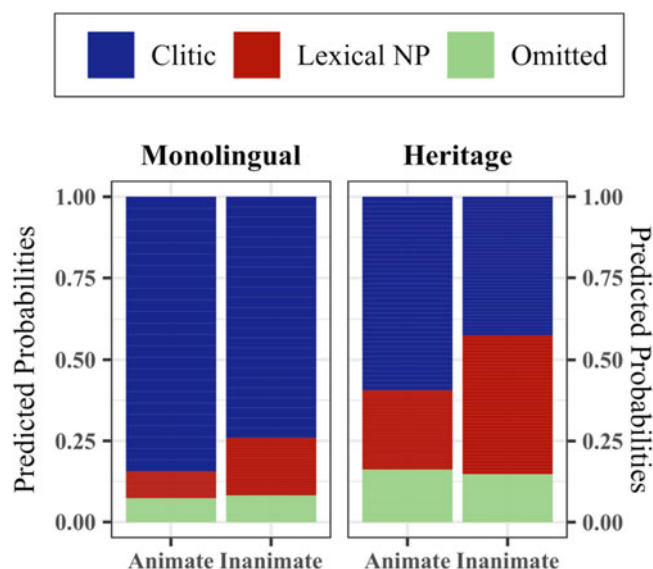


Figure 1. Predicted probabilities of direct object types by animacy, Monolingual and heritage children

research assistants, checked by the first author, and corroborated, where necessary, by the second and third authors.

The results are presented in two parts. Section 5 presents analyses investigating direct object types. Focusing on the three most common Spanish direct object types (Clitic, Lexical NP, Omission), Section 5.1 presents results from a mixed-effects multinomial logistic regression investigating the impact of Participant group (Monolingual, Heritage), Age, and Animacy on Direct object type. Section 5.2 presents results from a mixed-effects multinomial logistic regression investigating the impact of Animacy, Age, Language experience, and Lexical proficiency on child heritage speakers' Spanish direct object types. Section 5.3 presents the child heritage speakers' English responses, and compares rates of direct object types across the children's languages. Section 6 focuses on Spanish clitic gender, and presents results from a mixed-effects binary logistic regression and a conditional inference tree investigating the impact of Animacy, Noun gender, English experience, and English lexical proficiency on child heritage speakers' gender matching.

5. Results: Direct object types in Spanish

Table 1 presents the distribution of direct object response types produced by the monolingual and heritage children. The most frequent response type was clitics in non-doubled constructions.

These were primarily *lo/la*, but also included 59 tokens of *le* referring to direct objects in non-doubled constructions. The next two most common response types were lexical NPs and direct object omission. Together, these three types (clitics, lexical NPs, omission) comprised 97% of the data. The children also produced some doubling constructions with both a clitic and a lexical NP; they did not produce any doubling constructions with strong pronouns, as in (2e). Overall, monolingual and heritage speakers were similar in terms of the overall ranking of direct object types common to both groups: clitics>NPs>omission>doubling.

While both monolingual and heritage speakers produced direct object clitics, lexical NPs, and omissions, the latter were unique in that they also produced demonstratives, as in *afeitando ese* 'shaving that one', and strong (personal) pronouns, as in (6), produced by a 7;3-year-old child of Mexican descent (CB10). This use of strong pronouns in non-doubled constructions is unusual among monolingual Spanish speakers. Moreover, when monolinguals do produce direct object strong pronouns, these are preceded by differential object marker *a* (2e-f; Bautista-Maldonado & Montrul, 2019). In contrast, only 3 of the 25 direct object strong pronouns produced by the heritage speakers in our study were preceded by *a*.

- (6) Prompt: *¿Qué le está haciendo Carlota a su hija?*
 'What is Carlota doing to her daughter?'
 Response: *Cobijar ella*
 'Cover her'

5.1 Predicting direct object clitics, lexical NPs, and omission

Since 97% of the data was comprised of clitics, lexical NPs, and omissions, we isolated these response types for further analysis. A mixed-effects multinomial logistic regression was performed using the `mblogit` function in the `mclogit` package (Elff, 2021) in R (R Core Team, 2021) with Direct object type (Clitic, Lexical NP, Omitted) as the dependent variable, and Participant group (Monolingual, Heritage), Age in months, and Animacy (Animate, Inanimate) as predictor variables. The model included one random factor, Individual child. The model was set to predict lexical NPs and direct object omission (the baseline level was clitics). The baseline level for Participant group was Monolingual, and for Animacy it was Animate. Age in months was not significant, and thus was removed from the model.⁷

As shown in Table 2, inanimate reference significantly increased the likelihood of a lexical NP response, and heritage

⁷When the model only includes Age and the random factor Individual child, Age remains nonsignificant [Lexical NPs: $\beta = -.003$, SE: .01, $p = .77$; Omission: $\beta = -.009$, SE: .01, $p = .28$].

Table 3. Mixed effects multinomial regression analysis predicting lexical NP direct objects and direct object omission, Child heritage speakers only. Random factor = Individual child

	N Tokens	% Lex. NPs	Lexical NPs			Object Omission			
			β	SE	p	% omission	β	SE	P
Animate	370	26.5	—	—	—	17.3	—	—	—
Inanimate	444	42.8	.79	.18	< .0001	13.5	.01	.21	.95
Span vocabulary			-.05	.01	.0005		-.03	.01	.02
Eng vocabulary			.04	.02	.08		.04	.02	.03
Span experience			-.90	.40	.03		-.61	.35	.08
Eng experience			-.86	.35	.01		.04	.32	.91

children were significantly more likely to produce lexical NPs as compared to the monolingual children in Mexico. With respect to object omission, Animacy was not significant, but Participant group was: the heritage children were more likely to omit objects. Figure 1 plots the predicted probabilities of clitic, lexical NP, and direct object omission responses by Animacy for the monolingual and child heritage speakers.

5.2 The influence of age, language experience, and lexical proficiency on child heritage speakers' Spanish direct object types

To measure language experience, the child heritage speakers' parents completed a questionnaire in which they rated how often their children spoke Spanish to their relatives and friends (usage), as well as how often the children were spoken to in Spanish by their relatives and friends (exposure) (see Supplementary Materials, Appendix B). The same questions were asked about English to obtain a usage and exposure measure in each language. For each question, there were five possible answers; each was assigned a numerical value: *never* = 1, *almost never* = 2, *sometimes* = 3, *frequently* = 4, *very frequently* = 5. Average usage and exposure scores were calculated based on these values. Spanish usage and exposure scores were positively correlated [$r = .81, p < .0001$], as were English usage and exposure scores [$r = .76, p < .0001$]. Thus, we combined Spanish usage and exposure into one category labeled 'Spanish experience' and English usage and exposure into 'English experience'.

To measure lexical proficiency, the child heritage speakers completed standardized receptive vocabulary tests in Spanish (*Test de Vocabulario en Imágenes Peabody*, TVIP; Dunn, Lugo, Padilla & Dunn, 1986) and in English (Peabody Picture Vocabulary Test, PPVT; Dunn & Dunn, 2007). These tests are normed by age. The standardized mean is 100; scores between 85 and 115 are considered within the typical range.

Two children did not complete the Spanish vocabulary test, and one child's parents did not complete the questionnaire. These three children were excluded from analyses examining the impact of lexical proficiency and language experience on direct object types. The remaining 37 children's average Spanish experience score was 3.65, and their average English experience score was 3.20. Their standardized vocabulary scores ranged from 45–121 ($M = 83$) in Spanish from 78–124 ($M = 97$) in English.

A mixed-effects multinomial logistic regression was performed with Direct object type (Clitic, Lexical NP, Omission) as the

dependent variable, with Clitic as the baseline. The predictor variables included Animacy (Animate, Inanimate, with Animate as the baseline), Spanish experience, Spanish vocabulary, English experience, English vocabulary, and Age in months. These predictor variables are all continuous, except for Animacy, which is binary. Individual child was included as a random factor. Age in months was not significant, and thus was removed from the model.⁸ The results are presented in Table 3. Figure S3 in the Supplementary Materials (Appendix C) plots the predicted probabilities of Clitics, Lexical NPs, and Omission based on the model in Table 3. Each graph represents the effects of one continuous predictor variable at a time, showing effects for animate (red line) and inanimate (aqua line) referents, while holding the other predictor variables constant.

Table 3 demonstrates that inanimate reference increased the likelihood of Lexical NPs. This effect is depicted in Figure S3 in the Supplementary Materials: the probability of a lexical NP response is consistently higher for inanimates (aqua line) than animates (red line). The opposite holds for clitics: the probability of these responses is higher for animates than for inanimates. Spanish vocabulary and Spanish experience patterned similarly. The higher the Spanish vocabulary score, the less likely the response was a Lexical NP or direct object omission (the likelier the response was a clitic) (Figure S3a). Similarly, the higher the Spanish experience score, the less likely the response was a lexical NP (the likelier the response was a clitic) (Figure S3b). English vocabulary had the opposite effect: the higher the English vocabulary score, the more omissions and the fewer clitics (Figure S3c), although the impact of English vocabulary on lexical NP responses was not statistically significant. English experience patterned with Spanish vocabulary and Spanish experience: the more English experienced, the less likely the response was a lexical NP or direct object omission (Figure S3d), which is counterintuitive and contradicts the finding for English vocabulary. Setting aside English experience, the results in Table 3, illustrated by Figures S3a–c in the Supplementary Materials, are summarized as follows: among child heritage speakers, the more Spanish experience and the higher the Spanish vocabulary score, the more they produced clitics. The higher the English vocabulary score, the more they omitted direct objects.

⁸When only Age and the random factor Individual child are in the model, Age remains nonsignificant [Lexical NPs: $\beta = -.003$, SE: .01, $p = .85$; Omission: $\beta = -.02$, SE: .01, $p = .20$].

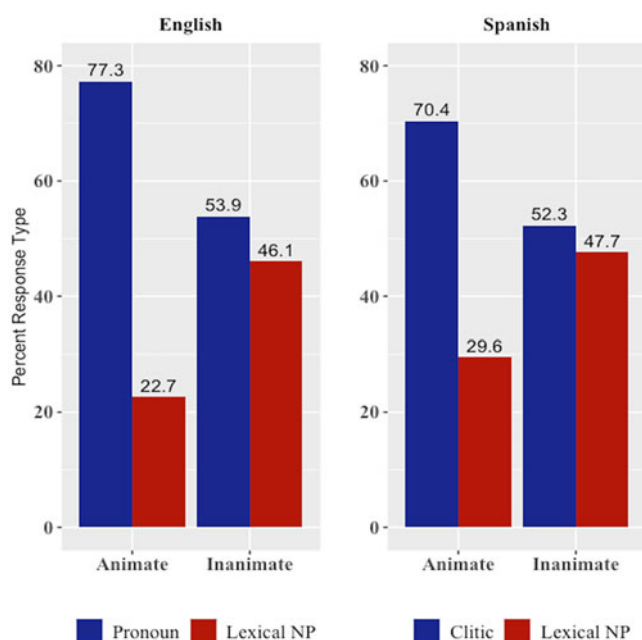


Figure 2. Child heritage speakers' expressed direct objects in English and Spanish by Animacy

5.3 English influence on Spanish

Several results reported above suggest English influence on the child heritage speakers' direct object instantiation in Spanish. Child heritage speakers produced strong pronouns as direct objects without an accompanying clitic and without differential object marker *a* (Example 6, P3a). This direct object type is unattested among the monolingual children in our study (Table 1). Child heritage speakers also produced more Spanish direct object lexical NPs as compared to monolingual children in Mexico (Table 1, Figure 1).

To further investigate English influence, we analyzed the children's English direct objects. Out of 926 responses to the elicited production task in English, 567 (61.2%) were postverbal pronouns, as in "buried it"; 317 (34.2%) were lexical NPs, as in "watching the movie", and 42 (4.5%) were omitted, as in "scraping" in response to "What is Helena doing with the cheese?" Figure S4 in the Supplementary Materials (Appendix C) plots the children's direct object types with animate and inanimate referents in both English and Spanish, side by side. The biggest difference between the children's English and Spanish is that they express significantly more direct objects in English and omit more in Spanish [$X^2(1) = 69.40, p < .0001$]. In English, their omission rates are 4% and 5% with animate and inanimate referents, respectively. In Spanish, their omission rates are 17.3% and 13.2% with animate and inanimate referents, respectively (Figure S4). Since the children rarely omit direct objects in English, their omission in Spanish is clearly not due to English influence.

Once we set aside object omission, the distribution of the children's direct object types is strikingly similar across the two languages (Figure 2). The similar distribution of pronominal and lexical forms across the two languages, coupled with the finding that the heritage speakers produce more lexical NPs than the monolinguals (Table 1, Figure 1), suggests that English influence on the children's Spanish manifests as an increase in direct object lexical NPs.

6. Results Part 2: Clitic Gender in Spanish

To investigate whether English influence manifests as increased reliance on masculine *lo* to refer to masculine and feminine referents alike (P3c), and that this *lo* extension is mediated by animacy (P3d), this section homes in on direct object clitics *lo/la* (Monolingual: 417, Heritage: 441⁹) and analyzes gender matching, where the clitic and the noun referent have the same gender (5a), versus gender mismatching, where the clitic gender differs from the noun referent (5b). The monolingual children rarely gender mismatched (3.6%). In contrast, the child heritage speakers gender mismatched at a significantly higher rate (23.4%) [$X^2(1) = 69.09, p < .0001$], thereby motivating further analysis. A mixed-effects binary logistic regression analysis was performed on the child heritage speakers' clitics using the `glmer()` function in the `lme4` package (Bates, Mächler, Bolker & Walker, 2015) in R, with Gender match vs. Gender mismatch as the dependent variable and the following predictor variables: Animacy (Animate, Inanimate, with Animate as the baseline), Noun gender (Masculine, Feminine, with Masculine as the baseline), English vocabulary (continuous, see section 5.2), and English experience (continuous, see section 5.2). Individual child was included as a random factor. The application value of the dependent variable was set to predict mismatches. Thus, positive coefficients indicate that a factor increases the likelihood of mismatching; negative coefficients indicate that a factor decreases the likelihood of mismatching. Inanimate referents and more English experience increased the likelihood of mismatches [Respectively: $\beta = 1.16, SE = .26, p < .0001$; $\beta = .50, SE = .25, p = .04$]. There was no main effect of Noun gender [Feminine: $\beta = .16, SE = .25, p = .53$], nor of English vocabulary [$\beta = -.02, SE = .01, p = .15$].¹⁰

To test the prediction that English experience results in the extension of *lo* to feminine referents (P3c) and, more specifically, to feminine inanimate referents (P3d), a mixed-effects binary logistic regression using the `glmer()` function was performed with Gender match vs. Gender mismatch as the dependent variable, Individual child as the random effect, and the three-way interaction term English experience*Noun gender*Animacy. This three-way interaction was not significant [$\beta = .22, SE = .83, p = .79$]. A model with the two-way interactions English experience*Noun gender, English experience*Animacy, and Animacy*Noun gender did not converge. Nevertheless, these three interactions were significant when the model included only one interaction at a time. The Animacy*Noun gender interaction [$\beta = 1.14, SE = .50, p = .02$] demonstrates the predicted *lo* extension effect: the children were likelier to mismatch when the referent was inanimate and feminine. Contra the predictions regarding English experience, more English experience decreased the likelihood of mismatching with feminine referents [$\beta = -.71, SE = .36, p = .048$] and with inanimate referents [$\beta = -1.16, SE = .41, p = .005$].

To further explore the interactions among English experience, Animacy, and Noun gender, a conditional inference tree was built using the `cree()` function in the `Party` package (Hothorn et al., 2006) in R. Conditional inference is a non-parametric method that recursively partitions the data to create maximally internally-homogenous subsets. The dependent variable was Gender match

⁹These counts include doubling with clitics *lo* or *la*. One token included both *lo* and *la*, and thus was excluded.

¹⁰There was no main effect of Noun gender or English vocabulary even in models with each as the only fixed effect [$\beta = .21, SE = .23, p = .36$; $\beta = .02, SE = .01, p = .22$].

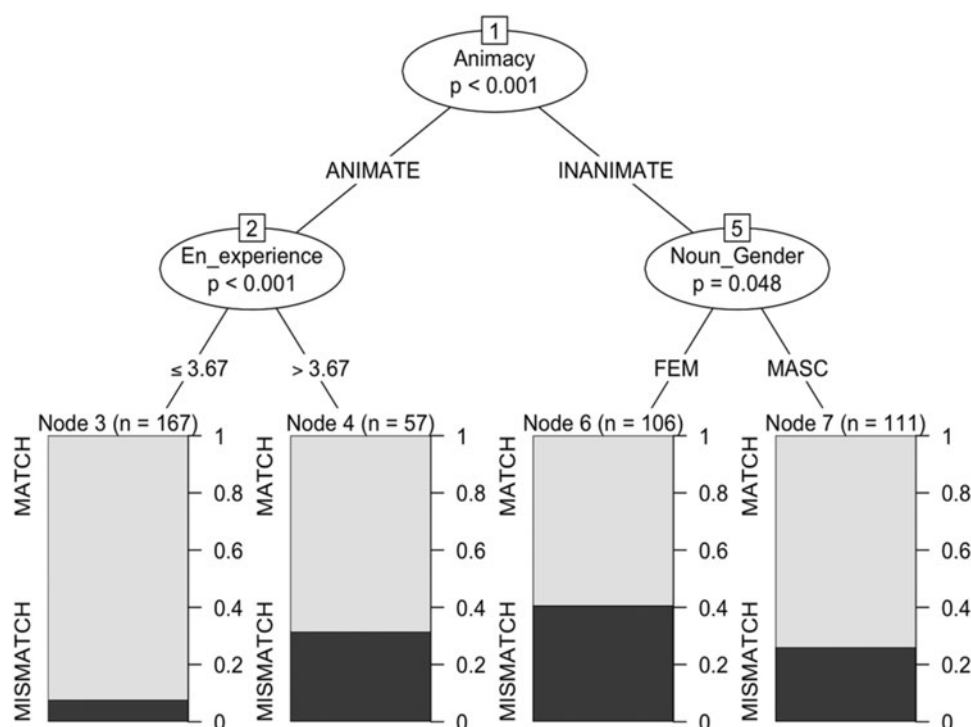


Figure 3. Conditional inference tree predicting child heritage speakers' Gender match vs. mismatch by Animacy, English experience, and Noun gender

versus Gender mismatch, and English experience, Animacy, and Noun gender were included as predictor variables.

Figure 3 demonstrates that the child heritage speakers gender mismatched most with inanimate referents (Node 1), confirming the importance of Animacy. Moreover, the children mismatched more with feminine inanimate referents (Node 6) than with masculine inanimate referents (Node 7), reflecting increased reliance on masculine clitic *lo*.

Turning to English experience, Figure 3 helps clarify the unexpected finding that the more English the children experienced, the more they gender mismatched with animate (rather than inanimate) referents (Node 2). The algorithm partitioned the data based on an English experience score of 3.67. Children whose score was higher than 3.67 mismatched with animate referents at a higher rate (Node 4) than those with lower scores (Node 3). At first glance, this contradicts Prediction 3d, which was that English would result in *lo* extension among inanimate, but not animate, referents. Nevertheless, a more complex story is revealed upon closer examination of the mismatches referring to animate referents ($n = 18$) produced by the ten children with scores over 3.67. Two children, both of Mexican descent, produced 61% of these mismatches. One (CB12) had a particularly interesting use of *la*. This child matched feminine clitic gender to the grammatical subject rather than the direct object referent. He applied this strategy with animate (7) and inanimate (8) direct objects.

(7) Prompt: ¿Qué le está haciendo Rosa a su hijo?
'What is Rosa doing to her son?'

Response: *La* está vistiendo.
ACC.F.3/SG is dressing.
'She is dressing her.'

(8) Prompt: ¿Qué hizo Ana con su teléfono?
'What did Ana do with her phone?'

Response: *La* dejó.
ACC.F.3/SG left.
'She left it.'

The second child (CB20) produced five tokens of *la* with animate masculine referents, but he also generally relied on *la*: 18/19 of this child's clitics were *la*. This tendency permeated the child's articles as well: he produced *la piano* and *la libro* 'the piano, the book' with feminine article *la* rather than masculine article *el*.

Although the conditional inference tree did not illustrate an interaction between English experience and Noun gender, the two unique children mentioned above also help explain the surprising regression result whereby more English experience decreased mismatching with feminine referents. Of the 39 mismatches produced by the ten children with English experience scores over 3.67, 22 were *la* with masculine referents, indicating *la* extension rather than *lo* extension. Nevertheless, 14 (64%) of those 22 tokens of *la* with masculine referents were produced by the two children discussed above. Thus, those two children increased the overall rate of mismatching with animate referents and with masculine referents among the group of children with abundant English experience.

Having identified the two children with unique strategies, the same mixed-effects binary logistic regression analysis was performed again, but this time the two unique children were excluded. As before, inanimate referents increased the likelihood of mismatches [$\beta = 1.62$, $SE = .30$, $p < .0001$] and so did feminine referents [$\beta = .74$, $SE = .27$, $p = .007$]. Animacy*Noun gender was again significant: inanimate feminine referents increased the likelihood of gender mismatching [$\beta = 1.19$, $SE = .58$, $p = .04$]. Different from before, however, with the two unique children excluded, neither English experience nor English vocabulary

predicted mismatching [respectively: $\beta = .31$, $SE = .30$, $p = .30$; $\beta = -.02$, $SE = .02$, $p = .17$].

To summarize, child heritage speakers gender mismatched significantly more often than monolingual children. Further, the child heritage speakers' data indicates 'lo extension', i.e., they mismatched more often with inanimate feminine referents. This suggests that *lo* may take on a gender-neutral function similar to English 'it' (P3c, P3d). The initial results at first glance contradicted the prediction that more English experience would promote more extension of *lo* to inanimate feminine referents. However, closer inspection of the children with abundant English experience revealed two children with unique strategies: one matched feminine clitic gender to the gender of the grammatical subject; the other relied on *la* in general. After excluding these two children, there was no significant impact of English experience. Thus, while the child heritage speaker group as a whole demonstrated *lo* extension to feminine inanimate referents, the study cannot confirm a relationship between *lo* extension and amount of English experience within the group of child heritage speakers.

7. Discussion

Our three research questions focused on how structured variation, language experience, and crosslinguistic influence impact child heritage speakers' direct objects. We discuss our findings within the context of the approach outlined in the introduction, highlighting not only differences, but also similarities between monolingual and heritage speakers. We underscore the importance of examining structured variation illustrated, in this case, by the impact of animacy on Spanish speakers' use of direct objects. We also argue that the child heritage speakers' direct object usage is shaped by their language experience and influence from English, and we propose that Sánchez's (2019) Bilingual Alignment model provides a promising way to capture crosslinguistic influence.

7.1 Structured variation: Direct objects and animacy effects

The first research question focused on what direct object forms child heritage speakers use and whether that usage is guided by animacy. Previous research had demonstrated that Spanish speakers mostly produce direct object clitics when referring to a noun mentioned in the previous clause (e.g., Castilla & Pérez-Leroux, 2010). In our study, clitics were the most frequent response type among both the monolingual and child heritage speakers (P1a). In fact, both participant groups showed the same ranking of direct object instantiation: clitics > NPs > omission > doubling. This ranking represents an important similarity between child heritage speakers and monolingual children.

We hypothesized that animacy would constrain the use of direct object forms. Indeed, animate reference promoted clitic use, whereas inanimate reference promoted lexical NP use (P1b; Table 2, Figure 1). This structured variation was evident among both participant groups, and thus represents a second similarity between child heritage speakers and monolinguals in Mexico. We interpret this as evidence that child heritage speakers, like all children, are sensitive to extralinguistic cues that probabilistically guide linguistic patterns – in this case, the tendency to pronominalize animate entities more than inanimate ones (Ariel, 1990; Branigan et al., 2008; Gundel et al., 1993; Mykhaylyk & Sopata, 2015).

The importance of animacy in our study concords with the observation that animacy shapes grammars across languages (Vihman & Nelson, 2019) and modulates forms related to direct object expression (Aissen, 2003; Mykhaylyk & Sopata, 2015; Ormazabal & Romero, 2013). There is evidence that children use animacy as a cue when learning linguistic patterns (Childers & Echols, 2004) and learners construct new grammatical patterns based on animacy. For example, Vihman, Nelson and Kirby (2018) presented 200 English-speaking adults with a miniature artificial language. The experiment included 20 transmission chains with 10 participants per chain. Each participant's output served as the stimuli for a subsequent participant, whose output became the stimuli for yet another participant. By the end of the chain, participants had introduced systematicity into the language: affixes were paired with nouns based on animacy, indicating that the participants relied on animacy as a cue for creating grammatical distinctions.

The animacy effect among the children in our study may be based on distributional patterns in the input. Previous research indicates that animacy affects adults' direct object types (Mykhaylyk & Sopata, 2015) and research on structured variation suggests that children attend to variable patterns in the input (Shin & Miller, 2022). It is also possible that learners have a predisposition to detect linguistic patterns based on animacy due to how salient a cue animacy is for pattern seeking in general (Childers & Echols, 2004).

7.2 Spanish language experience, lexical proficiency, and age

Our second research question focused on how Spanish language experience, lexical proficiency, and age shape child heritage speakers' direct object types. Previous research has demonstrated that amount of minority language experience impacts child heritage speakers' grammatical development (e.g., Flores et al., 2017; Rodina & Westergaard, 2017; Silva-Corvalán, 2014; Thordardottir, 2015; Unsworth, 2019). With respect to direct objects, previous studies indicate that omission is more common among children who experience restricted input, are younger, and have lower levels of lexical proficiency (Pérez-Leroux, Castilla & Brunner, 2012; Pirvulescu et al., 2014; Shin, 2022; Shin et al., 2019).

In our study, the higher the Spanish vocabulary score, the likelier the child heritage speakers were to produce clitics, and the less likely they were to produce lexical NPs or to omit objects (P2a; Table 3; Figure S3). Spanish experience also decreased the likelihood of using lexical NPs, which supports the conclusion that language experience shapes children's developing understanding of which forms to use in particular discourse contexts. Recognizing the important role of language experience has major ramifications for heritage language research. Given how varied their language experience is, we cannot treat heritage speakers as a monolithic group and instead must pay careful attention to how language experience shapes their grammatical development. There was no effect of age on object omission (contra P2b; Table 3). Thus, for the children in this study, whose mean age was 8;9, age was less of a determining factor of object omission as compared to language experience and lexical proficiency.

While language experience and lexical development affect children's acquisition of some grammatical structures, other structures appear to be less affected (Unsworth, 2019). Pérez-Leroux et al. (2012) argue that lexical development impacts Spanish-speaking children's acquisition of direct object clitics, but not articles. Shin et al. (2019) found that Spanish vocabulary scores correlated with

child heritage speakers' direct object expression, but not their rate of Spanish clitic gender matching. Although we did not set out to investigate the impact of Spanish experience and vocabulary on clitic gender, a follow-up analysis shows a significant impact of Spanish vocabulary. A binary logistic mixed-effects regression with Gender match vs. Gender mismatch as the dependent variable, and Animacy and Spanish vocabulary as the predictor variables demonstrates that the higher the vocabulary score, the less mismatching [$\beta = -.02$, $SE = .01$, $p = .02$]. Including Spanish experience instead of Spanish vocabulary in the model shows no effect of Spanish experience [$\beta = -.25$, $SE = .29$, $p = .39$]. These results suggest that, like direct object expression, acquisition of clitic gender matching may be related to lexical development. In a study of Spanish determiners and adjectives, adult heritage speakers gender matched more with higher frequency nouns (Hur, López Otero & Sánchez, 2020).

More research is needed to better understand which structures are most affected by lexical development and language experience. Grammatical patterns that are highly consistent and that have few lexically-based exceptions may be learned so rapidly that the window during which language experience and lexical development are relevant is short. Also, to better understand the role of language experience, we need to compare patterns in the input and the output (Daskalaki, Elma, Chondrogianni & Paradis, 2020). What exactly must children experience to learn the patterns of direct object instantiation? Do they need to experience specific verbs (Theakston, Lieven, Pine & Rowland, 2001), direct object pronouns (Pirvulescu et al., 2014), combinations of specific verbs and direct object types, or all of the above? We also need more research investigating what types of language experience matter. Experience with a variety of speakers appears to boost lexical development (Place & Hoff, 2016). Is the same true for grammatical development? Experiences in specific settings may also be relevant. Grammatical structures that are more common in school environments will naturally be less prominent among heritage speakers whose schooling is primarily conducted in the majority language (Otheguy, 2016). While we know that language experience matters, further research is needed to more precisely capture the types of experiences that matter for heritage language grammatical development.

7.3 English influence

With respect to our third research question, we hypothesized that English would influence child heritage speakers' Spanish direct objects. We predicted that child heritage speakers would produce post-verbal direct object strong pronouns in non-doubled constructions without the differential object marker *a* (P3a). There were 25 tokens of post-verbal strong personal pronouns (6), 22 of which occurred without *a*. The 25 tokens were produced by 11 different child heritage speakers; however, 12/25 were produced by one child (CB10). This child's complete set of responses ($n = 23$) included 12 strong pronouns, 9 lexical NPs, 1 demonstrative, 1 object omission, and no clitics. In other words, this child relied on strong pronouns to pronominalize direct objects, a strategy that, to the best of our knowledge, is unattested among monolingual Spanish speakers, and therefore suggests English influence.

Our second prediction was that English influence would manifest as an increase in post-verbal direct object lexical NPs in Spanish (P3b). This prediction was based on studies in which lexical NP direct objects were more frequent in English than in Spanish or French (Castilla & Pérez-Leroux, 2010; Pérez-Leroux

et al., 2008). As predicted, child heritage speakers produced more lexical NPs compared to the monolingual children in Mexico (Table 1, Figure 1). It is difficult to confirm the source of this quantitative difference given that both groups produced lexical NPs, and English vocabulary score and experience did not significantly increase the likelihood that child heritage speakers would produce lexical NPs (Table 3). At the same time, the child heritage speakers' lexical NP rates with animate and inanimate referents were strikingly similar across their two languages (Figure 2). We tentatively conclude that the child heritage speakers' higher rates of lexical NPs in Spanish are influenced by English.

What mechanism explains English influence on child heritage speakers' Spanish direct objects? Sánchez (2019) proposes that heritage speakers' usage patterns are the result of 'bilingual alignments', which consist of elements from different language components and which are used for comprehension and production processes, but do not necessarily result in fixed representations. She argues that these alignments arise so that bilinguals can access linguistic features for the purposes of communication even if their grammars do not have a fixed or stable representation for the features. The model offers a mechanism to account for crosslinguistic influence because bilingual alignments are permeable, and activation of features in one language can activate features in the other language (Kroll, Dussias, Bice & Perrotti, 2015). According to this model, child heritage speakers activate all direct object forms in both English and Spanish when using either language. Thus, they can access English features, which include strong pronouns, when speaking Spanish. Quantitative shifts, such as the higher rate of direct object lexical NPs in Spanish, can be explained by frequent activation of English lexical NPs, which strengthens this option in the child's memory and results in increased use of Spanish lexical NPs. The quantitative shifts could also be related to word order, which is prone to crosslinguistic influence (Hartsuiker, Pickering & Veltkamp, 2004). Activation of English direct objects is concomitant with activation of verb-object (VO) word order, which in turn could promote VO order in Spanish. Since clitics appear pre-verbally in most Spanish constructions, the activation of VO order could prompt the use of non-clitic forms, such as lexical NPs or strong pronouns. Regardless of whether the increased use of Spanish lexical NPs is due to activating English lexical NP forms or VO word order, Sánchez's bilingual alignment model provides a promising way to account for crosslinguistic phenomena among child heritage speakers.

While English appeared to impact the children's use of lexical NPs, it had no impact on object omission in Spanish. The child heritage speakers expressed objects more often in English than in Spanish (see Supplementary Materials, Appendix C, Figure S4). Thus, bilingual children omit more direct objects, but this 'bilingual effect' is not due to crosslinguistic influence. This conclusion concords with previous research showing that direct object expression in one language does not boost direct object expression in bilingual children's other language. Direct object expression is more frequent in child-directed English than in child-directed French (Pérez-Leroux et al., 2018). As such, English influence should yield increased object expression in bilinguals' French. Yet, bilingual French-English speaking children omit direct objects more often than French-speaking monolingual children (Pirvulescu et al., 2014). Learning to express direct objects appears to be related primarily to learning the transitivity frames for each verb and, as such, depends heavily on

lexical development in each language (Pérez-Leroux et al., 2012; Shin, 2022).

Our third and fourth predictions regarding English influence focused on clitic gender. Child heritage speakers produced higher rates of mismatches than monolingual children (23.4%, 3.6%, respectively). Results from analyses of the child heritage speakers' gender mismatches suggest English influence. Mismatching was more likely when referring to inanimate, feminine referents (P3c, P3d). This finding aligns with previous studies in which child heritage speakers produced *lo* for masculine and feminine referents alike (Goebel-Mahrle & Shin, 2020; Martínez-Nieto & Restrepo, 2022; Shin et al., 2019), and clarifies that this *lo* extension applies to inanimate referents in particular, which supports an English-influence explanation since English pronouns distinguish gender for animate but not inanimate singular referents. Frequent activation of English 'it' may result in increased activation of *lo*.

The clitic gender analyses revealed an interaction between animacy and English language experience: the more English the children experienced, the more they gender mismatched with animate (rather than inanimate) referents (Figure 3). At first glance, this appeared to contradict an English-influence explanation since English distinguishes between 'him' and 'her'. However, further analysis indicated that this surprising result was due to two unique children with interesting strategies. One child appeared to match clitic gender with the grammatical subject rather than the direct object. The second child mismatched with animate masculine referents because he relied on feminine *la*, regardless of the gender or animacy of the referent. Once these two children were removed from the data set, there was no effect for English experience. Overall, the clitic gender findings support English influence because, as a group, the child heritage speakers mismatched more with inanimate feminine direct objects (P3c, P3d). Nevertheless, there was no evidence that, within the group of child heritage speakers, more English experience or English vocabulary increased the likelihood of mismatching with inanimate feminine referents.

In sum, we argue that the child heritage speakers' use of strong pronouns as direct objects, increased rate of direct object lexical NPs, and clitic gender mismatching all suggest English influence. Nevertheless, future investigations that include other language pairs would help clarify the role of crosslinguistic influence. For example, if child heritage speakers of Spanish whose other language is French also show elevated rates of direct object lexical NPs and *lo* extension to inanimate feminine referents, we would not attribute these phenomena to French influence given that clitics are more prevalent than lexical NPs as direct objects in French (Pérez-Leroux et al., 2008), and French singular direct object clitics are marked for gender for both animate and inanimate referents. Thus, while our findings suggest that English plays a role in the child heritage speakers' use of Spanish direct object types and clitic gender, future research with other language pairs is needed to bolster this conclusion.

8. Conclusion

Our study responds to a call to increase research on language-internal variation in the field of heritage language acquisition (Flores & Rinke, 2020). Adopting an approach to the study of child heritage language acquisition that incorporates structured variation in addition to external factors such as language experience enabled us to uncover differences and similarities between

child heritage speakers in the U.S. and monolingual children in Mexico. As such, the study addresses the overamplification of differences in the field of heritage language acquisition (Polinsky & Scontras, 2020).

Our study investigated three research questions related to U.S. child heritage speakers' direct object instantiation in Spanish. The first was related to structured variation; we anticipated animacy would guide the use of direct object forms among child heritage speakers and child monolinguals in Mexico. Indeed, inanimate reference boosted the production of direct object lexical NPs for both participant groups, thereby revealing an important similarity between heritage and monolingual children. Our second question focused on Spanish language experience, lexical proficiency, and age. The more Spanish experience and the higher their Spanish vocabulary score, the more likely they were to produce clitics. Thus, language experience and lexical proficiency play a role in children's acquisition of direct object forms.

Our third question focused on English influence. We argued that the following phenomena reflect English influence on the child heritage speakers' Spanish: strong pronouns in non-doubled constructions, higher rates of lexical NPs in Spanish, and gender mismatching with inanimate feminine nouns. These results fit with Sánchez's (2019) bilingual alignment model of heritage language acquisition: bilinguals access features from both of their languages during language production, and crosslinguistic interaction may ensue. The analyses also uncovered interesting individual patterns, such as one child's tendency to match the gender of the clitic with the grammatical subject rather than the direct object referent. In summary, our study underscores the importance of analyzing language-internal structured variation, which revealed animacy effects among monolingual and heritage speakers alike, and presents new evidence showing the roles that language experience, lexical development, and crosslinguistic influence play in child heritage speakers' direct objects.

Competing interests. The authors declare none.

Data availability. The data that support the findings will be available in Open Science Framework at <https://osf.io/yr7ms/> following a 12 month embargo from the date of publication.

Supplementary Material. For supplementary material accompanying this paper, visit <https://doi.org/10.1017/S1366728922000694>

This article's online supplementary materials include:

1. Examples of question prompts and pictures used in the elicited production task (Appendix A, Figures S1 and S2)
2. The questions posed to caregivers to calculate Language Experience scores (Appendix B)
3. Figure S3 (Appendix C), which plots, for the child heritage speakers, the predicted probabilities of Clitic, Lexical NP, Omission responses based on multinomial model with four predictor variables: Spanish Vocabulary (A), Spanish Experience (B), English Vocabulary (C), English Experience (D). Each graph represents the effects of one continuous predictor variable at a time, showing effects for animate (red line) and inanimate (aqua line) reference, while holding the other predictor variables constant.
4. Figure S4 (Appendix C), which plots the child heritage speakers' rates of direct object types in English and Spanish by Animacy

Acknowledgments. The authors thank the children and their families for agreeing to participate in the current study. We also thank the College of Liberal Arts at Purdue University for providing funding to carry out this research (Exploratory Research in the Social Sciences Grant); partial

funding from Purdue University and the University of Illinois Chicago to support the Open Access fees; research assistants Ana María Molina, Jacob King, Mariana Marchesi, Rosa Prada Vivas, Krista Riegseker, and Laura Solano; as well as Leonardo Puertas Bustos and Nadia Neff for assistance with data analysis and visualization. Support for Neff was provided by ADVANCE at UNM (NSF EHR #1628471) through their Faculty Professional Development Support program.

References

- Aissen J (2003) Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21(3), 435–483.
- Allen S, Skarabela B and Hughes M (2008) Using corpora to examine discourse effects in syntax. In Behrens H (Ed.), *Trends in corpus research: Finding structure in data*. John Benjamins, pp 99–137.
- Ariel M (1990) *Accessing noun-phrase antecedents*. London: Routledge.
- Bates D, Mächler M, Bolker B and Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1), 1–48.
- Bautista-Maldonado S and Montrul S (2019) An experimental investigation of differential object marking in Mexican Spanish. *Spanish in Context* 16(1), 22–50.
- Branigan HP, Pickering MJ and Tanaka M (2008) Contributions of animacy to grammatical function assignment and word order during production. *Lingua* 118, 172–189.
- Callen MC and Miller K (2022) Linguistic variation in the acquisition of morphosyntax: Variable object marking in the speech of Mexican children and their caregivers. *Language Learning and Development* 18(3), 310–323. <https://doi.org/10.1080/15475441.2021.1977133>
- Castilla AP and Pérez-Leroux AT (2010) Omissions and substitutions in Spanish object clitics: Developmental optionality as a property of the representational system. *Language Acquisition* 17(1–2), 2–25.
- Castilla-Earls A, Restrepo MA, Pérez-Leroux AT, Gray S, Holmes P, Gail D and Chen Z (2016) Interactions between bilingual effects and language impairment: Exploring grammatical markers in Spanish-speaking bilingual children. *Applied Psycholinguistics* 37(5), 1147–1173.
- Castilla-Earls A, Pérez-Leroux AT, Martínez-Nieto L, Restrepo MA and Barr C (2020) Vulnerability of clitics and articles to bilingual effects in typically developing Spanish–English bilingual children. *Bilingualism: Language and Cognition* 23(4), 825–835.
- Childers J and Echols C (2004) 2 ½ -year-old children use animacy and syntax to learn a new noun. *Infancy* 5(1), 109–125.
- Cuza A (2016) The status of interrogative subject-verb inversion in Spanish–English bilingual children. *Lingua* 180, 124–138. <https://doi.org/10.1016/j.lingua.2016.04.007>
- Cuza A, Pérez-Leroux AT and Sánchez L (2013) The role of semantic transfer in clitic-drop among Chinese L1-Spanish L2 bilinguals. *Studies in Second Language Acquisition* 35(1), 93–125. <https://doi.org/10.1017/S0272263112000691>
- Cuza A and Pérez-Tattam R (2016) Grammatical gender selection and phrasal word order in child heritage Spanish: A feature re-assembly approach. *Bilingualism: Language and Cognition* 19(1), 50–68. <https://doi.org/10.1017/S1366728914000893>
- Daskalaki E, Elma BLOM, Chondrogianni V and Paradis J (2020) Effects of parental input quality in child heritage language acquisition. *Journal of Child Language* 47(4), 709–736.
- Dunn LM, Lugo DE, Padilla ER and Dunn LM (1986) *Test de Vocabulario en Imágenes Peabody*. Circle Pines, MN: American Guidance Service.
- Dunn M and Dunn LM (2007) *Peabody Picture Vocabulary Test—4*. Circle Pines, MN: American Guidance Service.
- Elff M (2021) Multinomial logit models with or without random effects or overdispersion. R package version 0.8.7.3. <https://CRAN.R-project.org/package=mclogit>
- Flores C and Rinke E (2020) The relevance of language-internal variation in predicting heritage language grammars. *Bilingualism: Language and Cognition* 23, 25–26.
- Flores C, Santos AL, Jesus A and Marques R (2017) Age and input effects in the acquisition of mood in Heritage Portuguese. *Journal of Child Language* 44, 795–828.
- Gathercole VCM and Thomas EM (2009) Bilingual first-language development: Dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition* 12(2), 213–217.
- Goebel-Mahrle T and Shin N (2020) A corpus study of child heritage speakers' Spanish gender agreement. *International Journal of Bilingualism* 24(5–6), 1088–1104. <https://doi.org/10.1177/1367006920935510>
- Gómez Seibane S (2013) *Pronombres átonos (le, la, lo) en español: aproximación histórica*. Madrid: Arco Libros.
- Gundel JK, Hedberg N and Zacharski R (1993) Cognitive status and the form of referring expressions in discourse. *Language* 69, 274307.
- Hartsuiker RJ, Pickering MJ and Veltkamp E (2004) Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish–English bilinguals. *Psychological Science* 15(6), 409–414.
- Hothorn T, Hornik K and Zeileis A (2006) Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15(3), 651–674. doi:10.1198/106186006X133933
- Hur E, López Otero JC and Sánchez L (2020) Gender agreement and assignment in Spanish heritage speakers: Does frequency matter? *Languages* 5(48). <https://doi.org/10.3390/languages5040048>
- Jia R and Paradis J (2015) The use of referring expressions in narratives by Mandarin heritage language children and the role of language environment factors in predicting individual differences. *Bilingualism: Language and Cognition* 18(4), 737–752.
- Kroll J, Dussias P, Bice K and Perrotti L (2015) Bilingualism, mind, and brain. *Annual Review of Linguistics* 1, 377–94.
- Labov W (1994) *Principles of Linguistic Change: Internal factors*. Mass: Blackwell.
- Martínez-Nieto L and Restrepo A (2022) Production and comprehension of grammatical gender by Spanish heritage speakers: Evidence from accusative clitic pronouns. *International Journal of Bilingualism*. <https://doi.org/10.1177/13670069211057318>
- Mayer E and Sánchez L (2017) Feature variability in the bilingual-monolingual continuum: clitics in bilingual Quechua-Spanish, bilingual Shipibo-Spanish and in monolingual Limeño Spanish contact varieties. *International Journal of Bilingual Education and Bilingualism* 22(7), 883–896. <https://doi.org/10.1080/13670050.2017.1322037>
- Montrul S and Sánchez-Walker N (2013) Differential object marking in child and adult Spanish heritage speakers. *Language Acquisition* 20(2), 109–132.
- Mykhaylyk R and Sopata A (2015) Object pronouns, clitics, and omissions in child Polish and Ukrainian. *Applied Psycholinguistics* 37(5), 1051–1082. <https://doi.org/10.1017/S0142716415000351>
- Ormazabal J and Romero J (2013) Object clitics, agreement, and dialectal variation. *Probus* 25(2), 301–344.
- Otheguy R (2016) The linguistic competence of second-generation bilinguals: A critique of “incomplete acquisition”. In Tortora C, den Dikken M, Montoya IL and O'Neill T (eds), *Romance Linguistics 2013. Selected papers from the 43rd Linguistic Symposium on Romance Languages (LSRL)*. Amsterdam: John Benjamins, pp. 301–320.
- Pérez-Leroux AT, Castilla A and Brunner J (2012) General and specific effects of lexicon in grammar: Determiner and object pronoun omissions in child Spanish. *Journal of Speech, Language, and Hearing Research* 55, 313–327.
- Pérez-Leroux AT, Pirvulescu M and Roberge Y (2008) Null objects in child language: Syntax and the lexicon. *Lingua* 118, 370–398.
- Pérez-Leroux AT, Pirvulescu M and Roberge Y (2018) *Direct objects and language acquisition*. Cambridge: CUP.
- Pescarini D (2021) *Romance object clitics. Microvariation and linguistic change*. Oxford University Press.
- Pirvulescu M, Pérez-Leroux A, Roberge Y, Strik N and Thomas D (2014) Bilingual effects: Exploring object omission in pronominal languages. *Bilingualism: Language and Cognition* 17(3), 495–510.
- Place S and Hoff E (2016) Effects and noneffects of input in bilingual environments on dual language skills in 2½-year-olds. *Bilingualism: Language and Cognition* 19(5), 1023–1041.

- Polinsky M and Scontras G** (2020) Understanding heritage languages. *Bilingualism: Language and Cognition* **23**, 4–20.
- Putnam MT and Sánchez L** (2013) What's so incomplete about incomplete acquisition?: A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism* **3**, 478–508. <https://doi.org/10.1075/lab.3.4.04put>
- R Core Team** (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Requena P** (2022) Variation versus deviation: Early bilingual acquisition of Spanish differential object marking. *Linguistic Approaches to Bilingualism*. <https://doi.org/10.1075/lab.21001.req>
- Rodina Y and Westergaard M** (2017) Grammatical gender in bilingual Norwegian-Russian acquisition: The role of input and transparency. *Bilingualism: Language and Cognition* **20**(1), 197–214.
- Sánchez L** (2003) *Quechua-Spanish bilingualism: Interference and convergence in functional categories*. Amsterdam: John Benjamins.
- Sánchez L** (2019) Bilingual alignments. *Languages* **4**(82). <https://www.mdpi.com/2226-471X/4/4/82>
- Schwenter S** (2006) Null objects across South America. In Face T and Klee C (eds), *Selected Proceedings of the 8th Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Proceedings Projects, 23–36.
- Shin N** (2022) *Está abriendo, la abrió*: The impact of verb type and verb form on Spanish-English bilingual children's direct object omission. *International Journal of Bilingualism*. <https://doi.org/10.1177/13670069221124475>
- Shin N and Miller K** (2022) Children's acquisition of morphosyntactic variation. *Language Learning and Development* **18**(2), 125–150. <https://doi.org/10.1080/15475441.2021.1941031>
- Shin N, Rodríguez B, Armijo A and Perera-Lunde M** (2019) Child heritage speakers' comprehension and production of direct object clitic gender in Spanish. *Linguistic Approaches to Bilingualism* **9**(4/5), 659–686. <https://doi.org/10.1075/lab.17029.shi>
- Silva-Corvalán C** (2014) *Bilingual Language Acquisition: Spanish and English in the first six years*. Cambridge: Cambridge University Press.
- Theakston AL, Lieven EVM, Pine JM and Rowland CF** (2001) The role of performance limitations in the acquisition of verb-argument structure: an alternative account. *Journal of Child Language* **28**, 127–152.
- Thordardottir E** (2015) The relationship between bilingual exposure and morphosyntactic development. *International Journal of Speech-Language Pathology* **17**(2), 97–114.
- Ticio E** (2015) Differential object marking in Spanish-English early bilinguals. *Linguistic Approaches to Bilingualism* **5**(1), 62–90.
- Unsworth S** (2019) Quantifying experience in heritage language development. In Schmid M and Köpke B (eds), *The Oxford handbook of language attrition*. Oxford University Press, pp. 434–45.
- Vihman V-A and Nelson D** (2019) Effects of animacy in grammar and cognition: Introduction to special issue. *Open Linguistics* **2019**(5), 260–267. <https://doi.org/10.1515/opli-2019-0015>
- Vihman V-A, Nelson D and Kirby S** (2018) Animacy distinctions arise from iterated learning. *Open Linguistics* **2018**(4), 552–565. <https://doi.org/10.1515/opli-2018-0027>
- Zagona K** (2000) *The Syntax of Spanish*. Cambridge: Cambridge University Press.
- Zdrojewski P and Sánchez L** (2014) Variation in accusative clitic doubling across three Spanish Dialects. *Lingua* **151**, 162–176. <https://doi.org/10.1016/j.lingua.2014.08.003>