

The ‘Big DP’ Hypothesis: New Evidence from Gender Agreement in a Shipibo-Spanish Language Contact Situation*

Liliana Sánchez

José Camacho

University of Illinois Chicago
sanchez.l.e.2000@gmail.com
jcamach@uic.edu

Elisabeth Mayer

Australian National University
elisabeth.mayer@anu.edu.au

Carolina Rodríguez Alzza

University of Texas at Austin
rodriguez.alzza@utexas.edu



Received: August 26, 2021

Accepted: August 11, 2022

Abstract

Previous evidence suggests clitics and determiners do not enter the same type of gender agreement relations in contact varieties of Spanish, despite proposals that treat clitics as determiners (big DP – Determiner Phrase – hypothesis) (Uriagereka 1995; Cecchetto 1999, 2000; Belletti 2005). We conducted a study on gender agreement between clitics and their antecedents, and determiners and nouns among adult Shipibo-Spanish bilinguals to answer the following question: Do Shipibo-Spanish bilingual speakers have similar patterns of local vs. long-distance gender agreement? Our results show gender agreement between D and N obtains categorically, but gender agreement between the clitic and its antecedent is at chance-level. We propose an alternative analysis for our data that does not assume the big DP hypothesis.

Keywords: big DP analysis; clitics; determiner-noun agreement; gender; language contact; Shipibo; Spanish

Resum. *La hipòtesi del SD gros: les dades de concordança de gènere en una situació de contacte de llengües shipibo-espanyol*

Els treballs anteriors suggereixen que els clítics i els determinants no entren en el mateix tipus de relacions de concordança de gènere en varietats de contacte de l'espanyol, tot i les propostes que tracten els clítics com a determinants (hipòtesi del SD gros) (Uriagereka 1995; Cecchetto 1999,

* We would like to thank Gesica Pérez for her work as a research assistant in this project, the participants in this study for entrusting us with their linguistic knowledge, and the Patria Nueva native Community for making this research possible. We would also like to thank the anonymous reviewers and the editors, whose comments have greatly improved the paper.

2000; Belletti 2005). Hem realitzat un estudi sobre la concordança de gènere entre els clítics i els seus antecedents, i els determinants i els noms entre els bilingües adults shipibo-espanyol per respondre a la pregunta: Els parlants bilingües shipibo-espanyol tenen patrons similars de concordança de gènere local vs. concordança de gènere a llarga distància? Els nostres resultats mostren que la concordança de gènere entre D i N es fa categòricament, però la concordança de gènere entre el clíctic i el seu antecedent es produeix de l'atzar. Proposem una anàlisi alternativa per a les nostres dades que no assumeixi la hipòtesi del SD gros.

Paraules clau: anàlisi del SD gros; clítics; concordança entre determinant i nom; gènere; contacte de llengües; shipibo; espanyol

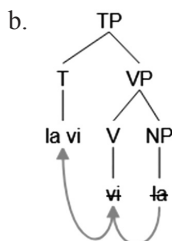
Table of Contents

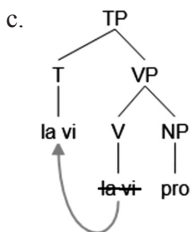
1. Introduction	5. The study
2. The big DP analysis	6. Results
3. Clitic gender agreement and DP-internal gender agreement in Contact Bilingualism	7. Discussion
4. Shipibo language	8. Conclusions
	References

1. Introduction

Clitics in Romance languages have been traditionally analyzed either as generated in argument position (Kayne 1975; Quicoli 1980; etc.) or as generated as inflectional heads matching or identifying the true argument respectively (Lenz 1920; Givón 1976; Jaeggli 1982; Suñer 1988; Sportiche 1996). In the first type of analysis, clitics like the one in example (1a) move from the argumental position to adjoin to the verb, and subsequently raise together with the verb to T/INFL, as schematized in (1b). In the second type of analysis, clitics are generated with the verb, or perhaps in a higher functional head, and subsequently, raise with the verb to T/INFL. Importantly, the argumental position is occupied by a null category, schematized in (1c) as *pro*.

- (1) a. L-a vi.
 CL-F see.1.SG.PST
 'I saw her.'





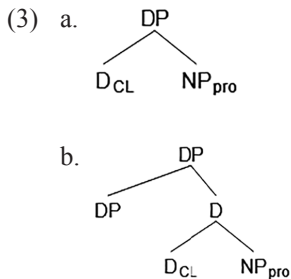
These illustrations are highly schematic, and they gloss over many important differences between the two analyses, for example, whether the initial movement of the clitic in (1b) targets a larger XP, with the subsequent movement of the clitic as a head (cf. Kayne 1989). Still, they provide a general idea of the main difference of interest for this study.

One instantiation of this proposal is the idea that clitics are generated as determiners of a “big DP” (Uriagereka 1995; Cecchetto 2000; Belletti 2005). This analysis builds on the phonological similarity between third-person clitics and determiners, as shown in Table 1, by proposing that clitics are definite determiners.

Table 1. Parallelism between determiners and clitics

	Determiner	Clitic
Sg.	el (M) /la (F)	lo (M.ACC) / la (F.ACC) / le (DAT)
Pl.	los (M) / las (F)	los (M.ACC) / las (F.ACC) / les (DAT)

The proposed structure for (1a) is shown in (3a), which includes a null NP as a complement; in the case of clitic-doubled arguments, the proposed structure is in (3b). In this last structure, the doubled element is in its specifier, corresponding to clitic-doubling structures.



In this paper, we present a novel exploratory study that suggests a way to experimentally test the “big DP” hypothesis in the Spanish grammar of bilingual speakers of Spanish in contact with Shipibo. Since Shipibo lacks pronominal clitics and gender, we compare gender agreement patterns inside traditional DPs (“local DP agreement”), with gender agreement between a clitic and *pro* (“long-distance agreement”), as shown in (4). If third-person clitics are determiners in the Spanish of these speakers, we would expect similar agreement patterns in (4a) and (4b). In contrast, dissimilar patterns would be more consistent with alternative analyses that do not equate clitics with definite determiners.¹

- (4) a. DP-internal (**local**) agreement b. Clitic-*pro* agreement
- | | |
|---|--|
| $\left[{}_{\text{DP}} \text{la}_F \left[{}_{\text{NP}} \text{casa}_F \right] \right]$ <div style="text-align: center; margin-top: 5px;"> </div> | $\left[{}_{\text{DP}} \text{la}_F \left[{}_{\text{NP}} \text{pro}_F \right] \right]$ <div style="text-align: center; margin-top: 5px;"> </div> |
|---|--|

Since *pro* is a null element, in order to assess the gender agreement between the clitic and *pro*, we target the referent or discursive antecedent of *pro*.

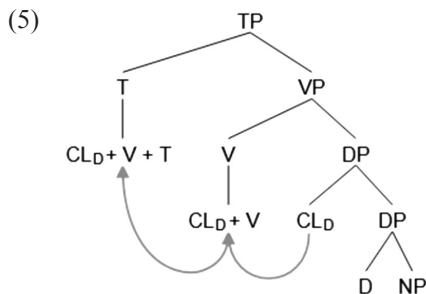
From a second/bilingual language acquisition perspective, gender is known to be a divergent property (Alarcón 2011; Keating 2010). The study of Shipibo L1-Spanish L1/L2 speakers allows us to test the acquisition of local vs. long-distance clitic agreement in bilinguals, assuming some level of feature reassembly of gender takes place in Spanish (Franceschina 2005; Lardiere 2017) due to contact with Shipibo. We first determine whether gender is used to generate masculine and feminine noun classes in their Spanish as evidenced by local agreement, and we then test whether long-distance clitic gender agreement is operative. On the basis of the results of two experimental tasks administered to L1 and L2 Spanish speakers in a Shipibo-Spanish contact situation, we propose different analyses for determiners and clitics. In particular, in the case of L2 Spanish speakers that exhibit gender mismatches between clitics and their discourse antecedents, we propose that clitics are argument markers not specified for gender (Sánchez et. al 2022). For those who exhibit gender matching we propose an analysis that allows for feature matching between the antecedent and the clitic rather than feature valuation at the sentential level.

2. The big DP analysis

Uriagereka (1995) proposed an analysis of clitic doubled expressions in which the clitic and the doubled DP formed part of the same constituent, as shown in (3) above. Cechetto (2000) and Belletti (2005) extend this analysis to unify cases of clitic doubling, clitic left-dislocation and floated quantifiers. The common idea behind all these constructions is that the clitic coexists with doubled elements because they initially form a larger constituent, a “big DP”, which is headed by the

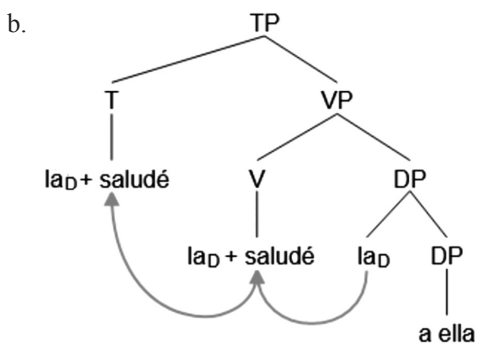
1. An anonymous reviewer points out that there is a possibility that agreement between the clitic and the DP may not be local if it takes place after movement. We are testing the big DP hypothesis under the assumption that agreement is local.

clitic. In Belletti's (2005) formulation, the doubled element appears in the complement position to the clitic determiner, as seen in (5). The clitic raises from the big DP and adjoins to V, together they raise to TP, as (5) shows.

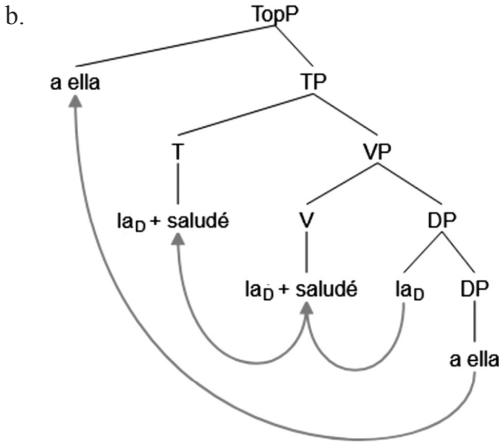


This analysis allows for the derivation of expressions in which a clitic and a doubled DP co-occur in a sentence from a single initial structure. In the simple case, this doubled DP is null and only the clitic appears. In more complex cases, the big DP hypothesis can also account for instances where the clitic is doubled by an overt DP. Specifically, it provides a unified analysis for both clitic doubling structures, as shown in (6) and clitic left dislocations structures, shown in (7).

- (6) a. L-a saludé a ella.
 CL-F greeted DOM her
 'I greeted her.'



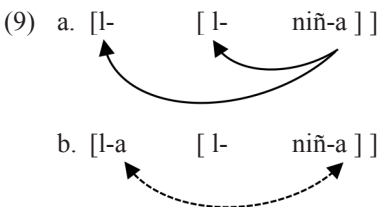
- (7) a. A ella, l-a saludé.
 DOM her, CL-F greeted
 ‘Her, I greeted.’



In addition to providing a unified analysis, the big DP Hypothesis has an advantage over other analyses of clitic doubling expressions that treat the doubled DP as an adjunct (Aoun 1981) which cannot easily account for the fact that in most Latin American Spanish varieties that allow for extended doubling of third-person direct objects, the clitic and the DP agree in gender and number (Jaeggli 1982; Suñer 1988; Di Tullio, Saab & Zdrojewski 2019; among others), as shown in

- (8) L-a vio a l-a niñ-a.
 CL-F saw DOM the-F girl-F
 ‘H/she saw the girl.’

In these cases, the big DP headed by clitic *la* is the local nominal domain in which gender and number agreement is not only possible but expected, just as in any other DP (cf. *las niñas* the-F-P girl-F-P). Technically, we can imagine two possible implementations of this local agreement: either N is marked for valued gender and agrees with all other categories within the DP that are unvalued for gender (clitic *la* and determiner *la*), as shown in (9a), or the clitic itself is valued as feminine, as well as N, as seen in (9b). Both values would then have to match.



3. Clitic gender agreement and DP-internal gender agreement in Contact Bilingualism

Gender expression in contact bilingualism generally exhibits high levels of variability cross-linguistically (Mayer & Sánchez 2017; Laleko 2018; Lohndal & Westergaard 2016). Previous studies of Spanish in contact with languages that lack gender have found a wide range of gender agreement (and lack thereof) options for clitics. These range from the use of invariable forms such as *le* or *lo*, unmarked for gender (Mayer & Sánchez 2017) to gender-specific forms such as *lo* (masculine) and *la* (feminine) for direct object clitics (Camacho & Sánchez 2002; Mayer 2010, 2017; Mayer & Sánchez 2016, 2017). In this section, we will present some previous studies conducted among Spanish-English heritage bilingual children and speakers of Spanish in contact with two indigenous languages of Peru, Quechua, and Shipibo. In all of these cases, Spanish is in contact with a language that lacks grammatical gender agreement between determiners and nouns, as illustrated by the contrast between Spanish and English shown in (12):

- (12) The house / L-a cas-a
 DEF-F.SG house-F.SG

Regarding gender assignment to pronouns, there are some differences. In the case of English, pronouns referring to animate entities show gender assignment and may refer to lexically gendered antecedents, as shown in example (13). As we will show below, in languages such as Shipibo, unlike in English, third-person pronouns are not lexically marked for gender.

- (13) John and *Mary* entered the room. I greeted *her*.

In a study on bilingual effects, Castilla-Earls et al. (2020) examined the performance of two groups of Spanish-English bilingual children (one with high and one with low English proficiency) in an oral task aimed to elicit sentences with clitics and DPs with definite determiners and nouns. The study's goal was to determine the accuracy of both groups regarding the overt production of clitics and determiners, in particular, their accuracy in gender, and number marking. The stimuli they used included questions such as (14a) to elicit clitics according to the antecedent in the question (14b), and (15a) to elicit DPs as in (15b):

- (14) a. ¿Qué hace el ratón con el queso?
 ‘What does the mouse do with the cheese?’
 b. Se l-o come.
 ASP DEF-M.SG eat.3.SG
 ‘(He/she) eats it.’

- (15) a. ¿A quién les lee la gallina?
 'To whom the chicken reads to?'
 b. A I-o-s pollit-o-s.
 DOM DEF-M-PL chicken-M-PL
 'To the chickens.'

Their results indicate that the bilingual children were overall more accurate with determiners than with clitics. Furthermore, they found the errors were minimal among a comparison group of children monolingual in Spanish. We take this as a first indication that bilingual individuals may exhibit some difference between determiners and clitics.

Shin et al. (2019) conducted a study on the production and comprehension of direct object clitics among child heritage speakers of Spanish. Their results showed that higher English vocabulary scores and more use of English in the home were related to higher frequencies of gender mismatches between the clitic and its antecedent in the production task they used, similar to the one used by Castilla-Earls et al. (2020) based on Grüter et al. (2012). We take this to indicate that there is an effect of crosslinguistic influence from English on gender agreement between the clitic and its antecedent. Furthermore, the results of their comprehension task (a picture selection task) showed evidence that the children did not attend to the clitic gender to select a referent.

Mayer & Sánchez (2017) analyzed the oral production of 17 adult Cuzco Quechua-Spanish bilinguals, 14 Huanuco Quechua-Spanish bilinguals, 24 Shipibo-Spanish bilinguals and three monolingual speakers of Spanish from the city of Lima. In this study, Mayer and Sánchez analyzed sentences with direct object clitics with and without doubling of a DP.³ Their findings showed a strong preference for the clitic *le* (unmarked for gender) over the masculine clitic *lo* among the three bilingual groups, but not among the monolingual contact individuals. Of particular interest is the hierarchy in the distribution of the clitics present in the Shipibo-Spanish group, where *le* was by far the most frequent direct object form, but not to the total exclusion of *lo*, *se*, and *la*, the latter two absent in the Quechua-Spanish groups:

- (16) $le > se > lo > la$
 $74\% > 7\% > 4\% > 3\%$

Mayer & Sánchez (2017) interpret this preference as an indication of variability in how features are mapped onto morphological forms such that these range from *le* as a form unmarked for gender to a more complex array of morphological forms available to the speakers including *lo* and *la* which indicates the emergence of gender marking.

While the strong preference for *le* may be taken as evidence of some type of neutralization in the mapping of gender features, it could also be attributed to dif-

3. Their study also focused on differential object marking on the doubled DP.

faculties in production. Unlike Shin et al.'s study, Mayer & Sánchez's (2017) study only had a production task leaving open the question of the extent to which the Shipibo group would be able to attend to differences in clitic gender in a comprehension task. Moreover, in contrast to Castilla-Earls et al.'s (2020) study, Mayer & Sánchez (2017) examined clitic production and gender matching with the discursive antecedents but it did not examine determiner-noun agreement. In this study, we examine comprehension data on gender (mis)matches between clitics and referents and we also provide data on determiner-noun agreement in a group of Shipibo-Spanish bilingual speakers in order to provide an answer to the question of whether the big DP analysis can be extended to adult Spanish in contact with Shipibo.

Before discussing the study, we present in the next section a brief sketch of some of the grammatical characteristics of Shipibo that are most relevant to this study.

4. Shipibo language

The Panoan language family is located in the Amazon basin and extends to Peru, Brazil and Bolivia. There are approximately 32 known Panoan languages (Fleck 2013), and one of these is Shipibo-Konibo or simply Shipibo. According to Valenzuela (2003), Shipibo is a predominantly agglutinating and suffixing language. Regarding gender, Shipibo lacks gender marking on pronouns, morphological gender marking, and gender agreement inside the Determiner Phrase (DP), as shown in (17), (18) and (19).

(17) Ja-ra jo-ke.
 3P:ABS-EVID come-ASP
 '(I have direct evidence that) he/she has come.'
 (Faust 1990: 7, our glosses)

(18) Bake ainbo /bake benbo
 child woman child man
 'daughter' 'son'
 (Valenzuela 2003)

(19) Maxko-shoko joni
 small-DIM person.ABS
 'small person'
 (Valenzuela 2003)

The sentence/clause mainly follows a basic order SOV/SV. Moreover, Shipibo presents an ergative-absolutive alignment with an ergative case marked by an affix at the end of the noun phrase in subject position and variable marking of absolutive case of direct objects. A salient characteristic is the variety of ergative case allomorphs: *-n*, *an*, *-en*, *-in*, *-kan*, *-ten*, *-tan*, *-man*, *-nin*, *-ton*, *-tonin* (Valenzuela 2010). In addition, this Panoan language also shows some other clitics as evidentials (see Valenzuela 2003). Unlike Spanish, Shipibo lacks pronominal clitics, as

illustrated in (20b) with a null object that corresponds to *tasa* ‘basket’, the direct object in (20a)⁴:

- (20) a. E-n-ra tasa jama-ke.
 1SG-ERG-EVID basket kick-COMPL
 ‘I kicked the basket.’
- b. E-n-ra jama-ke.
 1SG-ERG-EVID kick-COMPL
 ‘I kicked.’

5. The study

Given the previous evidence that suggests clitics and determiners do not enter the same type of gender agreement relations in contact varieties of Spanish, we conducted a study on gender agreement between clitics and their referents, and determiners and nouns among adult Shipibo-Spanish bilinguals. Our main goal was to answer the question: Do Shipibo-Spanish bilingual speakers have similar patterns of local vs. long-distance gender agreement? A positive answer would be consistent with the ‘big DP’ analysis. A negative answer, on the other hand, would be inconsistent with such analysis for this variety of Spanish in contact with Shipibo. In the next subsection, we present a description of the Shipibo community where we conducted the study, the participants, and the methods employed.

5.1. *The context of contact: The Shipibo-Konibo people at Patria Nueva*

In this section, we present information on the contact situation of the participants in this study which is vital in our understanding of how language contact affects language change. It also serves the purpose of documenting the emergence of the varieties of Spanish in this study.

The Shipibo-Konibo are one of the largest indigenous groups in the Peruvian Amazon region, with approximately 25,232 members (INEI 2017). These indigenous people are widespread along the Ucayali river. However, over the last decades, Shipibo-Konibo migration to urban cities has intensified. Nowadays, the cities of Pucallpa and Lima congregate one of the largest and most prominent Shipibo settlements. The continuity between the rural and urban areas has increased the locations and social network contributing to the development of bilingualism in Shipibo-Konibo and Spanish and the emergence of new varieties of Shipibo-Spanish.

4. An anonymous reviewer suggests that the distribution of non-nominal clitics in Shipibo might affect the use of nominal clitics in Spanish, for example “without movement or agreement”. The paper is, in some sense, an exploration of the distribution of nominal clitics in L1 Shipibo Spanish. However, the notion that second position clitics in Shipibo may affect the distribution of nominal clitics in Spanish presupposes that clitics form a natural class defined by properties that go beyond being syntactically and prosodically weak. While this is certainly possible, we do not assume that this is the default theoretical position.

The Native Community Patria Nueva de Mediación Callería is located close to the Callería river, one of the Ucayali's tributaries (Callería district, Coronel Portillo province, Ucayali region). In 1983, the Peruvian State recognized this territory and granted a property title of 723,8600 hectares, which was expanded in 1997 (IBC 2016: 129). According to Brabec de Mori (2015: 733), Patria Nueva was probably founded in 1972 with two or three houses of first settlers. Its growth in later years is associated with the establishment of the primary school taught mostly in Spanish which led to more Shipibo arriving from the Upper Ucayali area (Abel Mori, first authority in the community. Personal communication, 2019). Nowadays, the community also has a pre-school and a secondary school, with a Bilingual Intercultural Education program currently being developed with participation of Shipibo teachers.

The population in Patria Nueva, estimated at 271 in 2015, is mostly composed of Shipibo-Konibo people (Brabec de Mori 2015: 733) and to a lesser extent of individuals who the Shipibo call *nawa* (*mestizos* in Spanish). These are usually non-Shipibo men married to Shipibo women. Some *mestizos* working as teachers or health workers have a stable presence in the community. Other *mestizos* come to the community for seasonal fishing or sporadic tourism. There are also *mestizos* from evangelical churches participating in communal activities led by a Shipibo pastor.

Despite the relatively large population, it is almost impossible to precisely know how many people are living in Patria Nueva. This is mainly due to the population's high mobility within the Callería river basin in pursuit of fishing and agriculture. Additionally, the Shipibo-Konibo inhabitants of Patria Nueva regularly visit nearby cities in order to provide themselves with food to complement their diet based on the products they grow, to visit a relative or to settle temporarily. This results in their frequent contact with Spanish inside and outside the community.

5.2. Participants

Twenty-four Shipibo-Spanish bilinguals from Patria Nueva with an age range between 21 and 79 completed the two tasks using Psychopy, in addition to a linguistic background questionnaire and a proficiency task as part of a larger study. Among the people interviewed, 18 reported Shipibo as their first language (Shipibo L1 henceforth). Most participants in this group were born in the same community where the study was carried out. In contrast, others indicated that they came from native communities of the Upper Ucayali river, but that they had already lived in Patria Nueva for many years. Another 6 participants were *mestizos* that came from cities in the Lower Ucayali river and reported Spanish as their first language (Spanish L1 henceforth).

Participants were mainly male, with 8 Shipibo-Konibo and 6 *mestizos*. All female participants (10) were Shipibo-Konibo. As mentioned above, in the Patria Nueva Native Community, there is currently no presence of *mestizo* females. Still, there are *mestizo* males who have married Shipibo women or carry out some

specific activity in the community. All participants were compensated for their participation.

5.3. Methods

Participants underwent several training rounds to become familiar with the computer used to administer the three tasks of the Psychopy program. These tasks included: a lexical activation pre-task, a gender agreement picture-selection task (inspired in Grüter et al. 2012, but culturally adapted), and a DP-internal Agreement task.⁵ The three tasks were administered in the community of Patria Nueva to each individual participant by one of the authors and a research assistant who is a speaker of Shipibo. In addition to the Psychopy tasks, participants responded orally to a questionnaire about their language use and preference and completed two other tasks not reported here. The interviews were conducted entirely in Spanish, although some indications had to be clarified in the Shipibo language for the first group of participants. This was possible thanks to the support of the research assistant.

5.3.1. Lexical Activation Pre-Task

To ensure participants understood the meaning of the nouns, they were first shown drawings of 30 objects and given a bare noun to name the object in Spanish, as shown in Figure 1.

13 pictures represented objects with nouns ending in feminine *-a*, 7 in masculine *-o*, 4 in a consonant (3 masculine, 1 feminine) and 2 in *-e* (both masculine). The asymmetric distribution by gender and morphology is due to the fact these data were collected as part of a larger comparison with Ashaninka, so we had to ensure that words would be equally valid across the three languages.



Figure 1. Canasta 'basket'.

5. The stimuli for the Lexical Activation Pre-task, the Picture-Selection Task and the DP-internal Agreement Task are publicly available at the Center for Open Science Database (<<https://osf.io/m28t/>>) as a power point document. The questionnaire is also available there.

5.3.2. Long-distance Gender Agreement Picture-Selection Task

The picture-selection task items involved a pair of drawings that were shown side by side, chosen from the words already presented in the lexical activation task. The images correspond to culturally relevant objects that are part of the everyday life of the participants in the study. The task was a forced-choice task that did not allow for a third option. Participants had to pick one of the images, following an audio prompt that included an expression with a verb in imperative form and a clitic. Psychopy tasks were designed to choose using the keyboard left or right arrow. The imperative with a clitic referred to a masculine (21a) or feminine (21b) noun. This indicated clitic-antecedent gender agreement patterns.

- (21) a. Coge-l-o. b. Coge-l-a.
 Take-CL-M.SG Take-CL-F.SG
 ‘Take it.’ ‘Take it.’

Figure 2 shows an example of an item from this task in which participants had to choose one image after hearing an imperative verb form with a clitic as in (21a) or b. In this particular item, if the participant heard (21a), it was expected they would select the image with the comb as the word *peine* ‘comb’ is masculine in Spanish. Conversely, if they heard (21), they were expected to select the image with the basket as the word *canasta* ‘basket’ is feminine in Spanish. Half of the items included a masculine *lo* clitic, and half a feminine *la*.

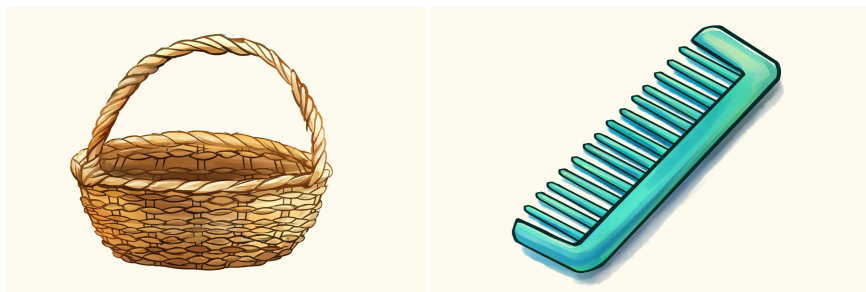


Figure 2. Long-distance Gender Agreement Picture-Selection Task.

5.3.3. DP-internal Gender Agreement Task

In the second task, participants saw 20 images from the same drawings they saw in the Lexical Activation Pre-Task. They were asked to provide a definite determiner for those objects orally. Figure 3 illustrates what the participants saw and what they were expected to produce.

The answers were annotated on a paper by hand. The determiner produced by the participant would indicate the assignment of the noun to a gender class through local agreement, which we could then match to the morphological ending of the noun. This task was given last to avoid priming.



- (22) l-a canast-a.
 DEF-F.SG basket-F.SG
 'The basket.'

Figure 3. DP-internal Gender Agreement Task.

6. Results

In the DP-internal agreement task, participants consistently assigned nouns to gender class (Table 2) whenever the noun ends in *-o* or *-a* ($\chi^2(5, N = 480) = 734.18$ $p < 0.001$). Nouns ending in a consonant and *-e* were excluded for two reasons: first, *-o* and *-a* are more transparently mapped onto masculine and feminine gender, potentially increasing the chances of agreement. Second, the nouns ending in a consonant and *-e* had very few tokens in the tasks.

Table 2. DP-internal Agreement Task: Percentage of Masc/Fem Ns indicated by determiner

	<i>-a</i>	<i>-o</i>
MASC	0%	84.5%
FEM	100%	15.5%

In the long-distance agreement task, if the clitic's gender matched the gender assigned to the corresponding noun when the participants produced the determiner in the DP-internal agreement task, it was coded as a case of successful agreement, otherwise, it was considered non-matching. The four possibilities are illustrated in Table 3.

Table 3. Clitic agreement options in Task 2

	Antecedent gender	Clitic gender
Matching agreement	Fem (<i>araña</i> , 'spider')	Fem (<i>la</i>)
	Masc (<i>cerro</i> , 'mountain')	Masc (<i>lo</i>)
Non-matching agreement	Fem (<i>araña</i> 'spider')	Masc (<i>lo</i>)
	Masc (<i>cerro</i> 'mountain')	Fem (<i>la</i>)

Figure 4 presents results for consistent vs. inconsistent gender between the antecedent and the clitic (target or non-target) by noun gender and by the first language of the participant where SP is Spanish (the group of mestizos) and SK is Shipibo.

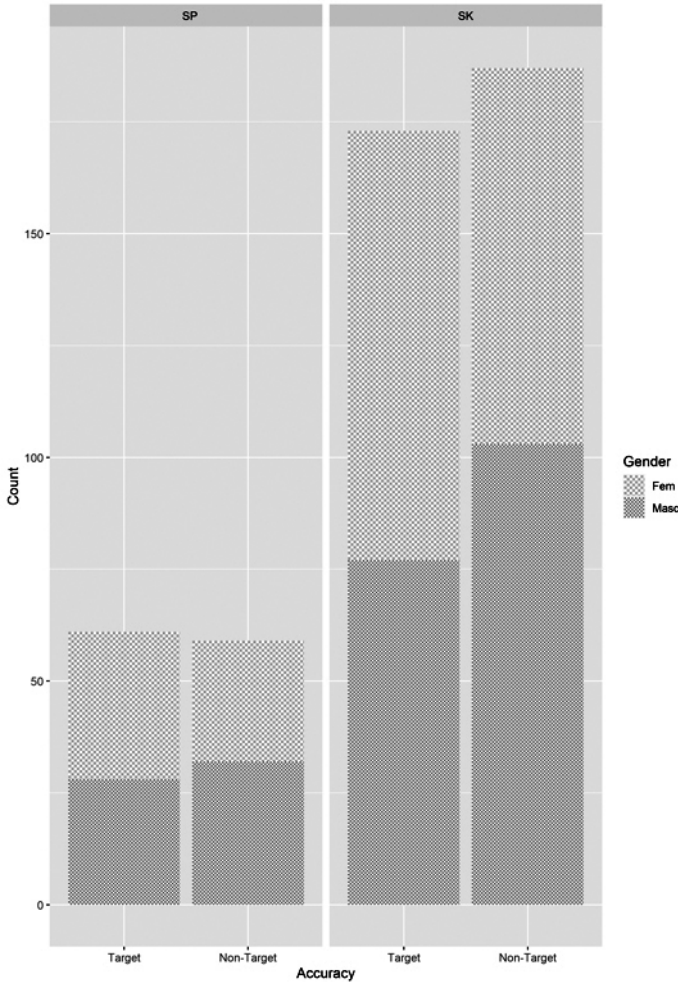


Figure 4. Clitic-antecedent gender match by gender and L1.

Target clitic-antecedent gender was almost random for both groups.⁶ For Spanish L1 speakers, target-like clitic-antecedent gender matching was 50.8% and for Shipibo L1 speakers, 48%, as shown in Table 4.

6. As the task was a forced-choice task, the data do not include the possibility for participants to state that they did not know which image to choose.

Table 4. Long-distance clitic-antecedent gender matching coding

	SP	SK
Target	51%	48%
Non-Target	49%	52%

($\chi^2(1, 480) = 0.178, p < .67$)

Overall, participants selected feminine antecedents slightly more target-like (53.7%) than masculine ones (43.75%). Table 5 shows the distribution of these forms among the two groups.

Table 5. Target and non-target antecedent selection

Spanish L1	<i>la</i> clitic	<i>lo</i> clitic
Target	55%	46%
Non-Target	45%	54%
Shipibo L1	<i>la</i> clitic	<i>lo</i> clitic
Target	53%	43%
Non-Target	47%	57%

A generalized linear model (R, lme4) with matching clitic-gender agreement (“Yes”, “No”) as the dependent variable and clitic gender (“Masc”, “Fem”) as independent variable revealed that feminine clitics were more highly associated with matching the antecedent than masculine clitics ($p = 0.03$).

Likewise, we tested whether the gender of the clitic would predict matching agreement patterns, but the generalized linear model with Clitic-gender agreement (“Yes”, “No”) as the dependent variable and Clitic gender (“Masc”, “Fem”) as predictor did not show clitic gender as significant, suggesting that the gender of the clitic does not determine whether the clitic will agree or not with its antecedent.

In sum, 1) nouns are assigned to gender classes, as signaled by local DP agreement; 2) long distance, CL-antecedent gender agreement is only a tendency, higher for feminine than for masculine antecedents; 3) clitic gender does not seem to affect gender agreement with the antecedent.

7. Discussion

The big DP hypothesis would lead us to expect that long-distance agreement (antecedent-clitic) and local (DP-internal) agreement should be similar. Our results show reliable local (DP-internal) agreement, indicating that Shipibo-Spanish speakers assign nouns to the right class and perform determiner-noun agreement involving feature valuation from N to D, as shown in (23):

(23) [_{DP} D_[_G] NP_[F]]

Agreement between the clitic and its antecedent, on the other hand, is at chance level (matching 51%, non-matching 49%) despite a slightly better matching for feminine (54%) than for masculine (45%) clitics and their antecedent. In contrast, assignment to gender class and local D-NP agreement is close to 100% accurate for feminine and 84.5% for masculine.⁷ These results are consistent with the findings of Shin et al. (2019), according to which bilingual children experience more difficulties with clitic antecedent agreement than with agreement between determiners and nouns. In the present study, this difference emerges despite the fact that the picture-selection task is a receptive task and the determiner-noun agreement task is a productive one. The difference suggests, first, that local and long-distance agreement patterns are different. Secondly, it suggests that the relationship between clitics and their antecedents in Shipibo-Spanish is not one of agreement, at least in gender, but more a relationship between a pronoun and its antecedent in sentences such as (24):

(24) Ayer vimos a *la niña-a_i*. Nos fuimos a jugar al parque con *ella-a_i*.
 ‘Yesterday, we saw the girl.F. We went to the park to play with her-F.’

The relationship between *la niña* ‘the girl.F’ and *ella* ‘her.F’ does not involve locally bound agreement, but perhaps long-distance concord, in that the pronoun in the second sentence in (24) is not bound. Similarly, the pronominal clitic does not establish a sentence-internal relationship that requires feature valuation but a relationship with an antecedent in discourse. If rather than feature valuation, a process of feature-matching were to take place between the clitic and its antecedent in the long-distance cases, as in (25), one could account for why non-identity features are allowed.

(25) DP_[F] CL/pro_[F]

If masculine is the default gender in Spanish (Harris 1991; Beatty-Martinez & Dussias 2019), then mismatches, when the clitic/pro is masculine, could be more frequent since speakers would assign the default to all cases where the concord relationship is uncertain. Alternatively, one could propose that, since Shipibo is an ergative language and different marking of the direct object and the subject is already a challenge, the acquisition of long-distance gender could prove to be an even greater challenge. The clitic would remain, for most L2 learners, an argument marker and

7. An anonymous reviewer points out that the tasks used in each case involved different levels of complexity in processing. While the determiner and the noun are in the same phrasal domain, the clitic and its antecedent are related in discourse. We acknowledge that this difference could play a role in the results, although the difference could be compensated by the fact that the determiner-noun agreement is tested in a production task which could in principle be cognitively more demanding than the receptive task with the clitic stimuli. Furthermore, as noted in Lewis and Vasisht (2005), other factors such as frequency of activation of gender features could be at play.

not a D with features. This would still not account for the lack of agreement between the clitic and its antecedent among the Spanish L1 group in the Shipibo community in the present study⁸ or for the instances of *lo* and *la* clitics found in Mayer & Sánchez (2017).

To account for the distribution of clitics in our results, we propose that in the picture selection task, the images evoke a linguistic representation that includes the lexical specification for gender of the antecedent in discourse, as is the case in (24), that is, there is no feature valuation. This specification matches the clitic and pro for speakers who show target-like long-distance agreement, as shown in (26):

(26) [canasta_[F]] [IP ... [VP [DP CL_[F] [DP pro_[F]]]]]

For Shipibo-Spanish speakers that do not have target-like long-distance agreement, we propose that the clitic lacks specified gender, so that again there is no phi-feature agreement with pro or the antecedent, nor matching between the clitic and the antecedent.⁹ This lack of specification is supported by results from previous work (Sánchez et al. 2022) showing that, in production, some L2 Shipibo-Spanish speakers show a preference for the clitic *le* (dative in other varieties) which is unmarked for gender. The antecedent is coreferential with the clitic/pro DP, but does not formally agree with it in gender:

(27) [canasta_[F]] [IP ... [VP [DP CL [DP pro]]]]

This analysis would better account for the distribution of the data in our study and could perhaps be extended to other contact varieties of Spanish such as the ones studied by Shin et al. (2019). For the present study, the grammar that has a clitic underspecified for gender would predict no systematic agreement pattern between the clitic and the antecedent, as found for some speakers. In fact, Shin et al. (2019)'s study found similar data. Children who used more English at home and who had higher English vocabulary scores had more gender mismatches and did not process the clitic's gender for referent identification, suggesting that the clitic may be unspecified for gender for some of those children.

8. Conclusions

In the bilingual variety of Spanish in contact with Shipibo, the clitic does not pattern like a determiner with respect to gender agreement. While gender agreement between D and N obtains categorically, gender agreement between the clitic and its antecedent is at chance level. Furthermore, all pro/CL are unmarked for gender,

8. Mismatches in gender have also been attested in other monolingual varieties of Spanish in contact with indigenous languages in Peru (Mayer & Sánchez 2017).

9. It is possible that individual speakers may have access to multiple grammars with different, potentially competing settings, or to a grammar with multiple specifications (e.g., see Adger & Smith 2005 and references therein). These results don't provide evidence for or against this possibility.

so no formal gender agreement takes place between the antecedent and pro/CL. We propose that DP-internal agreement is not the same operation as long-distance CL-antecedent agreement in these bilinguals. Shipibo-Spanish speakers systematically assign Ns to gender classes, as seen by agreement patterns between determiners and nouns, which also track morphological endings of nouns. However, clitic gender does not systematically match antecedent gender, particularly with feminine antecedents. These results are compatible with theories that do not assume clitics are generated in argument positions, inside a big DP, at least for this bilingual variety.

References

- Adger, D. & Smith, J. 2005. Variation and the minimalist program. In Cornips, L. & Corrigan, K. P. (eds.). *Syntax and variation: Reconciling the biological and the social*, 149-178. Amsterdam: John Benjamins.
- Alarcón, I.V. 2011. Spanish gender agreement under complete and incomplete acquisition: Early and late bilinguals' linguistic behavior within the noun phrase. *Bilingualism: Language and Cognition* 14(3): 332-350.
- Aoun, J. 1981. *The formal nature of anaphoric relations*. Doctoral dissertation, MIT, Cambridge, MA.
- Beatty-Martinez, A.L & Dussias, P. 2019. Revisiting masculine and feminine grammatical gender in Spanish: linguistic, psycholinguistic, and neurolinguistic evidence. *Frontiers in Psychology* 10.
- Belletti, A. 2005. Extended doubling and the VP periphery. *Probus* 17: 1-35.
- Brabec de Mori, B. 2015. *Die Lieder der Richtigen Menschen. Musikalische Kultur-anthropologie der indigenen Bevölkerung im Ucayali-Tal, Westamazonien*. Innsbruck: Helbling.
- Camacho, J. & Sánchez, L. 1995. The genitive clitic and the genitive construction in Andean Spanish. *Probus* 7: 133-146.
- Camacho, J. & Sánchez, L. 2002. Explaining clitic variation in Spanish. In Amberber, M. & Collins, P. (eds.). *Formal approaches to language universals & language variation*, 21-40. Norwood: Ablex.
- Castilla-Earls, A., Pérez-Leroux, A.T., Martínez-Nieto, L., Restrepo, M.A. & 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.
- Cecchetto, C. 1999. A comparative analysis of left and right dislocation in Romance. *Studia Linguistica* 53: 40-67.
- Cecchetto, C. 2000. Doubling structures and reconstruction. *Probus* 12: 93-126.
- Di Tullio, A., Saab, A. & Zdrojewski, P. 2019. Clitic Doubling in a Doubling World: The Case of Argentinean Spanish Reconsidered. In Gallego, A. (ed.). *The Syntactic Variation of Spanish Dialects*. Oxford: Oxford University Press.
<<https://doi.org/10.1093/oso/9780190634797.001.0001>>
- Escobar, A.M. 1990. *Los bilingües y el castellano en el Perú*. Lima: Instituto de Estudios Peruanos.
- Faust, N. 1990. *Lecciones para el aprendizaje del idioma Shipibo-Conibo*. Lima: Ministerio de Educación e Instituto Lingüístico de Verano.

- Fleck, D. 2013. *Panoan languages and linguistics*. (*Anthropological papers of the American Museum of Natural History*, no. 99). American Museum of Natural History.
- Franceschina, F. 2005. *Fossilized second language grammars: The acquisition of grammatical gender* (Vol. 38). Amsterdam: John Benjamins Publishing.
- Givón, T. 1976. Topic, Pronoun and Grammatical Agreement. In Li, C.N & Thompson, S.A. (eds.). *Subject and Topic*, 149-188. New York: Academic Press.
- Grüter, T., Lew-Williams, C. & Fernald, A. 2012. Grammatical gender in L2: A Production or a real-time processing problem. *Second Language Research* 28: 191-215.
- Instituto del Bien Común (IBC). 2016. *Directorio 2016 Comunidades Nativas del Perú*. Lima: Instituto del Bien Común.
- Instituto Nacional de Estadística e Informática (INEI). 2017. *Censos Nacionales 2017: XII de Población, VII de Vivienda y III de Comunidades nativas y comunidades campesinas*. Lima: Instituto Nacional de Estadística e Informática (INEI).
- Jaeggli, O. 1982. *Topics in Romance Syntax*. Dordrecht: Foris.
- Harris, J. 1991. The Exponence of Gender in Spanish. *Linguistic Inquiry* 22: 27-62.
- Kayne, R. 1975. *French syntax: The transformational cycle*. Current studies in Linguistics. Cambridge, MA: The MIT Press.
- Kayne, R. 1989. Null subjects and clitic climbing. In Jaeggli, O. & Safir, K. (eds.). *The null subject parameter*, 239-261. Dordrecht: Springer.
- Keating, G. 2010. The effects of linear distance and working memory on the processing of gender agreement in Spanish. In VanPatten et al. (eds.). *Research in second language processing and parsing*, 113-134. Amsterdam: Benjamins.
- Laleko, O. 2018. What is difficult about grammatical gender? Evidence from heritage Russian. *Journal of Language Contact* 11(2): 233-267.
- Lardiere, D. 2017. *Ultimate attainment in second language acquisition: A case study*. London: Routledge.
- Lenz, R. 1920. *La oración y sus partes* (No. 5). Madrid: Centro de Estudios Históricos.
- Lewis, R. & Vasishth, S. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29: 1-45.
- Lohndal, T. & Westergaard, M. 2016. Grammatical gender in American Norwegian heritage language: Stability or attrition? *Frontiers in Psychology* 7: 344. <<https://doi.org/10.3389/fpsyg.2016.00344>>
- Luján, M. 1987. Clitic-doubling in Andean Spanish and the theory of case absorption. In Matluck, J., Morgan, T., Lee, J. & VanPatten B. *Language and language use: Studies in Spanish*, 109-121. Lanham, MD: University Press of American.
- Mayer, E. 2010. *Syntactic variation of object arguments in Limeño Spanish contact varieties*. Unpublished Ph.D. dissertation, Australian National University.
- Mayer, E. 2017. *Spanish clitics on the move: Variation in time and space*. Berlin: Walter de Gruyter.
- Mayer, E. & Sánchez, L. 2016. Object agreement marking and information structure in monolingual and bilingual Andean Spanish. *Revista Española de Lingüística Aplicada* 29(2): 544-581.
- Mayer, E. & 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.

- Mayer, E. & Sánchez, L. 2020a. Feature selection in clitic expression in two bilingual Amazonian Spanish varieties. In Zariquiey, R., Valenzuela, P., Escobar, A. & Jara, M. (eds.). *Spanish Diversity in the Amazon: Dialect and Language Contact Perspectives*, 386-429. Brill Series Indigenous Languages of the Americas.
- Mayer, E. & Sánchez, L. 2020b. Clitics and argument marking in Shipibo-Spanish and Ashéninka-Perené-Spanish bilingual speech. In Fafulas, S. (ed.). *Amazonian Spanish: Language contact and evolution*, 128-154. Amsterdam: John Benjamins.
- Quicoli, A.C. 1980. Clitic movement in French causatives. *Linguistic analysis* 6(2): 131-185.
- Sánchez, L. 2003. *Quechua-Spanish Bilingualism*. Amsterdam and Philadelphia: John Benjamins.
- Shin, N., Rodríguez, B., Armijo, A. & Perara-Lunde, M. (2019). Child heritage speakers' production and comprehension of direct object clitic gender in Spanish. *Linguistic Approaches to Bilingualism* 9(4-5): 659-686.
- Sánchez, L., Camacho, J., Mayer, E. & Rodríguez Alzza, C. 2022. Gender Agreement in a Language Contact Situation. *Languages* 7(2): 81.
<<https://doi.org/10.3390/languages7020081>>
- Sportiche, D. 1996. Clitic constructions. In Rooryck, J. & Zaring, L. (eds.). *Phrase structure and the lexicon*, 213-276. Dordrecht: Springer.
- Suñer, M. 1988. The role of agreement in clitic-doubled constructions. *Natural Language & Linguistic Theory* 6(3): 391-434.
- Uriagereka, J. 1995. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26(1): 79-123.
- Valenzuela, P. 2003. *Transitivity in Shipibo-Konibo grammar*. Unpublished PhD dissertation, University of Oregon.
- Valenzuela, P. 2010. Ergativity in Shipibo-Konibo, a Panoan language of the Ucayali. In Gildea, S. & Quixalós, F. (eds.). *Ergativity in Amazonia*, 65. Amsterdam: Benjamins.
- Zdrojewski, P. & Sánchez, L. 2014. Variation in accusative clitic doubling across three Spanish dialects. *Lingua* 151: 162-176.