Language Variation and Change (2024), **36**, 1–24
doi:10.1017/S0954394523000248

CAMBRIDGE UNIVERSITY PRESS

Early acquisition of syntactic variation: Lexical conditioning of Spanish variable clitic placement

Pablo E. Requena

The University of Texas at San Antonio, USA
Email: pablo.requena@utsa.edu

Abstract

This paper examines how children acquire Spanish variable clitic placement (VCP), a lexically conditioned phenomenon whereby clitics may precede or follow complex verb phrases. Research on how children acquire truly syntactic variable phenomena suggests that they either generalize one variant initially or they match the variation in the input from the beginning. Here I examine how children acquire the lexical conditioning of Spanish VCP. A corpus study of naturalistic conversations between parents and young children suggests that from the earliest ages examined (2;0-3;0) children display lexically-specific patterns that seem to be fine-tuned by the early school years. Experimental results using two different elicitation techniques with children ages 4;0-7;0 provide further support for early acquisition of the lexical conditioning of VCP and some evidence for fine-tuning during this age window. Thus, methodological triangulation enables detection of variable use where children would otherwise show categorical use of variants with infrequent syntactic phenomena, such as Spanish VCP.

Keywords: syntactic variation; child language acquisition; clitic; lexical constraint

Two questions that have puzzled linguists for years are when and how children acquire the grammar of the language(s) they hear in their community. More recently, the focus has been placed on children's mastery of morphosyntactic variation. Existing research suggests a developmental path consisting of stages that range from sequential emergence of variants to the production of forms in overlapping contexts. Shin and Miller (2022) reviewed existing research and proposed a four-step pathway in the acquisition of morphosyntactic variation in which children's early use of variable forms is initially sequential, with one variant being used over the other almost exclusively. After this early regularization of one variant across the board, they find a stage where children use multiple variants, but in mutually exclusive contexts. Subsequently, children begin to show overlap in their use of variants in the same contexts, which approximates community patterns as the child is exposed to more input. Research in this vein also

shows that while children seem able to acquire morphosyntactic variation very early, the timing of acquisition of particular variable forms depends on the type and complexity of the morphosyntactic variable in question, its salience and use in the input, as well as children's cognitive development (e.g., Labov, 1989; Miller, 2013; Shin, 2016; Smith & Durham, 2019).

Most studies leading to such observations have focused on morphophonological or morphological variables that consist of the expression versus omission of a form (Hendricks, Miller, & Jackson, 2018; Kovac & Adamson, 1981; Miller, 2013; Miller & Schmitt, 2012; Shin, 2016) or the alternation between one or more overt forms (Miller, 2015; Smith & Durham, 2019; Smith, Durham, & Fortune, 2007). Fewer studies have concentrated on truly syntactic variables, such as word order, and the results of those that do exist are inconclusive. Some research has shown that children display an initial bias for one word order over another (Anderssen, Bentzen, Rodina, & Westergaard, 2010), while other work has found early input matching in variable word order acquisition (Anderssen & Westergaard, 2010). More recent studies show that some constraints are still developing between ages 6-8 (see Shin, 2021). It thus remains an empirical question whether children's acquisition of syntactic variation takes place early and whether it proceeds in the stages outlined in Shin and Miller (2022). The present study addresses this gap in the literature by examining the acquisition of lexically conditioned word order variation in Spanish variable clitic placement (henceforth, VCP).

Spanish clitics appear categorically before finite verbs [clitic+finite verb] or after nonfinite verbs and imperatives [nonfinite verb+clitic]. However, a number of [finite verb+nonfinite verb] constructions display variation that does not change the base meaning of the utterance. In example (1) the variation is found in the two ways speaker B may respond to speaker A's question. B1 exemplifies a preverbal clitic position (henceforth *proclicis*) and B2 a postverbal enclitic position (henceforth *enclisis*).

(1) A - *¿No vas a termina-r el pastel?*
NEG go-PRS.2SG to finish-INF the cake
'Aren't you going to finish the cake?'

B1 - ***Lo*** *voy a come-r después* (proclisis)
it-ACC-M3SG go-PRS.1SG to eat-INF later
'I am going to eat **it** later.'

B2 - *Voy a come-r=**lo** después* (enclisis)
go-PRS.1SG to eat-INF=**it-ACC-M3SG** later
'I am going to eat **it** later.'

While this alternation in clitic placement does not change the base meaning of the utterance, the choice between proclisis and enclisis also does not constitute a case of inconsistent variation. Instead, corpus studies of adult-to-adult speech indicate that VCP is systematically (and probabilistically) conditioned by register, by properties of the finite verb, and by semantic and discourse properties of the clitic referent, such as animacy and topic persistence (see Davies, 1995; Requena, 2020 and references therein;

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

Schwenter & Torres Cacoullos, 2014). Most notable, however, is that variationist studies across dialects of Spanish coincide in identifying the finite verb as the main factor conditioning VCP. While there seems to be a clear lexical effect (some verbs showing greater enclisis rates than others), Schwenter and Torres Cacoullos (2014) found evidence for a grammaticalization effect, such that more frequent and grammaticalized finite verbs probabilistically favored proclisis while infrequent and less grammaticalized finite verbs tended to favor enclisis (see Myhill 1988a; 1988b; Requena, 2020). Despite all that we know about how adult speakers use VCP, no research has examined whether such systematicity in VCP use is also present in child-directed speech and how young children acquire such patterns of VCP use.

The present study draws on both corpus and experimental methods to investigate when and how children acquire the variable distribution of clitics with different lexical items. Through an analysis of child-caregiver conversational speech data (children ages 2;0-5;0), I investigate clitic placement at the very earliest ages of acquisition. This corpus study provides an opportunity for comparing VCP in adult-to-adult speech to VCP in child-directed speech, which has not been done before. Secondly, through two elicited production tasks, I examine children's (ages 4;0-7;0) variable production in VCP constructions with *select* verbs that differ in adult speech in terms of how strongly each verb favors proclisis versus enclisis. Through methodological triangulation, I am able to examine individual children's knowledge of target VCP grammar through experimental tasks, an approach to the study of the acquisition of sociolinguistic variation that holds promise for the study of infrequent morphosyntactic phenomena.

Background

Acquisition of morphosyntactic variation

To acquire adult-like use of variable forms, children not only need to learn which variants coexist but also the patterns of use found in their language community. Kerswill (1996:199) noted that "exactly when a child acquires a feature of his or her first dialect depends on the linguistic level, the complexity of the conditioning, and the child's age." Accordingly, a comprehensive description of when and how children acquire adult-like use of variable forms remains an empirical question given the different levels where variation can be found (e.g., lexicon, phonology, morphology, syntax), as well as the diversity in the nature of the variable forms (e.g., involving production versus omission, substitution, placement) and in the number and types of constraints that condition the use of one variant over another one (e.g., linguistic, social). In what follows, I review the relevant research on the acquisition of variation leading us to highlight the limited existing knowledge about the acquisition of truly syntactic variation consisting of word order.

To date, there is a growing number of studies addressing children's acquisition of phonological variation but very few studies on syntactic variation. In addition, most studies investigate conversational data between children and their caregivers or another adult, but fewer have carried out elicitation tasks to obtain forms that are less frequently found in a corpus. Taking much of this previous work into consideration, Shin and Miller (2022) delineated four general phases in children's acquisition of variation. They noted that, taken together, much of the previous literature suggested a

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

four-step pathway whereby children initially produce only one of the variants of the variable form across all possible contexts of use (Step 1), followed by a period where they produce more than one variant but in mutually exclusive contexts (Step 2). Shin and Miller offered several explanations for Step 1, including children's documented tendency to regularize the grammar. However, it is also the case that Step 2 can also be the outcome of regularization across some contexts (for variant A) and other contexts (for variant B). For clarity, I will refer to this type of regularization in Shin and Miller's Step 2 as "regularization across some contexts." After these first two phases, children begin to show overlap in their use of variants in the same contexts (Step 3), although children's variable production may not completely match that found in their speech community. In Step 4, children's variable usage patterns more closely with their speech community.

An example comes from Miller's (2015) longitudinal study on *ain't* versus *isn't* in Sarah's production from the Brown corpus (Brown, 1973). Early on, from 2;0-4;0 years of age, Sarah initially only produced *isn't* with third person singular subjects, and most of those utterances occurred in declarative sentences. At a later period, from 4;0-5;0 years of age, Sarah produced both variants, *ain't* and *isn't,* but she did so in mutually exclusive contexts. She mostly produced *isn't* in interrogative constructions (especially tag questions) and *ain't* in declarative constructions. Miller (2015) noted that Sarah's usage was consistent with, yet more extreme than, patterns found in adult speech. Washington and Craig (2002), for example, found that adult caregivers rarely produced *ain't* in tag questions and that the children in their study, like Sarah, never produced *ain't* in these contexts. After these two initial phases, the authors predicted that Sarah would later show overlap in her use of *ain't* and *isn't*, and this would approximate more and more the patterns found in her community across age.

When it comes to syntactic variables, regularization has been reported in previous research on word order variation, particularly in cases where variants occur in complementary distribution or where one variant is more frequent than the other. Anderssen and colleagues (2010) investigated subject placement in Norwegian where subjects may variably precede or follow negation. Adult speakers more frequently produced the lower position [neg+S_{DP}] with lexical subjects (60/62, 97%) and the shifted position [S_{Pro}+neg] with pronominal subjects (758/864, 88%) (Anderssen & Westergaard, 2010). Children also chose the lower position with lexical subjects; however, as predicted by Shin and Miller (2022), children also initially regularized the lower position [neg+S] to pronominal subjects despite pronominal subjects being the more frequent subject-type in the input. Only later, by 2;6-3;0 years of age, children switched their preference toward the shifted position [S_{Pro}+neg] with pronominal subjects, similar to the patterns found in their speech community (for examples of regularization in artificial languages, see Saldana, Smith, Kirby, & Culbertson [2021] and references therein).

An effect of lexical verb was reported in a study of Spanish variable subject placement. Shin (2021) examined the SV-VS variation in naturalistic speech by monolingual children ages six to eight. The results indicated that by the beginning of their elementary education, children used properties of the subject (e.g., syntactic and pragmatic) to constrain SV-VS word order. While the semantic verb class effect (change of location verbs favoring VS) was not attested among the children, there was evidence for

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

more frequent verbs favoring VS compared to less frequent verbs within this category. These results may suggest that larger effects that characterize the target grammar may develop among children through fine-tuning impacting particular lexical items (e.g., high-frequency items) first.

The two variants in Spanish VCP (enclisis and proclisis) seem to emerge simultaneously instead of sequentially—as predicted by Step 1 in Shin and Miller (2022) (Rodríguez Mondoñedo, Snyder, & Sugisaki, 2004). In a sentence repetition study with children ages 3;0-6;4, Eisenchlas (2003) reported full grammatical competence in clitic placement since age three and documented a preference for proclisis over enclisis overall. But the acquisition of the lexical constraint on the variation was not examined in those studies. Through sociolinguistic interviews, Shin, Requena, and Kemp (2017) showed that children are sensitive to the verb lexeme in their variable clitic placement preferences between 6;3-11;9 years of age. Although this study was very useful in documenting the later stages of acquisition, the results were not reported by participant, nor by age, making it difficult to determine whether younger school-aged children exhibit signs of regularization across some contexts (lexical constructions for VCP) or not, as discussed in Step 2 in Shin and Miller (2022).

Here I ask how VCP might be instantiated within Shin and Miller's four-step pathway to the acquisition of variation. My question is not only whether acquisition of VCP is consistent with this pathway or not, but I ask how the present data might further inform the various phases of their proposed pathway. Spanish-speaking children do not seem to go through a regularization phase (Step 1; Rodríguez Mondoñedo et al., 2004). So, here I focus on questions that arise once both variants are part of the child's grammar (i.e., Steps 2 onward) and ask:

1. Do children go through a phase when they use proclisis and enclisis in restricted rather than overlapping contexts (Step 2)? Specifically, do children initially regularize proclisis to a set of verbs and enclisis to a different set of verbs?
2. Once children begin to show use of both VCP variants in overlapping contexts (Step 3), what differences between children's variable usage and that of the adults in their speech community remain? In other words, does VCP become target-like with some lexical constructions before others?

Spanish variable clitic placement (VCP)

In Spanish, when clitics are used as objects to the nonfinite verb in [finite verb+nonfinite verb] constructions, two available positions exist.[1] The clitic may precede the finite verb (proclisis) or follow the nonfinite verb (enclisis)—as shown in (1)—without resulting in any change in the base meaning of the utterance.

The main constraint identified by VCP research is lexical. Particular finite verbs systematically exhibit different rates of enclisis, which are consistent across dialects (see, for example, Davies [1995] & Requena [2020] for Argentine Spanish, and Schwenter & Torres Cacoullos [2014] for Mexican Spanish). As reported by Davies (1995), across varieties of Spanish there is a continuum-like distribution of verbs according to their frequencies of VCP (see Figure 1).

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

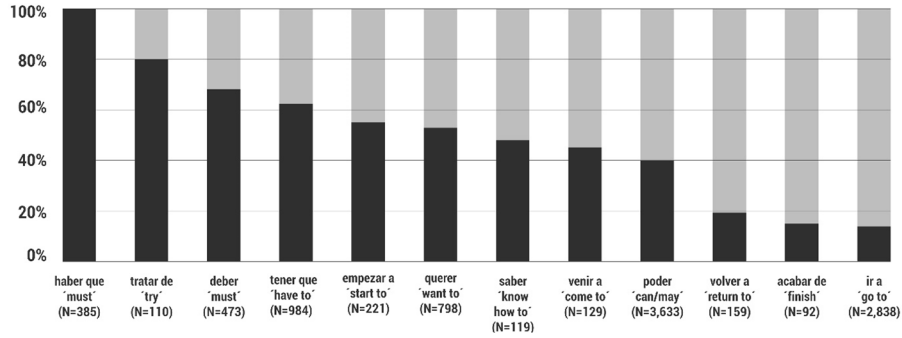

Figure 1. Rate of enclisis (black) and proclisis (gray) by finite verb in spoken Spanish (adapted from Davies, 1995:374).

One proposal for how to account for this continuum invokes the degree of grammaticalization of such verbs. Myhill (1988a, 1988b) observed that finite verbs heading [finite verb+nonfinite verb] constructions that frequently appear in proclisis have grammaticalized meanings. For example, *estar* 'be' in [*estar*+gerund][2] or *ir* 'go' in [*ir a* 'go to'+infinitive] have progressive and future meanings, respectively, and both favor proclisis (gray area in Figure 1). In contrast, *querer* 'want' or *tratar de* 'try' in similar constructions have meanings that are more lexical in nature, and they both favor enclisis (black area in Figure 1). Corpus on adult-to-adult speech data support Myhill's grammaticalization account (e.g., Davies, 1995; Schwenter & Torres Cacoullos, 2014), but one apparently exceptional case has also been reported. Despite allowing VCP, the highly frequent and grammaticalized verb *tener que* 'have to' favors enclisis in adult-to-adult speech. When examining this particular verb, Requena (2020) proposed that the behavior of the [*tener que*+infinitive] variable construction can be accounted for by the fact that *tener* is a relatively recent instance of grammaticalization and that, as such, it retains some analyzability and paradigmatic links with elements outside the variable construction (see Bybee, 2010).

Studies

Following the variationist tradition, Study 1 is a corpus study of naturalistic child-caregiver conversations that describes the distribution of clitics with specific lexical verbs at the earliest stages in language development. These are the most ecologically valid data we can examine to describe early VCP use. In Studies 2 and 3, I investigate how children use VCP across lexical constructions through an elicited production and a sentence repetition task.

Study 1. Naturalistic production

Corpus and data extraction and coding

A total of 125 hours of spontaneous conversations between Mexican children and their caregivers ($n = 25$; ages = 1;06-5;03) from the *Mexican Child-Caregiver Corpus* (Miller & Schmitt, 2012)[3] was included in the analysis. Families were from Mexico

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

City, and recordings of child-caregiver dyads were made over multiple sessions in the children's homes while they played with their caregivers. First and second pass transcriptions were later carried out by native or near-native speakers using the *CLAN* program (MacWhinney, 2000).

To examine VCP, I extracted all contexts in which third-person (3p) direct object (DO) clitics occurred in proclisis or enclisis in [finite+non-finite verb$_{(\text{gerund, infinitive})}$] constructions. Exact repetitions were excluded as well as other cases (see Appendix A in the Supplementary Materials for a list of exclusions). This process of data extraction yielded a total of 1,120 tokens of 3p DO clitics in variable contexts, as in examples (2) and (3). Of these, 776 tokens were produced by the caregivers, and 344 tokens were produced by the children.

(2) ***los*** *voy* *a* *rescata-r*
them.ACC.M3PL go.PRS.1SG to rescue-INF
'I'm going to rescue **them**.'
(Sami_4;3c, line 711 - child)

(3) *voy* *a* *rescata-r=**los***
go.PRS.1SG to rescue-INF=**them.ACC.M3PL**
'I'm going to rescue **them**.'
(Sami_4;3c, line 812 - child)

After exclusions, five children were left with no tokens of clitics in variable contexts (four were younger than 2;5 years old) and were thus not included in further analyses, although their caregivers' data were included. That left twenty children (2;2 to 5;3) and twenty-five adults for analysis, listed below in Table 1.

As with other syntactic variables, VCP is relatively infrequent in naturalistic production. However, the present study constitutes the best attempt at the study of VCP in early child language by extracting all instances of VCP found in the entire corpus. Each token was coded for *Speaker* (child versus caregiver), *Age* (2;0, 3;0, 4;0, 5;0, adult), *Finite verb* (e.g., *querer* 'want,' *ir a* 'go to,' etc.), and *Clitic position* (proclisis versus enclisis). Hortative uses of *ir a* were identified as 1pl forms (*vamos a…*) that could be translated as *let's* into English. Given that, analyses by individual age groups 2;0, 3;0, 4;0, and 5;0 would render groups with very small data volumes, so ages 2;0 and 3;0 were combined to form a group of younger children ($n = 8$), as were ages 4;0 and 5;0 to form a group of older children ($n = 12$) (see darker line on Table 1 for information on individual participants in each age group).

Results

Overall distributions

The overall rate of enclisis in the child-directed speech was 28% (218/776). This rate resembles the rate reported for Mexican adult-to-adult speech (27% in Schwenter & Torres Cacoullos, 2014). Overall, children produced 29% (98/342) enclisis, which matches the distribution in their input. This rate of enclisis was found in both the younger children (57/200) and the older children (41/142).

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

Table 1. Production of clitics in variable contexts by participant

#	Child	Age	Instances of VCP in child speech	Instances of VCP in caregiver speech
1	Raquel	1;6	–	25
2	Roberto	1;9	–	35
3	Mario	2;2	1	55
4	Augusto	2;4	–	46
5	Florencia	2;5	–	35
6	Sergio	2;8	1	36
7	Carla	2;8	3	78
8	Marcos	2;11	4	14
9	Sabrina	2;11	27	29
10	Martín	2;11	43	83
11	Lorena	2;11	21	22
12	Elizabeth	2;11	20	20
13	Flavia	2;11	36	21
14	Andy	2;11	5	7
15	Alicia	3;5	38	30
16	Santiago	3;8	2	13
17	Gaspar	4;1	21	21
18	Sami	4;3	16	47
19	Marcela	4;8	19	34
20	Eduardo	4;10	19	27
21	Antonella	4;10	25	26
22	Rocío	5;0	2	16
23	Andy	5;2	25	19
24	Daniel	5;2	15	22
25	Juan	5;3	–	15
TOTAL			342	776

The rate of enclisis in caregiver production varies considerably among caregivers (ranging from 5% to 54% enclisis) and even more so among children (ranging from 7% to 85% enclisis). Figure 2 displays the thirteen dyads where the children produced at least fifteen instances of VCP, following recommendations for minimum number of tokens to conduct individual analyses (Guy, 1980:20). Data from seven children who only produced between one and five VCP contexts were excluded (see Appendix B for data on these children). The dyads appear from left to right organized by rate of enclisis in the caregivers' speech.[4] As the linear trend lines in Figure 2 show, children's overall rates of enclisis do not match the rates of enclisis found in their own caregiver's speech. Indeed, there was no correlation between the two ($r = –.178$, $df = 12$, $p = .561$).

Upon further examination, differences between dyads emerge. Of the thirteen children included in this analysis, seven produced enclisis at different frequencies

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

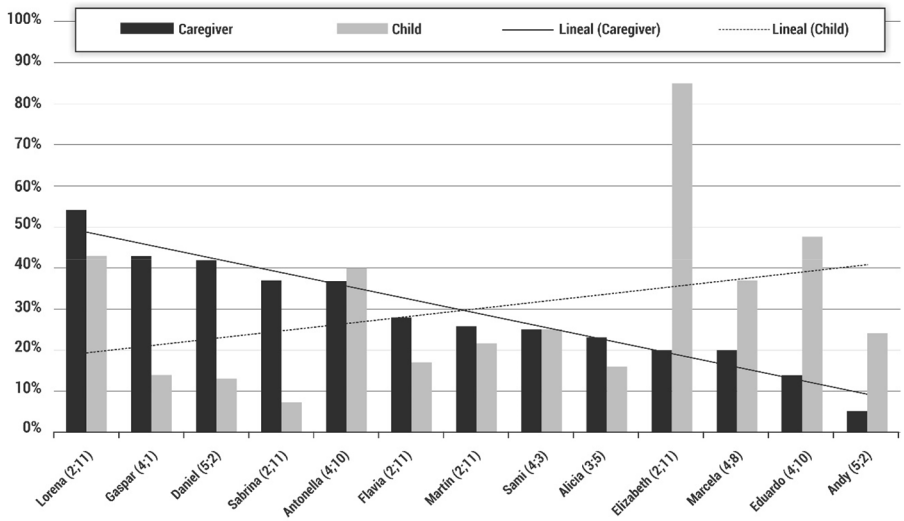

Figure 2. Rate of enclisis by individual caregiver-child dyads for children who produced more than fifteen tokens of VCP.

than their caregivers. Three children (Gaspar, Daniel, and Sabrina) used enclisis less frequently than their respective caregivers and four children (Elizabeth, Marcela, Eduardo, and Andy) used enclisis more frequently than their respective caregivers. The rest of the children (Lorena, Antonella, Flavia, Martin, Sami, and Alicia) produced overall rates of enclisis that approximated those of their caregivers. The question this raises is why do half of the children in the corpus (7/13) not match the enclisis rate in the input? Do some of them overextend the use of enclisis (Elizabeth, Marcela, Eduardo, Andy) while others do so with proclisis (Gaspar, Daniel, and Sabrina)?

Since some finite verbs favor enclisis more than others, it might be that children who produced more enclisis than their caregivers did so because they used more finite verbs that favor enclisis. Table 2 shows the types of verbs children and their caregivers produced, verbs that tend to favor or disfavor enclisis in adult speech. Larger values in the rightmost column correspond to larger differences in the number of enclisis-favoring contexts produced by caregiver and child. The dyads at the top of the table correspond to the children who matched their input in overall enclisis rate. As the rightmost column shows, these children differed less from their own caregivers (in the amounts of enclitic-favoring verb contexts) than the dyads at the bottom of the table. Interestingly, the four children whose overall rate of enclisis greatly exceeded that of their caregivers (Elizabeth, Marcela, Eduardo, Andy) produced more tokens of enclisis-favoring verbs (see positive numbers on the rightmost column, which indicate that, compared to their caregivers, these children produced relatively more enclisis-favoring verbs). In contrast, the three children whose overall rate of enclisis was lower than that of their caregivers (Gaspar, Daniel, and Sabrina) produced fewer enclisis-favoring verbs relative to their caregivers (see negative values).

To examine the age at which children's productions are conditioned by lexical verb, Figure 3 shows the rate of enclisis for the caregivers and the younger and older

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

Table 2. Rate of enclisis-favoring contexts (out of all VCP produced by each participant in dyads where children match and did not match the input)

		Percentage of enclisis-favoring contexts (out of total VCP contexts)		
	Name of child in dyad	*Child*	*Caregiver*	Difference (percentage points child vs. adult)
Children who matched the input in overall enclisis rate	Lorena	43%	54%	−11
	Antonella	40%	37%	2
	Flavia	17%	28%	−11
	Martin	22%	26%	−4
	Sami	25%	25%	0
	Alicia	16%	23%	−7
Children who did not match the input in overall enclisis rate	Gaspar	14%	43%	−29
	Daniel	13%	42%	−29
	Sabrina	7%	37%	−30
	Elizabeth	85%	20%	65
	Marcela	37%	20%	17
	Eduardo	47%	14%	33
	Andy	24%	5%	19

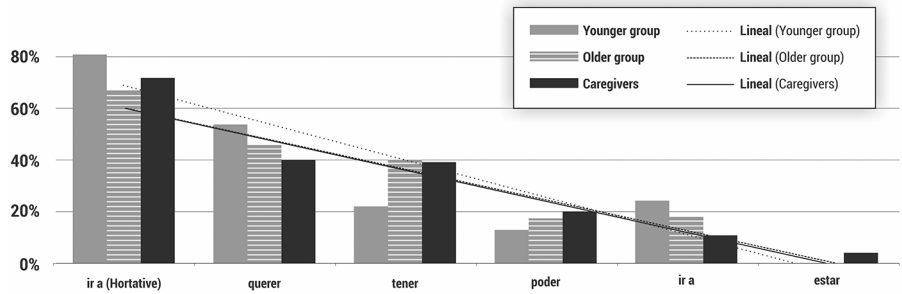

Figure 3. Rate of enclisis by frequent constructions in both child groups and in caregivers.

children groups. As with previous corpus studies, *ir a*, *estar*, and *poder* display lower rates of enclisis, and *querer* and *tener* display higher rates of enclisis in both children and adults. Nonetheless, older children pattern more closely to adults than younger children do, and a developmental pattern is observed such that children seem to reduce their use of enclisis with hortative *ir a*, *querer*, and *ir a* as they increase in age, and they also seem to increase their use of enclisis with *tener* and *poder* across age. There is

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

some indication, therefore, that children may fine-tune these lexically specific variable patterns throughout the preschool years.

Due to the lack of statistical power, regression models were deemed inappropriate for the analysis of the children's enclisis/proclisis with individual verbs. Therefore, the analysis of individual verbs rests on the descriptive data presented above, which suggest very early target (or near-target) clitic placement by verb and, possibly, a process of fine-tuning of lexically specific patterns of VCP with some verbs between ages of 2;0-3;0 and 4;0-5;0. Still, analyses that collapse verb lexemes according to their clitic placement preferences are still possible because the size of the dataset would make them more reliable. Shin et al. (2017) grouped proclisis-favoring verbs on the one hand (*ir*-future, *estar*) and enclisis-favoring verbs on the other (*querer*, *tener*, infrequent verbs). Their results showed that rates of enclisis were higher with the latter, as expected if children aged 6;0-11;0 know the clitic placement patterns displayed by these verbs. For the purposes of the present study, data were recoded grouping verbs that favor proclisis (*estar*+gerund, *ir*+gerund, *poder*, *ir a*-future) into one category and verbs favoring enclisis (*deber*, *ir a*-hortative, *querer*, *saber*, *tener que*, *volver*) into another category. A generalized linear mixed model (GLMM) analysis was performed in *IBM SPSS Statistics* (IBM Corp., 2021) with *Clitic Placement* as dependent variable. *Verb class* (recoded as binary variable) and *Age* were included as explanatory variables. The initial model also included their interaction as well as pairwise comparisons. *Speaker* was included as a random intercept. Neither the interaction nor *Age* were significant in the initial model, and they were removed one at a time. The final model, with the best AIC and BIC, included just *Verb* recoded as an explanatory variable (See Appendix C for estimates). A significant effect of this factor ($p < .001$) indicates that speakers were more likely to use enclisis with enclisis-favoring verbs than with proclisis-favoring verbs. Study 1, thus, shows that very young Spanish-speaking children appear to use information on the verb to guide VCP in naturalistic production. This study extends the existing research to much younger children. Inferential statistics reveal that very young children associate some verbs with enclisis more than others, mirroring the input (Schwenter & Torres Cacoullos, 2014). Descriptive statistics by individual verbs suggest that some knowledge about particular verbs' patterns with VCP may be already acquired by the youngest group. I also found indication that these younger children may engage in a process of fine-tuning of already pretty close-to-target patterns. In order to examine individual verbs in a way that analysis by verb are reliable, experimental techniques were developed in Studies 2 and 3 and tested with children ages 4;0-7;0.

Study 2. Elicited production

Participants

Sixty-two children between 4;1-7;0 ($M = 5;7$) were recruited from private preschools in Córdoba, Argentina. Of these sixty-two children, fifty-one produced variable clitic structures and were thus included in the analysis. Table 3 shows the classification of child participants by age group. Eleven adults from the same local community also participated in the study. One was excluded from the analysis for not producing any

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

Table 3. Age groups based on composite score means and standard deviations for child participants ($n = 51$)

Group	Age
4;0 ($n = 13$)	$M = 4;5$ ($SD = 0;2$)
5;0 ($n = 17$)	$M = 5;5$ ($SD = 0;2$)
6;0 ($n = 21$)	$M = 6;6$ ($SD = 0;3$)

variable contexts and, as such, the analysis was carried out with data from the ten remaining adult speakers.

Stimuli and procedure

The task used in this study was adapted from Thomas (2012). Six large-sized cards were created (see sample in Figure 4), two for each verb condition: *ir a* 'to go,' *querer* 'to want,' and *tener que* 'to have to.' Each card introduced a pair of familiar cartoon/TV characters that were immediately visible in the middle of the card. Each card also contained two folded ends (referred to during the task as "windows") that served to cover thought bubbles depicting what each character was "going to do" (using *ir a*), "wanted to do" (using *querer*), and "had to do" (using *tener que*) with an object or animal, depending on the trial. The corresponding trial for Figure 4, eliciting *ir+a* constructions, is shown in (4) (see Appendix D for the complete set of situations).

(4) **Experimenter:** *La Chilindrina y el Chavo van a tener un perrito. Contame qué van a hacer con el perrito.* ('Chilindrina and Chavo are going to get a puppy. Tell me what they are going to do with the puppy.')

Experimenter: *¿Qué va a hacer la Chilindrina con el perrito?* ('What is Chilindrina going to do with the puppy?')

Participant: (Child lifts window on the left to reveal Chilindrina's thoughts of bathing the puppy) *La Chilindrina lo va a bañar/va a bañarlo.* ('Chilindrina is going to bathe it.')

Experimenter: *¿Qué va a hacer el Chavo con el perrito?* ('What is Chavo going to do with the puppy?')

Participant: (Child lifts window on the right to reveal Chavo's thoughts of drying the puppy) *El Chavo lo va a secar/va a secarlo.* ('Chavo is going to dry it')

Responses were coded by noting first whether participants produced a variable [finite verb+infinitive] construction and a clitic pronoun. If this was the case, the response was further coded for the finite verb produced. If a participant produced VCP with one of the three verbs tested here (*ir a*, *querer*, and *tener que*), regardless of the verb used in that particular prompt, the response was included in the analysis.

Adult participants produced 132 answers. Of those, exclusions consisted of sixteen cases of invariable contexts (5); fourteen cases that did not contain a clitic pronoun, but a full NP direct object instead (6); and twenty-one cases where participants made

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

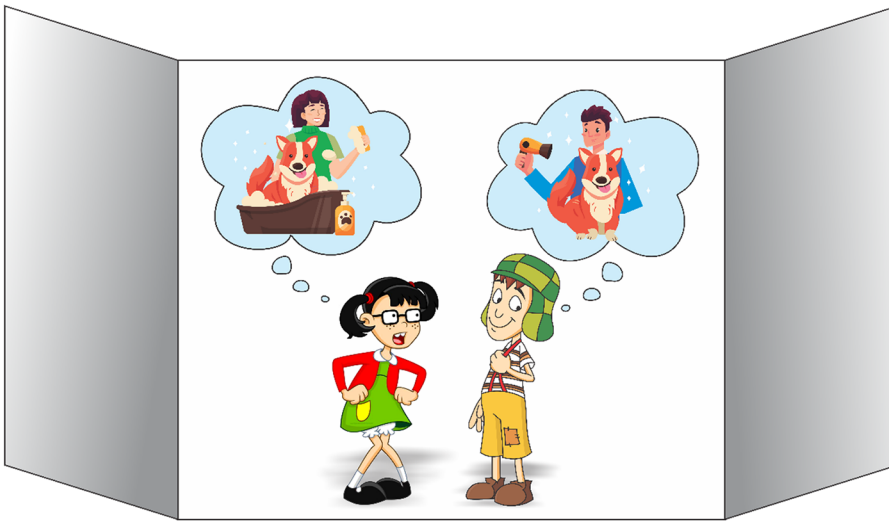

Figure 4. Sample visual support for elicited production.

substantive changes in the lexical construction, such as the insertion of a second non-finite element (7). This resulted in a total of eighty-one tokens being included in the final analysis.

(5) *Donald* ***lo*** *abre.*
Donald **it.ACC.M3SG** open
'Donald opens it'
(Participant 127)

(6) *El Chavo va a secar el perrito.*
the Chavo go.PRS.3SG to rinse the puppy
'Chavo is going to rinse the puppy'
(Participant 129)

(7) *La Chilindrina* ***lo*** *va a intentar abrir*
the Chilindrina **it.ACC.M3SG** go.PRS.3SG to try open
'Chilindrina is going to try to open it'
(Participant 122)

Children produced other nontarget, albeit felicitous, responses. Data cleaning resulted in a total of 304 tokens being included in the final analysis (see Appendix E for more details).

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

Results

Figure 5 presents the rates of enclisis by verb in each of the age groups. As can be seen, VCP was dependent on the finite verb construction among adults: they produced more enclisis with *querer* and *tener que* than with *ir a*. The descriptive data indicate that, overall, children also showed this pattern, but children's usage also became more like adult usage as children's age increased, especially with *ir a*.

To test whether the finite verb predicted the probability of enclisis, I used a GLMM with a logit-link and binomial error distribution. The binary response variable was clitic placement (proclisis versus enclisis). The explanatory variables were *Finite Verb* (*ir a*, *querer*, *tener que*) and *Age* (4;0, 5;0, 6;0, 7;0). I included an *Age*Verb* interaction in order to test whether VCP with particular verbs differed by age. I also included *Participant* and *Trial* as random intercepts. The postestimation settings included the Residual approximation as well as robust estimation in tests of fixed effects and coefficients, which helps manage violations of model assumptions. Model comparison using the Akaike information criterion (AIC) and the Bayesian Information Criterion (BIC) suggested dropping the *Age* explanatory variable and the *Trial* random intercept as well as the insignificant interaction from the final model. Therefore, I will report on the model with *Finite Verb* as an explanatory variable and *Participant* as a random intercept.

Results from the GLMM showed an observed association of *Finite Verb* ($p < .001$). Use of enclisis was lower for *ir a* compared to *tener que* but not for *querer* compared to *tener que*. Model estimates are provided in Appendix F.[5] Given the drop in enclisis with *ir a* (Figure 5), I followed a reviewer's suggestion to check for age differences with this finite verb only. I first used a GLMM that produced a Hessian warning due to a lack of between variance. As a result, a generalized linear model was used to analyze the data. *Age* was entered as a predictor, and pairwise comparisons were included.

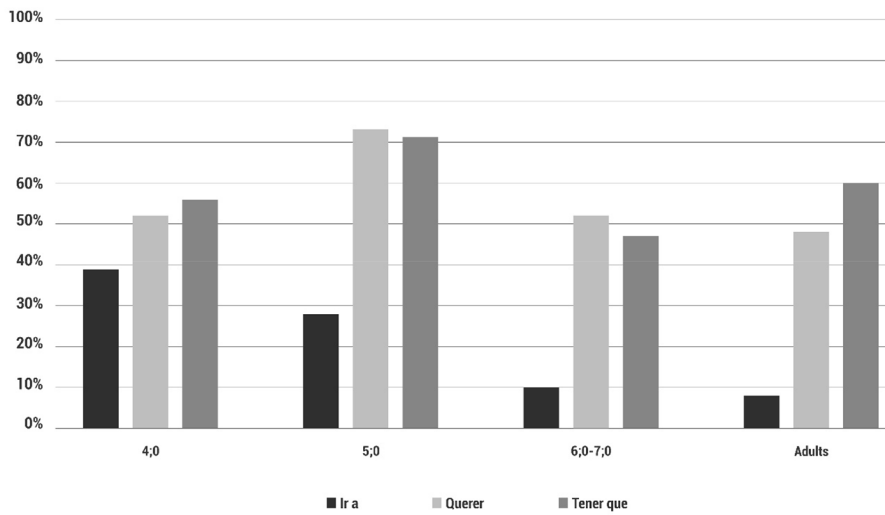

Figure 5. Rates of enclisis by verb in each of the age groups in the elicited production task.

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

The results indicate a significant main effect for age groups in clitic placement with *ir a* (Wald $\chi^2 = 7.049$, $df = 2$, $p < .05$). Pairwise comparisons show no significant difference between four- and five-year-olds or between four- and six-year-olds, but the six-year-olds used enclisis significantly less than the five-year-olds with this verb. An analysis at the individual level, however, reveals that many participants (both children and adults) responded using only one clitic position across all trials (see Appendix G for details). Of the children who produced at least five tokens, fourteen produced only one clitic position across trials (eight were categorically proclitic, and five were categorically enclitic).[6] Such nonvariable production is puzzling. Given that both variants have been attested in much younger children (Rodriguez Mondoñedo et al., 2004; Study 1 here), I doubt that this is indication of Step 1 in the acquisition path outlined by Shin and Miller (2022)—that is, regularization of a single variant across the board. Instead, it is more likely, in my opinion, that the findings of Study 2 may display task effects. To find out, Study 3 tests the same participants in a sentence repetition task. If children are regularizing across the board, I may find that children who categorically produced just one variant in Study 2 will produce the same variant throughout Study 3, regardless of clitic position in the repetition prompts. This would suggest that those participants operate with just one variant (and thus that they are not at the stage of interest in this study, which is when children already use both variants, or Step 2 in Shin & Miller [2022]).

Study 3. Sentence repetition

Participants

Participants were the same as in Study 2. All sixty-two children produced VCP as part of their repetitions. Table 4 shows a distribution of child participants according to age. However, due to technical problems with the recordings, data from two adult participants were excluded. Thus, the adult group in Study 3 consisted of nine participants.

Table 4. Age groups based on composite score means and standard deviations for child participants ($n = 62$)

Group	Age
4;0 ($n = 17$)	$M = 4;5$ ($SD = 0;2$)
5;0 ($n = 20$)	$M = 5;4$ ($SD = 0;2$)
6;0 ($n = 25$)	$M = 6;6$ ($SD = 0;3$)

Stimuli and procedure

Participants were asked to repeat sentences in two conditions (proclisis and enclisis). Each condition contained twelve experimental sentences divided by three finite verbs: *ir a*, *querer*, and *tener que*. Referent animacy and sentence length were controlled for across conditions. All stimuli were prerecorded by a native speaker from the same local area as the participants. Recordings were auditorily checked to avoid pauses and salient peaks in intonation.

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

Each stimulus sentence was preceded by a short preamble (read aloud by the experimenter and accompanied by visual support). The preamble introduced a masculine, singular, indefinite noun that became the referent of the DO clitic in the repetition stimulus. The clitic *lo* 'him/it' was used throughout the experiment since the clitic *lo* has the highest frequency among all 3p DO clitics in naturalistic production corpora (for sample stimuli, see examples [12] and [13] in next section).

Scoring

All responses were transcribed by a native Spanish-speaking research assistant who was from the same local area as the children and later checked for accuracy. For the purposes of this paper, I will focus only on inaccuracies in repetitions that involve changing the placement of the clitic, which I call *clitic repositioning*. Examples of forward and backward repositioning are provided in (12) and (13), respectively.

(12) *Forward repositioning (clitic repositioning into proclisis)*
Stimulus: *A la siesta Mulán va a ver**lo** con sus amigos.*
'At nap time Mulan is going to see **it** with her friends'
Imitation: *A la mañana Mulán **lo** va a ver con sus amigos.*
'In the morning Mulan is going to see **it** with her friends'
(P12, 4;8)

(13) *Backward repositioning (clitic repositioning into enclisis)*
Stimulus: *Mañana el hada madrina **lo** tiene que ver de cerca.*
'Tomorrow the Fairy Godmother has to see **it** closely'
Imitation: *Mañana tiene que ver**lo** de cerca.*
'Tomorrow has to see **it** closely'
(P5, 5;0)

Results

Adults produced inaccurate repetitions in only 4% of the trials (9/210), but the children produced inaccurate repetitions in 47% (694/1488) of the trials. Of these inaccurate repetitions, 60% (420/694) involved errors repeating the clitic, which is in line with previous studies (Eisenchlas, 2003). These errors included clitic copying in both positions ($n = 15$), clitic substitution ($n = 99$), clitic omission ($n = 66$), and considerably more errors of clitic repositioning ($n = 240$). Clitic repositioning errors thus amounted to 35% of all inaccurate repetitions and those are the focus of this study.

Figure 6 shows clitic placement in children's imitations by condition. As can be seen, children exhibited high accuracy rates when repeating clitics in both positions. However, children moved the clitic to a preverbal position in the Enclisis Condition more often than they moved the clitic to a postverbal position in the Proclisis Condition. This general pattern mirrors previous studies (Eisenchlas, 2003; Pérez-Leroux, Cuza, & Thomas, 2011). With respect to particular verbs, a look at the

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

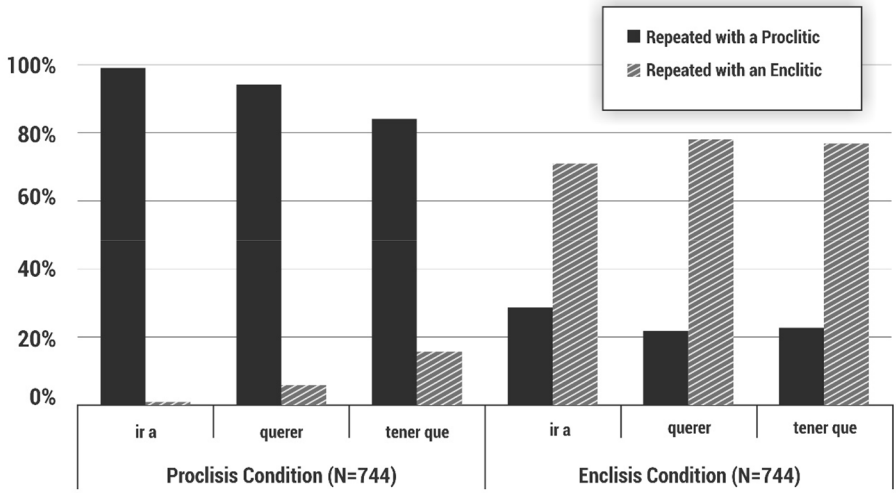

Figure 6. Clitic placement in children's imitations by condition (enclisis versus proclisis).

rightmost section of Figure 6 shows that when presented with enclisis, children moved clitics occurring with *ir a* to a preverbal position (i.e., forward repositioning) slightly more frequently than clitics occurring with the other verbs. Conversely, when presented with proclisis (see leftmost section of Figure 6), children almost never moved clitics with *ir a* to a postverbal positioning (i.e., backward repositioning). Backward repositioning was more frequent with *querer* and *tener que*.

To determine if differences in clitic repositioning by verb reached statistical significance, I ran a generalized linear mixed model (GLMM) with imitation of placement versus repositioning as the dependent variable. Since backward repositioning was only found in two trials with *ir a* in the dataset (and in fourteen trials with *querer* and thirty-nine trials with *tener que*, out of 248 responses with each verb), I excluded *ir a* from the analysis, since it was clear that this verb almost never prompted backward repositioning (in line with the strong tendency this verb shows toward proclisis). The analysis thus contrasted *querer* versus *tener que*. *Repositioning Type* (backward versus forward), *Finite Verb* (*querer, tener que*), and *Age Group* (4, 5, 6) were entered as explanatory variables. The random structure included *Participant* and *Trial*. All possible interactions were tested. The three-way interaction, as well as the interaction between *Age Group* and *Repositioning Type* were removed from the final model because they did not reach significance nor did they improve model fit.

The analysis found significant main effects of *Finite Verb* ($p < .001$) and *Repositioning Type* ($p = .003$), as well as significant interactions between *Finite Verb* and *Age Group* ($p = .039$) and *Finite Verb* and *Repositioning Type* ($p = .002$). The probability of any type of repositioning was greater with *tener que* (.178) than with *querer* (.106). Pairwise contrasts with Bonferroni adjustment indicate that the probability of repositioning between these two verbs was significantly different ($p = .000$). The probability of any type of repositioning with the two verbs tested (*querer* and *tener que*) was greater when forward repositioning was possible (i.e., in sentence originally presented

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

in enclisis; .215) than when backward repositioning was possible (i.e., in sentence originally presented in proclisis; .086), in line with overall preference for proclisis. Pairwise contrasts with Bonferroni adjustment indicate that the probability of repositioning was significantly different depending on the type of repositioning possible ($p = .003$).

Pairwise contrasts for the interaction between *Finite Verb* and *Age Group* indicated that the probability of repositioning the clitic in any direction was significantly greater for *tener que* than for *querer* only for the five-year-old group, whereas it was marginally significant ($p = .058$) for the four-year old group and not significant for the oldest group ($p = .367$) (see Figure 7). Pairwise contrasts for the interaction between *Finite Verb* and *Repositioning Type* indicated that the probability of repositioning the clitic was significantly greater for *tener que* than for *querer* only for backward repositioning (see Figure 8).

The statistical analysis of clitic repositioning in sentence repetition thus replicates the repositioning preference into proclisis, but it also reveals that children reposition *tener que* more than *querer* (especially the younger groups), and that the lexical difference in repositioning between these two verbs is most noticeable in backward repositioning.

At the individual level, most children were variable in their production of clitic placement, producing at least one token of both variants. Six participants (24, J22, P5, P6, P8, and P9) only produced accurate repetitions of proclisis in Study 3. Four of them (P5, P6, P8, and P9) belonged to the youngest age group and only two of them (24 and P6) had been fully categorical in Study 2, producing only proclisis. The rest of these categorical participants in Study 3 had either produced both clitic positions in Study 2 (P8 and P9) or produced categorical enclisis in Study 2 (P5 and J22), meaning that their grammar showed evidence of both variants. Therefore, for Studies 2 and 3 together, we find just two cases of regularization of one variant (proclisis) across the

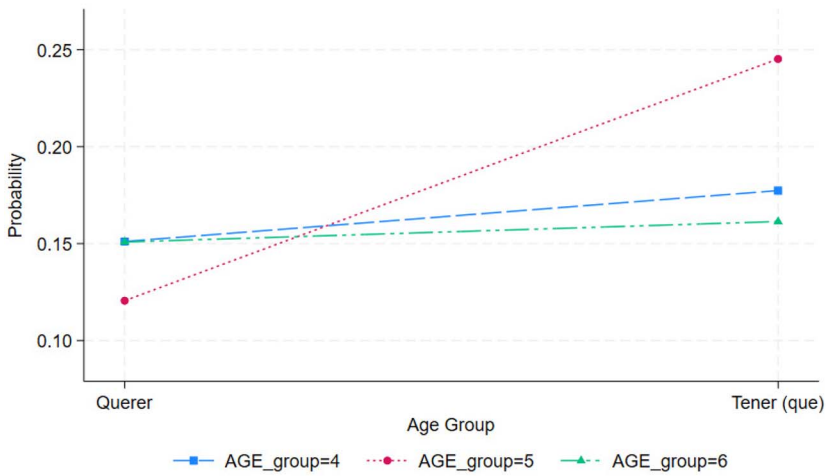

Figure 7. Significant interaction between Finite Verb and Age Group.

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

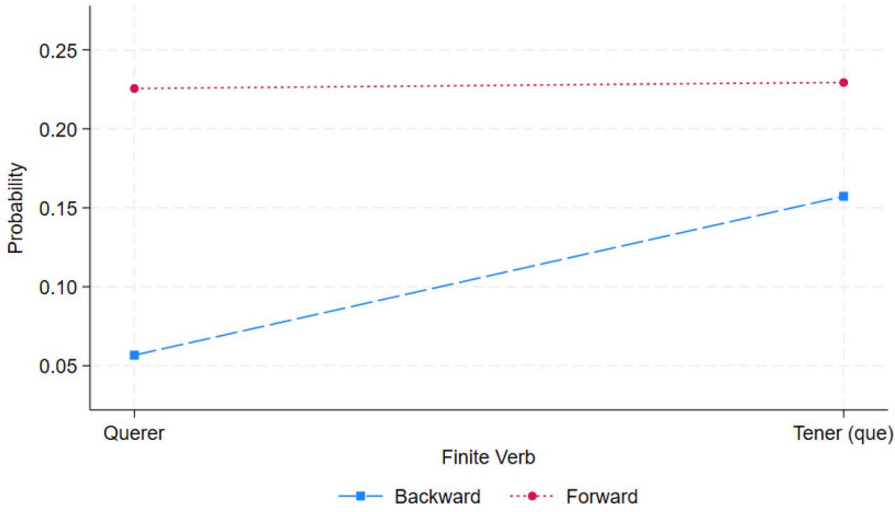

Figure 8. Significant interaction between Finite Verb and Repositioning Type.

board by children who only produced proclisis in Study 2 and only accurate repetitions of proclisis (as well as not backward repositioning errors) in Study 3.

I then examined individual child data emerging from both studies by verb. Taking categorical proclisis with *ir a* (the pattern that is expected to be the strongest given the use of proclisis in VCP overall and the clear bias of this verb toward proclisis), I found nine children who only produced proclisis in Study 2 and who repeated accurately only sentences containing proclisis in Study 3 with this verb. None of these children produced backward repositioning errors with *ir a* either. This means that 9/51 children displayed categorical proclisis with *ir a* across experimental studies. To examine whether these children may exhibit evidence in line with Step 2 (use of variants in nonoverlapping contexts, that is, only proclisis with *ir a* and only enclisis with the other verbs), I then looked at how these children used VCP with the two verbs that favor enclisis (*querer* and *tener que*). I did not find any of these nine children who used enclisis categorically across studies with either *querer* or *tener que*. In summary, the combined data from both experimental studies on the same children indicate that all but two children showed evidence that their grammar allows both clitic positions with the three verbs tested. This is expected of learners at Step 3 of the pathway by Shin and Miller (2022).

Discussion

Corpus and experimental data provide different windows through which we can observe children's patterns of syntactic variation. In the present study, all these data together have shed light on the later steps discussed in Shin and Miller's (2022) pathway proposal to the acquisition of morphosyntactic variation. Below I address the main questions that guided this study in turn.

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

With respect to the first research question concerning an initial phase of regularization, the results of Study 1 by age group indicate that, for the most part, children used both variants across verbs. Still, interesting cases were found. For example, both groups of children (ages 2;0-3;0 and 4;0-5;0) categorically used proclisis with the lexical construction [*estar*+gerund], which greatly disfavors enclisis in the input (4%). This behavior is in line with overextension of highly skewed distributions (Anderssen et al., 2010). Data from Studies 2 and 3 combined indicate that for most children both VCP variants (proclisis and enclisis) were available as part of their linguistic repertoire. I only found two children for whom enclisis was not attested at all. But in order to answer whether children use proclisis or enclisis in restricted rather than overlapping contexts, VCP by verb was examined.

Individual analyses of experimental data show that 18% (9/51) of the child participants (evenly distributed across the three age subgroups) used proclisis categorically with *ir a* across both experimental studies. Of those individuals, three also used proclisis categorically with the other verbs: *querer* (24, P6, and P9) and *tener que* (24 and P6). Again, in line with findings in Anderssen and colleagues (2010), this suggests that the strong skewness toward proclisis in a particular context in the input (adults produced 89% and 92% proclisis with *ir a* in Studies 1 and 2, respectively) may result in regularization of that variant in that context, albeit in a limited number of children. This behavior constitutes weak evidence for categorical use of variants in restricted lexical contexts, highlighting some individual variation among the children. For most children, however, both variants were attested with the verbs tested either within or across studies. Therefore, in the case of lexically restricted syntactic variation found in VCP, only a small subset of the children extended the high distribution of proclisis toward categorical use. Most of the children used both variants in overlapping contexts across studies, suggestive of a more advanced stage in the acquisition of this variation (Step 3 in Shin & Miller [2022]).

The use of different methodologies to investigate child acquisition of syntactic variation is an innovation of this study stemming from the realization that investigating variable use of infrequent syntactic structures among children may require triangulation to avoid misconstruing children's grammatical knowledge (see Requena, forthcoming). Most existing child language corpora are limited in size and do not allow researchers to address individual production due to insufficient data (as happens in many studies addressing the acquisition of morphosyntactic variation). This creates a tension between the gold standard methodology in variationist sociolinguistics (some type of naturalistic production) and the goals of language acquisition (to describe how acquisition proceeds in the individual learner). I encountered those limitations in Study 1, where analyses of the use of variants with each of the verbs *by individual child* were not possible due to scarce data. Such level of granularity is not feasible for most acquisition studies of morphosyntactic variables using the type of corpus data available. To study how children acquire grammatical variation and to test specific predictions at the individual level—such as those stemming from Shin and Miller's pathway—requires as much evidence as we can get. Therefore, in the absence of denser corpora, the present study resorted to methodological triangulation to address questions about what variants are part of the child's knowledge of Spanish VCP with particular verbs. Controlled designs can successfully reveal the variants that are part of

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

the child's internal grammar. Far from rote repetition, elicited imitation requires that children form syntactic and semantic representations (Crain & Thornton, 1998:76) and allows the researcher to precisely identify the target structure that the child is attempting to use (Lust, Flynn, & Foley, 1996:67). Therefore, as a methodology, sentence repetition taps into the child's grammar by eliciting specific target structures and thus revealing which variants are part of the child's language competence and whether the child's grammar allows for both variants to occur in a particular context.

With respect to the second research question on whether VCP becomes target-like with some lexical constructions before others, the results from Study 1 revealed that, when taken as communities of speakers, both younger and older child groups display differential use of variants according to finite verb and that this use approximates (for the most part) the distributions in the input. Very young children distinguish enclisis-from proclisis-favoring verbs. But observation of Figure 3 suggests differences by particular verbs and a process of fine-tuning of lexical preferences that may take place through the preschool years. Evidence from the verb *tener que*, albeit with few datapoints, provides some indication that by ages 4-5 the preference of use of VCP with this verb seems to shift in the direction of the input. A follow-up analysis of *ir a* in Study 2 also found that the probability of producing enclisis is significantly lower for the oldest group compared to the youngest groups, possibly signaling fine-tuning with these high-frequency verbs (see Shin, 2021).

To the best of my knowledge, processes of fine-tuning in the acquisition of constraints on variable use are not explicitly addressed in Shin and Miller's proposed pathway, which focuses mostly on the emergence of variants. Assuming that fine-tuning of sociolinguistic patterns led by high-frequency items is common in L1 acquisition of variation, it could begin to operate concurrently with the emergence of both variants in some overlapping contexts (Step 3 in Shin & Miller's proposal). Alternatively, Step 4 could be reformulated as consisting of fine-tuning of variable use across levels of a given conditioning constraint (or factors in a factor group) to match the input. For example, as shown in Figure 5, children seem to engage in a subtle process of fine-tuning that increasingly disfavors enclisis more and more with *ir a*. With *querer* and *tener que*, Study 2 does not reveal a clear pattern of fine-tuning. These two verbs are clearly enclisis-favoring (a finding also attested in younger children in Study 1). The clitic backward repositioning patterns found in Study 3 suggest an already attained greater level of fine-tuning by which the same children may be associating enclisis more strongly with *tener que* than with *querer*, which is in line with corpus research with adult speakers (e.g., Davies, 1995). The present study finds very early acquisition of most lexical biases but suggests that fine-tuning may be also at play during this process of acquisition.

A surprising difference between child and caregiver speech was young children's initial overall preference for proclisis with the *tener que* in Study 1 (compared to older children and caregivers). This could indicate something special about *tener que*. This observation is based on very few tokens and could not be confirmed statistically at the individual verb level, so future research should corroborate this finding. It is possible, however, that this observation could reveal children's attempt to learn how VCP patterns with a finite verb that is grammaticalized and relatively frequent (features that characterize proclisis-favoring verbs) but that favors enclisis in the input (see Requena,

https://doi.org/10.1017/S0954394523000248 Published online by Cambridge University Press

2020). Future studies may reveal how young children navigate acquisition of lexically conditioned variation for items that could provide conflicting evidence in the input.

Conclusion

The present study contributed novel data about child language acquisition of syntactic variation. In particular, through the combination of corpus and experimental techniques, I examined the acquisition of Spanish variable clitic placement (VCP), a low frequency syntactic variable that is lexically conditioned in naturalistic production among adults. Analyses of corpus data from the very first stages of language development indicate early knowledge of this lexical conditioning. By the beginning of the school years, children distinguish the three verbs that were tested across studies based on their particular patterns of clitic placement. Categorical clitic placement found in individual tasks was put into perspective through methodological triangulation, which revealed a fuller picture of children's variable grammar. While few individual children displayed signs of overextension, the grammar of almost all the children tested allowed both variants with each of the frequent verbs examined here (as expected from Step 3 of Shin & Miller's [2022] pathway). But beyond variable use in overlapping contexts, the data suggested a process of fine-tuning by which very young children seem to adjust rates of enclisis with particular finite verbs. While the possibility of fine-tuning and its place in developmental pathways should be investigated in future research, the present results constitute evidence for very early acquisition of syntactic variation in the input that is lexically conditioned. Furthermore, the findings point to ways in which children may arrive at such detailed knowledge of subtle patterns of language use in their input. Research on syntactic variables that are not lexically conditioned would shed light on the extent to which probabilistic associations between lexical items and syntactic variants are responsible for the very early acquisition reported here.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0954394523000248.

Acknowledgments. The work presented here was funded by Penn State University dissertation awards (STAR, RGSO, and Center for Global Studies) as well as a *Language Learning* Dissertation Grant. The author would like to thank Hannah Forsythe, Ana Ferrer, Victoria Tissera, and Victoria Bognanno for their help with data extraction/transcription and coding. Also, he would like to thank the administrators, teachers, students, and parents at Colegio William C. Morris, Colegio 25 de Mayo, and Colegio Santo Tomás for their assistance and participation.

Competing interests. The author declares none.

Notes

1. VCP also occurs with indirect object clitic pronouns. However, the focus of this study will be 3sg direct object pronouns.

2. [*Estar*+gerund] is not included in Figure 1 because Davies (1995) only looked at VCP in infinitival constructions.

3. The Mexican Child-Caregiver Corpus contains approximately 125 hours of conversational interactions between twenty-five Mexican children and their caregivers. The recordings were collected from May to July 2008, and they are part of a larger collaborative project that compares acquisition of grammatical morphology in contexts of variable input.

4. Data in Figure 2 are organized by caregiver's rate of enclisis in decreasing order.

5. With respect to the verb *ir a*, where descriptives suggested very high use of enclisis by the youngest children, pairwise comparisons using Bonferroni adjustment indicated no significant difference with adults (Contrast estimate = .278, *SE* = .140, *p* = .280, 95% CI = –.092, .648).

6. 30/51 children were categorical in Study 2. Fourteen of these children produced five tokens or more. Of these, 8/14 produced proclisis categorically, and 5/14 produced enclisis categorically. This finding begs the question about whether some children may be at a stage of clitic placement regularization across the board. Evidence form Study 3 will show this is not the case and will provide more clarity about lexical preferences.

References

Anderssen, Merete, Bentzen, Kristine, Rodina, Yulia, & Westergaard, Marit. (2010). The acquisition of apparent optionality: Word order in subject and object shift constructions in Norwegian. In M. Anderssen, K. Bentzen, & M. Westergaard (eds.), *Variation in the Input*. Springer: Dordrecht. 241–270.

Anderssen, Merete, & Westergaard, Marit. (2010). Frequency and economy in the acquisition of variable word order. *Lingua* 120(11):2569–2588.

Brown, Roger. (1973). *A first language: The early stages*. Cambridge MA: Harvard University Press.

Bybee, Joan. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.

Crain, Stephen, & Thornton, Rosalind. (1998). *Investigations in Universal Grammar: A guide to experiments on the acquisition of syntax and semantics*. Cambridge MA: MIT Press.

Davies, Mark. (1995). Analyzing syntactic variation with computer-based corpora: The case of modern Spanish clitic climbing. *Hispania* 78(2):370–380.

Eisenchlas, Susana. (2003). Clitics in child Spanish. *First Language* 23(2):193–211.

Guy, Gregory R. (1980). Variation in the group and the individual: The case of final stop deletion. In W. Labov (ed.), *Locating language in time and space*. New York: Academic. 1–36.

Hendricks, Alison E., Miller, Karen, and Jackson, Carrie. (2018). Regularizing unpredictable variation: Evidence from a naturalistic setting. *Language, Learning and Development* 14:42–60.

IBM Corp. (2021). IBM SPSS Statistics for Windows, Version 28.0. Armonk: IBM Corp.

Kerswill, Paul. (1996). Children, adolescents, and language change. *Language Variation and Change* 8(2):177–202.

Kovac, Ceil, & Adamson, Douglas H. (1981). Variation theory and first language acquisition. In D. Sankoff & H. Cedergren (eds.), *Variation omnibus*. Edmonton: Linguistic Research, Inc. 403–410.

Labov, William. (1989). The child as linguistic historian. *Language Variation and Change* 1(1):85–97.

Lust, Barbara, Flynn, Suzanne, & Foley, Claire. (1996). What children know about what they say: Elicited imitation as a research method for assessing children's syntax. In D. MacDaniel, C. McKee, & H. S. Cairns (eds.). *Methods for assessing children's syntax*. Massachussets: MIT Press. 55–76.

MacWhinney, Brian. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah NJ: Lawrence Erlbaum Associates.

Miller, Karen. (2013). Acquisition of variable rules: /S/-lenition in the speech of Chilean Spanish-speaking children and their caregivers. *Language Variation and Change* 25:311–340.

Miller, Karen. (2015). Children's production of ain't. In P. Donaher & S. Katz (eds.), *Ain'thology: The history and life of the taboo word ain't*. Cambridge: Cambridge Scholars Publishing. 96–112.

Miller, Karen, & Schmitt, Cristina. (2012). Variable input and the acquisition of plural morphology. *Language Acquisition: A Journal of Developmental Linguistics* 19(3):223–261.

Myhill, John. (1988a). The grammaticalization of auxiliaries: Spanish clitic climbing. *Berkeley Linguistics Society* 14:352–363.

Myhill, John. (1988b). Variation in Spanish clitic climbing. In T. J. Walsh (ed.), *Synchronic and diachronic approaches to linguistic variation and change (GURT '88)*. Washington: Georgetown University Press. 227–250.

Pérez-Leroux, Ana Teresa, Cuza, Alejandro, & Thomas, Danielle. (2011). Clitic placement in Spanish–English bilingual children. *Bilingualism: Language and Cognition* 14(2):221–232.

Requena, Pablo E. (2020). A usage-based account of Spanish variable clitic placement. *Languages* 5(3):33.

Requena, Pablo E. (forthcoming). Methodological triangulation in the study of acquisition of morphosyntactic variation *Language Learning and Development*.

Rodríguez Mondoñedo, Miguel, Snyder, William, & Sugisaki, Koji. (2004). Clitic-climbing in child Spanish and the theory of parameters. *Poster presented at the* Boston University Conference on Language Development *(BUCLD29).* Boston, Massachusetts, November 5-7. Retrieved from http://faculty.human.mie-u.ac.jp/~sugisaki/Papers/Rodriguez-Mondonedo_et_al_2005_BUCLD.pdf.

Saldana, Carmen, Smith, Kenny, Kirby, Simon, & Culbertson, Jennifer (2021). Is regularization uniform across linguistic levels? Comparing learning and production of unconditioned probabilistic variation in morphology and word order. *Language Learning and Development* 17(2):158–188.

Schwenter, Scott, & Torres Cacoullos, Rena. (2014). Competing constraints on the variable placement of direct object clitics in Mexico City Spanish. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics* 27(2):514–536.

Shin, Naomi. (2016). Acquiring patterns of morphosyntactic variation: Children's Spanish subject pronoun expression. *Journal of Child Language* 43(4):914–947.

Shin, Naomi L. (2021). Acquiring constraints on variable morphosyntax: SV-VS word order in child Spanish. In M. Díaz-Campos (ed.), *Handbook of variationist approaches to Spanish*. New York: Routledge.

Shin, Naomi, & Miller, Karen. (2022). Children's acquisition of morphosyntactic variation. *Language Learning and Development* 18(2):125–150.

Shin, Naomi, Requena, Pablo E., & Kemp, Anita. (2017). Bilingual and monolingual children's patterns of syntactic variation: Variable clitic placement in Spanish. In A. Auza & R. Schwartz (eds.), *From typical language development to language disorders in Spanish-speaking children: Language processing and cognitive functions*. Dordrecht: Springer. 63–88.

Smith, Jennifer, & Durham, Mercedes. (2019). *Sociolinguistic variation in children's language: Acquiring community norms*. Cambridge: Cambridge University Press.

Smith, Jennifer, Durham, Mercedes, & Fortune, Liane. (2007). Mam, my trousers is fa'in doon!: Community, caregiver, and child in the acquisition of variation in a Scottish dialect. *Language Variation and Change* 19:63–99.

Thomas, Danielle L. T. (2012). *Grammatical optionality and variability in bilingualism: How Spanish-English bilinguals limit clitic-climbing*. Doctoral dissertation, University of Toronto.

Washington, Julie A., & Craig, Holly K. (2002). Morphosyntactic forms of African American English used by young children and their caregivers. *Applied Psycholinguistics* 23:209–231.

Cite this article: Requena PE (2024). Early acquisition of syntactic variation: Lexical conditioning of Spanish variable clitic placement. *Language Variation and Change* **36**, 1–24. https://doi.org/10.1017/S0954394523000248