



Emerging DOM patterns in clitic doubling and dislocated structures in Peruvian-Spanish contact varieties

Author

Mayer, E, Sánchez, L

Published

2021

Book Title

Differential Object Marking in Romance: The third wave

Version

Version of Record (VoR)

DOI

[10.1515/9783110716207-005](https://doi.org/10.1515/9783110716207-005)

Downloaded from

<http://hdl.handle.net/10072/417134>

Griffith Research Online

<https://research-repository.griffith.edu.au>

Elisabeth Mayer and Liliana Sánchez

Emerging DOM patterns in clitic doubling and dislocated structures in Peruvian-Spanish contact varieties

Abstract: In this chapter we explore the expression of Differential Object Marking (DOM) in monolingual and bilingual Spanish in contact with typologically different languages. We focus on how DOM patterns are expressed in bilingual and monolingual clitic doubling and dislocated structures by investigating the effect of typological differences in case marking and the effects of definiteness, animacy and thematic role. Our findings show that while definiteness, animacy and thematic structure are relevant factors in the production of DOM across bilingual groups, there are also differences in DOM frequency related to typological distinctions in the L1 of the bilingual groups. Finally, our findings also show variability in the monolingual data of three individuals raised in a contact situation as well as the raise of topicality as a possible factor that contributes to DOM in the L2 varieties under study.

Keywords: Differential Object Marking (DOM), clitic doubling, bilingual acquisition, typological differences, Andean and Amazonian Spanish

1 Introduction

Differential Object Marking (DOM) is a variable and widespread argument marking system (Bossong 1991; 2003; Aissen 2003; Dalrymple/Nikolaeva 2011; Witzlack-Makarevich/Seržant 2017; a.o.). It is a widely attested strategy across many genetically unrelated languages used to mark a select range of specific and/or topical direct objects (Bossong 1991; 2003; Aissen 2003; Dalrymple/Nikolaeva 2011; a.o.). In the emergence of DOM in Romance, three different but interacting facts have been identified (Laca 2006; Mayer 2017, 101). One is the need to differ-

Acknowledgements: We would like to thank all participants in the study and acknowledge the contribution of the following research assistants: Yoshidaira Garcia for help with Huánuco Quechua, Caleb Cabello Chirisente for Asháninka, and Enrique Espinoza for Shipibo, and Julio López Otero for help with the statistics [R].

Elisabeth Mayer, Griffith University Queensland, e-mail: e.mayer@griffith.edu.au
Liliana Sánchez, University of Illinois at Chicago, e-mail: lesanche@uic.edu

entiate between subject and object arguments (Bossong 1991). Secondly, separation of dative and accusative marking is relevant, and finally variability in verbal argument structure in certain verbs (Laca 2006). These factors have given rise to multiple competing theories on how DOM emerged. The emergence of a new accusative form (Givón 1997; Bossong 2003) versus a single plus/minus dative case system (Alsina 1996) and the transition from neutralization of dative and accusative marking (Lapesa 2000) to the “cannibalistic” dative theory (Company 2001; 2003). The third and last point refers to topic marking (Givón 1976) and its extension to secondary topic marking (Dalrymple/Nikolaeva 2011). While DOM preferably marks the argument itself, Differential Object Indexing (DOI), another related and very similar strategy, preferably marks the verb. Notably, both strategies can cooccur in one language (Iemmolo 2011).

Once DOM emerges in the grammar, it may be conditioned by a wide range of semantic and syntactic properties of the marked direct object. Factors that have been argued to trigger DOM are the interaction of the semantic-pragmatic properties of the object such as animacy and definiteness (Bossong 1991; 2003; Aissen 2003, for a counter argument cf. Sinnemäki 2014), and other information structural properties that can only be interpreted at higher constituent levels such as topicality (Leonetti 2008), and telicity (Torrego 1998).

In this study, we investigate emerging patterns of DOM in Clitic Doubling (CLD) and the two related information structures Clitic Left Dislocation (CLLD) and Clitic Right Dislocation (CLRD). We focus on two referential factors: animacy and definiteness as well as on the verbal thematic roles *PATIENT* and *THEME* of the direct object noun phrase which have been less explored as a factor involved in DOM. Both thematic roles are subject to different degrees of affectedness, while *PATIENT* objects undergo a change, *THEME* objects remain unchanged (Naess 2004). In order to understand how some of these factors interact with each other in a continuum of language contact, we conducted a study among three groups of bilingual speakers of one of three indigenous languages spoken in Peru: Asháninka, Quechua, Shipibo and Spanish as well as of three monolingual speakers of Spanish exposed to contact varieties living in a continuum of language contact situations.

It has been shown that DOM in clitic doubling and dislocated structures¹ exhibits high levels of variability across Spanish dialects more generally and particularly in contact varieties (Mayer/Sánchez 2016; 2017). This variability may be due to the specific properties of the languages with which it is in contact, spe-

¹ We focus on DOM in these structures as DOM exclusively occurs in them in the indigenous languages.

cifically in languages that lack DOM as is the case of Asháninka, Quechua and Shipibo. The study of the development of DOM in these contexts is of relevance in the understanding of how DOM emerges and evolves and in the understanding of the main semantic and syntactic features that trigger it. In the next Sections, we present a brief overview of the framework we assume for language contact situations, a brief description of case marking in each of the indigenous languages and Spanish, the study design, its results and conclusions.

2 Contact and SLA continuum

It is well known that in language contact situations, divergence at the interfaces and typological differences in morphological patterns may result in new systems (Matras 2010) due to a range of ecological factors (Mufwene 2002). These new systems reflect different grammaticalization stages and different choices in the selection from a feature pool (Mufwene 2001; 2002; Matras 2010; Mayer/Sánchez in press) depending on factors such as levels of proficiency and frequency of exposure to bilingual varieties as well as frequency of interaction and activation of features between bilinguals and monolinguals (Putnam/Sánchez 2013). Following Mufwene (2001; 2002) we assume that idiolectal output systems are the result of the selection of the best fitting feature available for communicative purposes from the competing features in any given pool, see Figure 1 below. In the case of the bilingual groups under investigation, contact with xenolectal input contributes different alignment systems to the feature pool in a continuum that goes from very different to closer to Spanish. The feature pool is also impacted by differences in input to Spanish due to varying access to educational opportunities. Bilingual as well as monolingual speakers find themselves confronted with a complex relationship between idiolectal input, xenolectal input in the form of language or dialect contact, as in the case of the monolinguals – and access to educational opportunities. Idiolectal output systems are the result of the selection of a subset of features in individuals and may be subject to crosslinguistic convergence (Sánchez 2003).

Furthermore, differences in the mapping of syntactic features onto morphology are a main factor in differential bilingual development. In second language acquisition studies this has been identified as difficulties in assembling or reassembling features in the L2 that are not activated in the L1. The Feature Reassembly Hypothesis states that these difficulties stem from the need to reassemble features and associate them with new morphological forms in the L2 (Lardiere 1998; 2005). More recently, the Bottleneck hypothesis in SLA (Slabakova 2008; Jensen et al. 2017) has identified functional morphology as the locus of greater levels of difficulty in

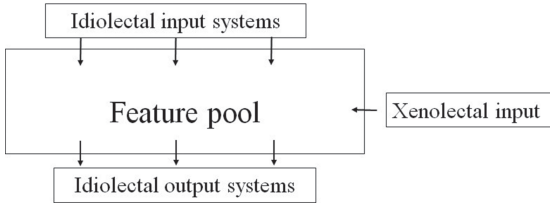


Figure 1: Feature pool: input varieties and output varieties.

second language acquisition. Monolingual child acquisition of DOM in monolingual contexts has been described as straightforward (Rodríguez-Mondoñedo 2008) and somewhat variable among L2 speakers but as variable among heritage speakers (Montrul 2004). One source of this variability lies in differences between the acquisition of DOM as triggered by object inherent semantic features like animacy and definiteness versus structural features like case (Blake 2001). The former has been identified as more difficult to learn than the latter. Even within semantic features, animacy has been identified as more learnable as a DOM trigger than the discourse-related features definiteness and specificity (Guijarro-Fuentes 2011; 2012).

In this study, we take into account the inherent variability of language acquisition when analysing bilingual and monolingual DOM clitic doubling and information structure systems in contact situations, as these involve individual differences that result from variability in the feature pool and sensitivity to different semantic features.

3 The study

In previous work (Mayer/Sánchez in press), we examined the distribution of clitics and DOM in clitic doubled constructions in Amazonian Spanish in contact with Shipibo and Ashéninka Perené – a different data set from the Asháninka data set used here. In this chapter, we focus on crosslinguistic effects and DOM in oral production data from bilingual Spanish of speakers of three typologically different indigenous languages in Peru across urban and rural settings. As a point of comparison, we also include data from a longitudinal study of monolingual Spanish speakers in Lima (LSCV-Lima Spanish contact varieties).² We explore how DOM patterns are expressed in bilingual and monolingual clitic doubling

² The speakers of LSCV whose data we analyzed are part of a continuum of Quechua and Spanish contact that ranges from passive knowledge of Quechua to monolingualism in Spanish.

and information structures. In order to do so, we investigate the effect of contact on case marking based on the conditioning features animacy and definiteness in definite and indefinite object Determiner Phrases (DP) as well as transitivity expressed through the thematic roles PATIENT and THEME in marked and unmarked direct object arguments in clitic doubled and dislocated information structures. The selection of these information structures stems from the fact that, in our contact data, DOM preferably occurs in those structures and our data sets show practically no evidence of DOM in simple transitive clauses.

The following examples reflect the preferred DOM strategies of each of the three bilingual data sets and the monolingual data set. The bilingual datasets consist of data from two Amazonian bilingual groups and one Andean bilingual group. The monolingual data comes from a longitudinal study of three individuals in Lima representing Limeño Spanish Contact Varieties (LSCV).

Expression of DOM in Quechua-Spanish from a rural Andean area in (1) is strikingly different from the other data sets by marking close to 100% of all direct objects, based on the features animacy and definiteness with 90% each. In terms of thematic roles, the preference for marking THEME over PATIENT is significant. The high percentage for both features of the object DP can be analyzed as the result of a shared nominative/accusative alignment system.

- (1) *Le acarici-a a su perr-o*
 CL.3SG caress-PRES.3SG DOM POSS dog-M.SG
 ‘(S)/He caresses her/his dog.’

Huánuco Quechua-Spanish (Sánchez Dataset 2005)³

Data from rural Asháninka-Spanish show a different scenario being the only group with a preference for unmarked DPs (55%) over marked DPs (45%). DOM

³ The following abbreviations are used in this chapter: A = agent, ABS = absolutive, ACC = accusative, AUX = auxiliary, CAUS = causative, LEJ/CIS = cislocative, CL = clitic, CMPL = completive, DEM = demonstrative, DET = determiner, DIM = diminutive, DOM = Differential Object Marking, DP = determiner phrase, ERG = ergative, EV/EVID = evidential, F = feminine, GER = gerund, HSY = hearsay, IMPF = imperfective, INC = incorporation, INDEF = indefinite determiner, INF = infinitive, IRR = irrealis, LOC = locative, M = masculine, O/OBJ = object of transitive verb, P/pat = patient, 1 = first person, 2 = second person, 3 = third person, PART = participle, PAS/PAST = past tense, PFV/PRF/PERF = perfective, PL = plural, P/POSS = possessive, PP2 = past completive particle, PP = positive polarity, PR = prospective, PRES = present tense, PRO = pronoun, PRT = preterit, REAL = realis, S = subject of intransitive verb, SG = singular, SUB = subordinate, SS = same subject, TOP = topic, T/TR = transitive, TH = theme. For the examples from South American indigenous languages we keep the original glosses as used by the authors in the publications. Differences with the abbreviations used in Spanish and additional abbreviations are noted in bold.

as in (2) occurs preferably with definite and animate DPs, PATIENTS are strongly preferred over THEMES. Given the fact that Asháninka is a nominative/accusative language with fluid/split transitivity and the availability of gender specific bound morphemes as argument markers it is possible that these factors support acquisition of DOM in clitics and the coding of information structure.

- (2) *Lo bot-aron a la rana*
 CL.3.M.SG kick-PERF.3PL DOM DET.F.SG frog.F.SG
 ‘They kicked out the frog.’

Asháninka-Spanish (Mayer Dataset 2016)

Shipibo-Spanish shows exactly the opposite distribution from Asháninka-Spanish for marked (55%) and unmarked objects (45%). The distribution of DOM follows a clear preference of animacy closely followed by the thematic role PATIENT with little attention to the feature definiteness and thematic role. Ergative alignment in Shipibo seems to be a possible cause for the lack of DOM with a human and definite object in (3).

- (3) *Le molest-a Ø el niño*
 CL.3SG bother-PRES.3SG Ø DET.M.SG child-M.SG
 ‘He bothers the child.’

Shipibo-Spanish (Sánchez Dataset 2002)

Finally, monolingual Spanish (LSCV) shares a similar distribution as Shipibo-Spanish in terms of marked and unmarked objects however with a slightly bigger margin between marked (60%) and unmarked objects (40%). LSCV shares the hierarchical distribution of features and thematic roles with Shipibo-Spanish however with two important differences. For one, the features definiteness and animacy (61.5% and 56.1% respectively) play a significantly stronger role in marking than the thematic roles PATIENT and THEME (26.2% and 23.8% respectively). Extension of DOM to inanimate, definite, patient direct object in clitic doubled DPs as in (4) and dislocated constructions is common in LSCV. These kinds of extensions of DOM can be linked to LSCV as an acquisitional and mainly oral variety, where speakers are located on a continuum dependent on variable access to formal instruction.

- (4) *Lo licu-o a-l ajo*
 CL.3.M.SG blend-PRES.1SG DOM-DET.M.SG garlic.M.SG
 ‘I blend the garlic.’

LSCV (Mayer 2017, 65)

A unifying factor of our different data sets is a complex interaction of direct or indirect contact with typologically different languages, language ecological factors (Mufwene 2002) and stages of language acquisition (Lardiere 2005). We find common syntactic properties across a continuum of speakers and variability in morphological marking, lack of DOM with human objects and general extension to definite inanimate objects in combination with non-agreeing clitics. Overall, definiteness is found to work as a semantic restriction for topic-worthy objects overriding animacy constraints. Due to these facts and as DOM preferably occurs in clitic doubled and dislocated structures, we will analyze the interaction of the above factors as a strong tendency to mark information-structure (Mayer/Sánchez 2016; Mayer 2017).

Given the typological differences between the languages in contact with Spanish and the complex array of syntactic conditions and semantic features that yield great variability of DOM and its relationship to clitic doubling and dislocated structures in Spanish, we formulate the following two research questions:

1. Are there differences in the Spanish DOM patterns found in clitic doubling and clitic dislocated structures in contact bilingual varieties depending on the typological characteristics of the contact language? Do monolingual contact varieties differ too?
2. What is the effect of animacy, definiteness and thematic role on the presence or absence of DOM in bilingual Spanish clitic doubling and dislocated structures in contact bilingual and monolingual varieties?

We formulate the following hypotheses:

1. DOM patterns in clitic doubling and clitic dislocated structures in contact varieties are affected by differences in the configuration of the morphology of the contact language. We expect the Quechua group to exhibit more DOM than the other groups as a result of similarities with Spanish in nominative/accusative alignment. We expect the Shipibo and the Asháninka groups to exhibit lower frequencies of DOM due to differences in alignment between these languages and Spanish.
2. We expect animacy to have an important role in the acquisition of Spanish due to its relevance in the input. Given that none of the indigenous languages has overt definite determiners we expect animacy to play a greater role than definiteness in all bilingual groups but possibly less so in the contact monolingual individuals given the availability of definite determiners. Finally, we expect thematic role to play a more salient role in Asháninka Spanish given the sensitivity to thematic role in the morphology of Asháninka.

In the next Section, we describe the typological differences in the argument marking systems of Shipibo, Huánuco Quechua, Asháninka and Spanish and address the difficulties arising for bilingual speakers from typologically different systems.

4 Typological differences in argument marking systems

4.1 Argument-marking in Shipibo

In this Section we present the different properties of case marking in each of the indigenous languages and in Spanish. As shown below, each of the three languages differs from Spanish with respect to case marking in significant ways although of the three, Quechua is closer in terms of case alignment.

Shipibo is part of the Panoan language family, it is an agglutinative language with ergative/absolutive alignment in its nominal and pronominal systems. It has null objects and SOV word order (Valenzuela 2010). In an ergative/absolutive system, pronominal subjects of most intransitive verbs as well as overt objects of most transitive verbs are marked as ergative. In (5a), the pronominal form *jato* is the subject (S) of an intransitive verb and in (5b) the object (O) of a transitive verb.

- (5) a. *Moa-ra jato bo-kan-ai*
 already-EV 3P:ABS go.NON.SG-PL-INC
 ‘They are leaving already.’

(Valenzuela 2010, 71)

- b. *Ja-n-ra jato keyyo-ke*
 3-ERG-EV 3P:ABS finish-CMPL
 ‘S/he exterminated them.’

(Valenzuela 2010, 71)

As shown in (6), subjects of transitive verbs have a different pronominal form *jabaon*.

- (6) *Jaská-a-xon-ki ja-baon no-a onan-ma-a iki*
 SO-DO.T-A-HSY2 3-PL:ERG 1P-ABS know-CAUS-PP2. AUX
 ‘And then they (our grandmothers) taught us (the activities that are proper of women).’

(Valenzuela 2010, 71)

When a third person subject co-occurs with a third person and object, verbal subject and object agreement may be null altogether as shown in (7).

- (7) *Nima_i oin-xon-ra, Jose-kan kena-Ø-_ike*
 Nima see-SS.TR.PRT Jose-ERG call-Ø-PERF
 ‘When he saw Nima, José called him.’

(Loriot/Lauriault/Day 1993, 56)

Given the ergative/absolutive alignment typically found in Shipibo, there is to the best of our knowledge no evidence of Differential Object Marking.

4.2 Argument-marking in Asháninka

Asháninka is an Arawakan language with VS, VO basic constituent order, nominative-accusative alignment in transitive clauses and fluid or split transitivity in intransitive clauses (Payne/Payne 2005; Mihás 2015, 5). Fluid or split transitivity refers to the lexical meaning of the verb root being open to an increase or decrease in valency that determines transitivity. Bound morphemes mark obligatorily A/S (agent/subject) and P/O (patient/object) arguments on the verb. A/S arguments can occur pre- or postverbally and are marked according to semantic role. Of the three indigenous languages discussed here, Asháninka is the only one to mark gender with non-masculine gender as the default.

In transitive clauses in nominative/accusative alignment as in (8), the basic constituent order is VO, with compulsory marking of the subject A with a verbal prefix and the object O with a verbal gender specific suffix (*-ri* masculine, *-ro* non-masculine). Note that these gender-specific forms also mark the semantic role of theme in transitive constructions (Mihás 2015, 200).

- (8) *No-pos-ak-i-ri*
 1SG.A-hit-PFV-REAL-3M.O
 ‘I hit him.’

(Mihás 2015, 441)

In transitive clauses co-referential overt pronouns may co-occur as topics additionally to the postverbal object marker as shown in (9) but they are not marked for case.

- (9) *Iri no-pos-ak-e-ri ir-ako-ki*
 3M.TOP 1SG.A-hit-PFV-REAL-3M.O 3M.POSS-arm-LOC
 ‘I hit him on his hand.’

(Mihás 2015, 193)

Finally, Asháninka also exhibits ambitransitive transitive clauses which lack fixed transitivity thus allowing for increase or decrease of their valency even without object morphological marking on the verb as in (10). Such clauses exhibit fluid transitivity as explained above.

- (10) *n-a-ak-i kaniri*
 1SG.S-take-PFV-REAL manioc
 ‘I obtained manioc roots.’

(Mihás 2015, 194)

Once again, crucial for our present study is that the dislocated object is not marked for accusative case. This is also the case in (11) where the left dislocated object is unmarked too. Gender-specific suffixes appear on both verbs but not on the object.

- (11) *pakitsa ari a-ñ-ak-e-ri apaata a-shiyakant-ak-e-ri*
 eagle PP 1PL.A-see-PFV-IRR-3M.O later.on 1PL.A-take.picture-PFV-IRR-3M.O
 ‘The eagle, there we’ll see it, when we later on take pictures of it.’

(Mihás 2015, 616)

As in the case of Shipibo, to the best of our knowledge, there is no evidence of DOM in Asháninka.

4.3 Argument marking in Huánuco Quechua

Huánuco Quechua is a nominative/accusative, agglutinative language with SOV word order as shown in (12):

- (12) *Juan Tumas-ta maga-n*
 Juan Tomas-OBJ hit-3
 ‘Juan hits Tomas.’

(Weber 1996, 51)

In this sentence, the direct object is marked with the accusative suffix *-ta* and there is no overt nominative marking. Like Shipibo, Asháninka and Spanish, Huánuco Quechua has null subjects as shown in:

- (13) *Maga-ma-ra-n*
 hit-1-PAS-3
 ‘S/He hit me.’

(Weber 1996, 249)

When first or third person subjects co-occur with a third person object, the latter is not marked on the verb as shown in (14a) and (14b):

- (14) a. *Maga-Ø-n*
 hit-Ø-3
 ‘S/He hits (him, her, it).’
 b. *Apa-mu-na-:-paq ka-yka-sah-:*
 bring-LEJ-SUB-1P-PR be-IMP-PRF-1
 ‘I was about to bring (it).’

(Weber 1996, 168)

While there is no evidence of DOM in Huánuco Quechua of the type that is sensitive to semantic features such as definiteness, animacy or thematic role, some objects inside subordinate clauses may remain unmarked for accusative case in Huánuco Quechua as in:

- (15) *Cristobal-Ø asi-q aywa-ska-:*
 Cristobal search-SUB go-PRF-1
 ‘(I) went to look for Cristobal.’

(Weber 1996, 250)

This indicates the possibility of some alternation between direct object marking and lack of direct object marking on the noun. Finally, there is evidence in Huánuco Quechua of left and right topic dislocated structures as shown in (16). The dislocated element is marked with the topic suffix *-ga*.

- (16) a. *Hatun wasi-ta-ga muna:*
 big house-OBJ-TOP want-1
 ‘I want a big house.’

- b. *Wasita muna: hatunta-qa*
 House-OBJ want-1 big-OBJ-TOP
 ‘I want a big house.’

(Weber 1996, 515)

The existence of these dislocated structures may facilitate the acquisition of clitic doubled and clitic doubling structures.

4.4 Spanish argument marking

Spanish is a nominative/accusative language with minimal case marking on personal pronouns, some relative pronouns, clitics and a range of direct objects. In (17), the subject bears unmarked nominative case, the definite and human direct object receives accusative case through DOM.

- (17) *La niñ-a am-a a-l niñ-o*
 DET.F.SG girl-F.SG love-PRES.3SG DOM-DET.M.SG boy-M.SG
 ‘The girl loves the boy.’

(Mayer 2017, 58)

Direct object arguments can be marked by feature-agreeing verbal clitics as anaphors in the examples in (18). Clitic features include person, number, gender (in accusative/direct object clitics) or case (in dative/indirect object clitics). They are phonologically bound to their verbal host and can appear as proclitics with finite verbs (18a) and enclitics in non-finite contexts (18b). As subject information is encoded in verbal inflection, the overt expression of pronominal subjects depends on their informational status (18a).

- (18) a. *(Ella) lo am-a*
 PRO.3.F.SG CL.3.M.SG love-PRES.3SG
 ‘She loves him (the boy).’
 b. *Quier-e am-ar-lo*
 want-PRES.3SG love-INF-CL.3.M.SG
 ‘She wants to love him.’

As mentioned in the Introduction, DOM in Spanish marks a range of objects ranked for prominence on a two-dimensional scale based on the interaction of animacy and definiteness and also semantic role (Bossong 1991; 2003; Aissen 2003). According to Leonetti (2008), in most varieties of General Spanish, human

and animate patient arguments as in (19a) are generally marked, they are considered highly topical; indefinite animate objects may be optionally marked because of their specificity/identifiability status as in (18b), and inanimate non-specific core object arguments are excluded from marking (19c).

- (19) a. *Amalia abraz-ó a su tí-a*
 Amalia hug-PERF.3SG DOMDOM POSS aunt-F.SG
 ‘Amalia hugged her aunt.’
- b. *El ladrón mat-ó (a) un perr-it-o*
 DEF.M.SG thief.M.SG kill-PERF.3SG DOM INDEF.M.SG dog-DIM-M.SG
 ‘The thief killed a little dog.’
- c. *María compr-ó *a un pian-o nuev-o*
 María buy-PERF.3SG DOM INDEF.M.SG piano-M.SG new-M.SG
 ‘Maria bought a new piano.’

In clitic doubling and related configurations expressing information structure such as CLLD and CLRD clitics play an important role in Spanish syntax in object-verb agreement as head markers in conjunction with dependent marking by DOM (Nichols 1986).

This marking strategy is characterized by a continuum of DOM and CLD with both marking strategies exhibiting great diachronic and synchronic variability across time and space. In some varieties of Spanish, DOM in CLD is restricted to pronominal direct objects. Historically, the extension of optional marking to inanimate proper nouns has been documented in the Old Spanish text *Cantar de mio Cid* and it has been related to a specific relationship between the subject and the object which allows for referential identification (Melis 1995; Bresnan/Aissen 2002, 91). Contemporary liberal clitic doubling varieties such as Buenos Aires Spanish and Lima Spanish extend DOM to animate specific (20a) (Mayer 2008; 2017; Zdrojewski/Sánchez 2014) and in the case of Buenos Aires Spanish to inanimate specific direct object arguments (20b).

- (20) a. *La bes-o /ve-o a la maestr-a*
 CL.3.F.SG kiss-PERF.1SG /see-PERF.1SG DOM DET.F.SG teacher-F.SG
 ‘I kiss/see the teacher.’
- b. *Lo quier-o mucho a este arbol-ito*
 CL.3.M.SG like-PRES.1SG much DOM DEM.M.SG tree-DIM.M.SG
 ‘I like this little tree very much.’

(Suñer 1989, 379)

Extension of optional DOM in contemporary varieties of Spanish in (19b) points to information structural marking (Dalrymple/Nikolaeva 2011; Mayer/Sánchez 2016; Mayer 2017). Clitic configurations such as left dislocated structures (CLLD) as in (21a) and right dislocated structures (CLRD) as in (21b) also known as topicalized structures show the same patterns of head and dependent marking.

- (21) a. *A la maestr-a la salud-é ayer*
 DOM DET.F.SG teacher-F.SG CL.3.F.SG greet-PERF.3SG yesterday
 ‘The teacher, I greeted her yesterday.’
- b. *La salud-é ayer a la maestr-a*
 CL.3.F.SG great-PERF.3.SG yesterday DOM DET.F.SG teacher-F.SG
 ‘I greeted her, the teacher yesterday.’

The syntactic and semantic factors that determine marking of direct objects including clitic agreement in most varieties of Spanish are summarized in Table 1. They show that DOM depends on the interaction of the referential features animacy and definiteness allowing for variability in animacy with no significant role for thematic roles.

Table 1: Spanish direct object marking, case, clitic agreement (CL) and thematic roles (adapted from Mayer 2017, 61).

	Direct Object	
	marked direct object	unmarked direct object
Syntax		
Case	accusative <i>a</i>	∅
CL Agreement	yes	no
Semantics		
Object type	PAT/THEME	PAT/THEME
	[±anim/+def]	[-anim/-def/-spec]
	[+spec/+top]	

Given the complex array of syntactic conditions and semantic features yielding variability of DOM and its relationship to clitic doubling in Spanish, two questions arise immediately. These are a) how do bilinguals in language contact situations map functional features onto morphology, and b) how do monolingual contact speakers navigate the variability arising from the complex settings for DOM. As we will discuss in the rest of the chapter, most varieties of contact Spanish show even greater variability than non-contact varieties as they may lack DOM or extend case marking to new contexts.

4.5 The bilingual puzzle

In terms of differences and similarities of the bilingual systems, differently from Spanish, Shipibo exhibits case marking in nouns but the verb is unmarked for subject and object agreement. Important for clitic doubling structures is the fact that while Spanish clitics have person, number and case/gender features, the Shipibo pronominal system lacks gender and definiteness. In order to acquire clitic structures, Shipibo-Spanish bilinguals need to master the assembly of pronominal agreement with multiple features into one single morpheme (Lardiere 2005). Furthermore, they need to acquire sensitivity to semantic features such as definiteness and animacy as triggers of DOM.

Quechua argument marking is closest to Spanish argument marking in relation to case marking (nominative/accusative) as expressed by the accusative suffix *-ta* in Quechua and the DOM marker *a* in Spanish. As our data show, common nominative/accusative alignment with Spanish facilitates the acquisition of DOM for Quechua-Spanish bilinguals despite the morphological differences between the Spanish preposition *a* and the Quechua case marking suffix *-ta*. Like Shipibo, Quechua lacks marking of (in)definiteness and gender when the subject and the object are third person.

Asháninka, unlike Shipibo, shares two important similarities with Spanish, which could contribute significantly to the acquisition of Spanish argument marking in terms of DOM and clitic structures by Asháninka-Spanish bilinguals. These are a set of feature specific bound morphemes that mark objects as verbal affixes and a set of free personal pronouns to mark information structure in preverbal or postverbal position. Given the fact that the first one of these sets is specified for person, number and gender, it could play an important role in the acquisition of Spanish clitics. Also, the fact that these two sets of bound and free forms can co-occur and allow thus for the possibility to restructure word order pragmatically, could potentially facilitate the acquisition of clitic doubled and dislocated structures.

In the next Section, we address the ethnographic description, methodology, and data collection procedures.

5 Ethnography and methodology

In order to address this question, we analyzed the data sets of 3 bilingual L2 Spanish speaking groups in contact with typologically different languages living in their communities in rural and urban areas. By way of comparison, the analysis

was extended to a monolingual Spanish acquisitional variety in Lima. The focus is on the interaction of animacy and definiteness in conjunction with the thematic roles, PATIENT and THEME, in Differential Object Marking in clitic doubled and dislocated arguments structures.

5.1 Participants

For the bilingual argument structures, the data sets are based on fieldwork with L2 learners of Spanish with a Shipibo community in Lima, with a Quechua community in Huánuco, a rural Andean area, and with two Asháninka communities in the central Peruvian Amazon in the Junín province. The monolingual LSCV data set is a longitudinal study of three months with two sisters and the daughter of one of them living in two socioeconomically different neighbourhoods in Lima. Participant details in terms of numbers, gender, age range, education and place of data collection are specified in Table 2 below.⁴

Table 2: Ethnographic details of participants.

<i>Communities</i>	Σ	<i>Gender</i>	<i>Age</i>	<i>Education</i>	<i>Place</i>
<i>Shipibo</i>	19	10F, 9M	29–56	P:9F,3M S:1F,6M	Cantagallo, Lima
<i>Huánuco Quechua</i>	14	8F, 6M	22–45	P:6F,4M S:3F,2M	Chaglla, Huánuco
<i>Asháninka</i>	21	13F, 8M	18–88	P:5F,1M S:4F,3M PS:4F,4M	Arizona, Puerto Ocopa, Junín
<i>Monolingual LSCV</i>	3	3F	16–45	S:3F	Ventanilla, Surco, Lima

5.2 Methodology

For the Quechua and Asháninka groups a picture-based elicitation task was used to elicit a narration. The task was based on Sanchez's (2003) adaptation of Mayer and Mayer's (1992) frog story. A very similar story was used to collect the Shipibo

⁴ For the abbreviations in education, P refers to primary schooling, S to secondary schooling and for the Asháninka participants PS to postsecondary which in this case refers to an Agricultural Technical Institute located in the Puerto Ocopa community.

Spanish data using a figure-based narrative elicitation task. In order to investigate the research questions, firstly an ethno-biographical survey about language history, preferences for language use including language attitudes towards their indigenous languages and Spanish was conducted, followed by the oral elicitation tasks. Both were orally administered and digitally recorded. In the case of the Asháninka group these were followed by a short written Spanish language proficiency test adapted from the cloze section of the DELE used in Cuza et al. (2013). Older and illiterate participants were read the proficiency test and they completed it orally. An important point to make here is that specifically in Andean and Amazonian bilingual Spanish rurality and orality are predominant.

The Shipibo narratives were collected in 2002 in Lima, Peru, the Huánuco Quechua narratives in 2005 in Chaglla, Huánuco, Peru and the narratives of two Asháninka bilingual groups in Arizona Portillo and Puerto Ocopa, Junín, Peru in 2016. In the case of the Asháninka groups, argument structures were elicited in order to test the effect of similarities in case marking, gender marking and affixal verbal subject and object marking.

Data collection with the three monolinguals took place in early 2006 in Lima over the course of three months in the form of recorded elicited narratives about their lives, their regular activities and general life stories. Two participants are siblings, born to Quechua-speaking parents in Iquitos who migrated to Lima at ages 14 and 16. Their inclusion is warranted by the fact that the group exhibits gender and DOM and presents as such an invaluable opportunity for reasons of comparison. Data for all groups was transcribed using the CHILDES system. The Asháninka group data was transcribed using ELAN and PRAAT.

6 Data sets and results

In this Section, we present the comparative results for the features under study of the three bilingual groups and the three monolingual individuals followed by a discussion of these results and brief conclusion.

6.1 Data sets and coding

The size of the data set in terms of individual tokens in accordance with the distribution of clitic doubling and dislocated structures across all four groups as well as the overall tokens is shown in Table 3 below.

Table 3: Distribution of clitic doubling and dislocated structures across all groups.

	CLD (%)	CLLD (%)	CLRD (%)	Total
Shipibo	91% (62t)	3% (2t)	6% (4t)	100% (68t)
Asháninka	72% (73t)	23% (23t)	5% (5t)	100% (101t)
Huánuco	77% (40t)	13% (7t)	10% (5t)	100% (52t)
LSCV	54% (76t)	26% (38t)	20% (28t)	100% (142t)

For all four groups all transitive verbs were coded according to each one of the CLD, CLLD and CLRD structures for i) clitics and absence and presence of DOM, and ii) animacy, definiteness and thematic role of the object. The individual occurrences for each structure and the overall occurrences are shown in Table 3. In terms of animacy we counted humans and animals as animate and everything else as inanimate.⁵ Objects with a definite determiner were coded as definite and objects with an indefinite determiner or bare nouns were coded as indefinite. Thematic roles were coded depending on transitivity of the verbs, e.g. PATIENTS undergo visible changes and THEMES remain unchanged. Doubtful cases that could not be resolved by double checking audio files and PRAAT were excluded from the analysis all together.

6.2 Overall production of DOM

As shown in Figure 2, all bilingual groups and the monolingual individuals produce DOM to varying degrees and pattern in interesting ways. Expression of DOM in the Huánuco Quechua-Spanish dataset is strikingly different from the other data sets by marking 75% of all direct objects produced within clitic structures and leaving 25% unmarked. LSCV exhibits a similar distribution with 66% marked and 34% unmarked direct objects. Data from rural Shipibo-Spanish and rural Asháninka-Spanish show a different scenario. The Shipibo group exhibits the highest percentage of unmarked DPs (71%) over DOM marked DPs (29%), closely followed by the Asháninka Spanish group exhibiting a preference for unmarked DPs although less pronounced (56% unmarked vs. 44% marked). Overall, the groups show a significant difference in their distribution ($\chi^2(3, N = 400) = 7.8, p < .005$).

⁵ We acknowledge that animacy in indigenous languages may have other scales.

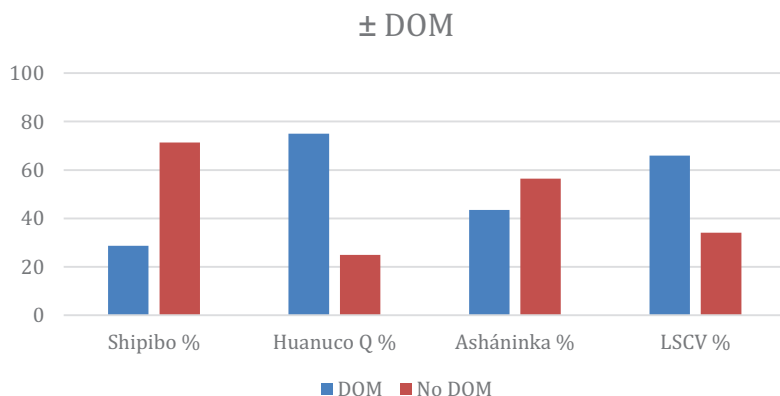


Figure 2: Expression of DOM across groups in CLD, CLLD and CLRD structures.

When comparing the production of DOM in clitic doubling constructions in Asháninka-Spanish bilinguals and Quechua-Spanish bilinguals, a general linear model showed a difference ($\beta = 3.05$, $SE = .43$, $z = 7.05$, $p < .001$). That difference was not found when comparing the Shipibo-Spanish and the Asháninka-Spanish bilinguals ($\beta = .34$, $SE = .41$, $z = .81$, $p = .42$).

6.3 DOM and conditioning factors

Looking at three of the data sets (Huánuco Quechua, Asháninka, and Shipibo)⁶ a general lineal model was fitted for definiteness, human, animacy and thematic role as fixed factors, in order to determine their effect on DOM production. There was a main effect of definiteness ($\beta = 2.24$, $SE = .49$, $z = 4.55$, $p < .001$), human ($\beta = 2.25$, $SE = .62$, $z = 3.66$, $p < .001$), animacy ($\beta = 3.01$, $SE = .51$, $z = 5.90$, $p < .001$), and thematic role ($\beta = -0.59$, $SE = .27$, $z = -2.19$, $p = .028$). Given these main effects found for each conditioning factor, in this Section, we present the distribution of DOM according to these factors to the exclusion of human, which we will treat as subsumed under animate.

⁶ The LSCV individuals data were not included in the model.

6.3.1 DOM and definiteness

The distribution of DOM with definite and indefinite DPs according to group is provided in Figure 3 below. The Huánuco Quechua group exhibits the highest percentage of DOM with definite DPs (73%), followed by the LSCV monolinguals (61.4%), the Asháninka group (41.6%), and the Shipibo group (35.8%).

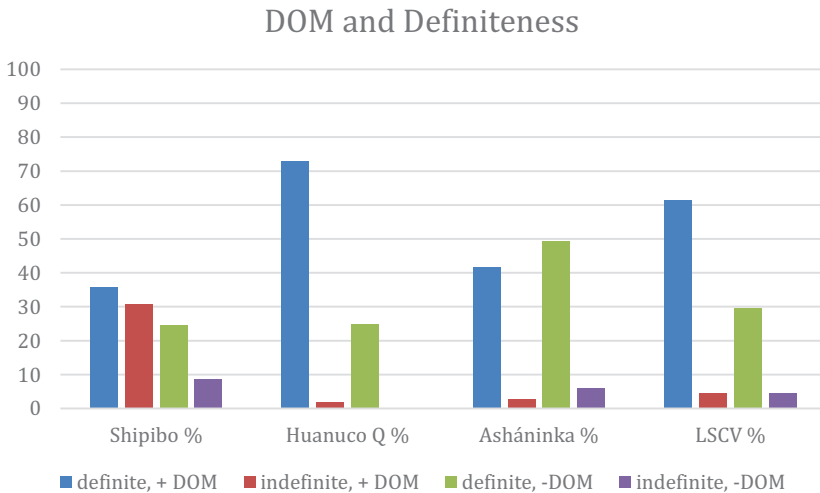


Figure 3: DOM with (in)definiteness across groups.

The Asháninka group shows the highest percentage of unmarked definite DPs (49.5%), followed by the LSCV (29.5%) and the Shipibo group (24.7%). The Huánuco Quechua group and the Shipibo group's percentage of unmarked definite DPs are practically equal and the lowest of all groups (25% and 24.7% respectively). Extension of DOM to indefinites is highest in the Shipibo group (30.9%) followed by very low numbers from LSCV (4.5%), the Asháninka group (3%) and the Huánuco Quechua group (1.9%). Again, independence of preference is proven by a chi-square test $\chi^2(6, N = 300) = 12.5, p <<.0003$.

A general linear model (excluding LSCV data) found interactions between definiteness and language in the case of Huánuco Quechua ($\beta = 5.02, SE = 0.87, z = 5.7, p = 0.00$) and Shipibo ($\beta = 2.39, SE = 1.02, z = 2.34, p > 0.01$). When compared to the Asháninka data, the probability that a speaker of Huánuco Quechua or Shipibo would produce DOM with a definite DP is higher, especially among the Huánuco Quechua speakers.

The following examples illustrate the use, extension or lack of DOM with definite and indefinite DP's representative of each of the language varieties under study.

DOM with definite DP in CLD

- (22) *Le ve-mos también a-l sapo*
 CL.3SG see-PRES.1PL also DOM-DET.M.SG toad.M.SG
 ‘We also see the toad.’ (H15, 82)
 Huánuco Quechua-Spanish (Sánchez Dataset 2005)

DOM with indefinite DP in CLD

- (23) *Sí, le he ayud-ado a una señora.*
 Yes CL.3SG have.PERF.1SG help-PART DOM INDEF.F.SG lady.F.SG
 ‘Yes, I have helped a Lady.’ (LSCV1)
 LSCV (Mayer Dataset 2006)

DOM with definite DP in CLRD

- (24) *Pued-e morder-le el perr-o a-l niñ-o*
 can-PRES.3SG bite.INF-CL.3SG DET.M.SG dog-M.SG DOM-DET.M.SG
 boy-M.SG
 ‘The dog can bite the boy.’ (H5,13)
 Huánuco Quechua-Spanish (Sánchez Dataset 2005)

DOM with definite DP in CLLD

- (25) *Yo a mi sobrin-ita la recib-ía acá*
 PP.1SG DOM POSS.1SG niece-DIM.F.SG CL.3.F.SG receive-IMP.F.3SG here
 ‘I received my niece here with me.’ (LSCV1)
 LSCV (Mayer Dataset 2006)

Lack of DOM with definite DP in CLD

- (26) *El niño lo agarr-ó Ø el sap-ito*
 the boy CL.3.M.SG grab-PERF.3SG DOM DET.M.SG toad-DIM.M.SG
 ‘The boy grabbed the little toad.’ (A 15)
 Asháninka-Spanish (Mayer Dataset 2016)

Lack of DOM with indefinite DP in CLD

- (27) *Le sac-ó un lor-o*
 CL.3SG took out-PERF.3SG INDEF.M.SG parrot-M.SG
 ‘(He) took out a parrot.’ (SH2)
 Shipibo-Spanish (Sánchez Dataset 2002)

Lack of DOM with definite DP in CLRD (the only occurrence)

- (28) *Un gente le sac-amos la caja*
 INDEF.SG person CL.3SG take.out-PRES.1PL DET.F.SG box.F.SG
Ø este su lor-o
 DOM DEM.M.SG POSS parrot-M.SG
 ‘Somebody took this parrot out of the box.’

Shipibo-Spanish (Sánchez Dataset 2016)

Lack of DOM with definite DP in CLLD

- (29) *Y el sap-ito lo llev-ó*
 and DET.M.SG toad-DIM.M.SG CL.3.M.SG carry-PERF.3SG
Ø la tortuga
 DOM DET.F.SG turtle.F.SG
 ‘And the little toad carried the turtle.’ (PO41)

Asháninka-Spanish (Mayer Dataset 2016)

6.3.2 DOM and animacy

In the overall results for animacy patterns shown in Figure 4, the Asháninka group leads with the highest percentage of animates (84.3%) and lowest for inanimates (15.7%) followed by the same pattern by the Shipibo group with higher frequency of animates (61%) than inanimates (39%). The Huánuco Quechua group presents a similar pattern to the LSCV monolinguals with a higher production of inanimates (57%) than animates (43%). The monolingual individuals exhibit a reversal from the definiteness results with a higher (59.6%) percentage of inanimate than animate DPs (40.4%). Again, independence of preference is corroborated by $\chi^2(3, N = 400) = 7.8, p <<.0067$.

The distribution of DOM and animacy for the four groups shows that in all four groups, DOM occurs with animate DPs led by the Huánuco Quechua group with a very high result (71.2%), followed by the LSCV group (56.1%) and the Asháninka group (38.6%), with the Shipibo group presenting the lowest percentage (28.7%).

The Shipibo group stands out from the other three groups with close percentages for DOM with animate DPs (28.7%) and lack of DOM with animate DPs (23.5%). There is no overextension of DOM to inanimates in the Shipibo group and inanimates without DOM reach 47.8%. All other groups extend DOM to a very low extent to inanimates led by LSCV (9.8%), followed by Asháninka (5%) and Huánuco Quechua (3.8%) and again, a chi-square test corroborates independence of preference $\chi^2(6, N = 300) = 12.5, p <<.0029$.

Animacy and DOM

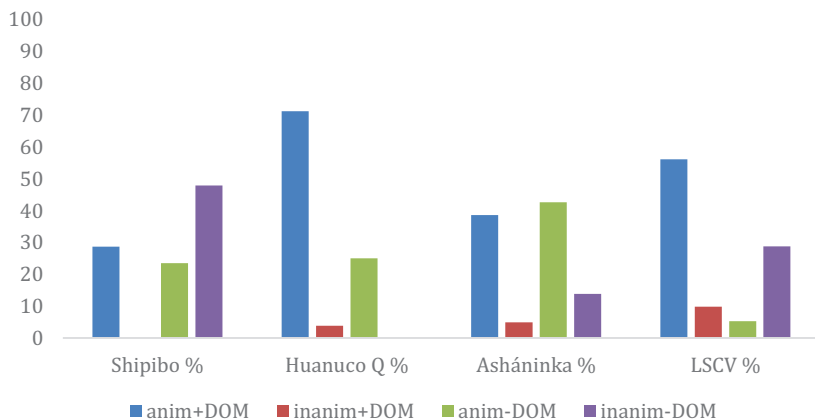


Figure 4: Comparative data for DOM and animacy across groups.

The following examples illustrate the unexpected patterns of lack of DOM with animate DPs (30), overextension of DOM to inanimate DPs (31) and lack of DOM with animate DP in CLLD (32).

Lack of DOM with animate DP in CLLD

- (30) *El niñ-o le riñ-e Ø ese perr-o*
 The boy-M.SG CL.3SG tell.off-PRES.3SG DOM DEM.M.SG dog-M.SG
 ‘The boy tells off that dog.’

Shipibo-Spanish (Sánchez Dataset 2002)

DOM with inanimate DP in CLLD

- (31) *A-l aji la hag-o hervir*
 DOM-DET.M.SG chili.M.SG CL.3.F.SG make-PRES.1SG boil-INF
 ‘I’ll boil the chili.’(LSCV2)

LSCV (Mayer Dataset 2006)

Lack of DOM with animate DP in CLLD

- (32) *El otr-o le han dej-ado cre-o*
 DET.M.SG other-M.SG CL.3SG PERF.3PL left-PART think-PRES.1SG
 ‘The other one, I think they left him.’ (PO50)

Asháninka (Mayer Dataset 2016)

The only instances of DOM marking inanimate DPs in CLRD and CLLD dislocated structures were found in the monolingual LSCV data in LSCV1. The other three bilingual groups did not show extension of DOM to those information structures.

DOM with inanimate DP in CLLD

- (33) *Por ejemplo, yo a-l fresco*
 for example PRO.1SG DOM-DET.M.SG juice.M.SG
le he licu-ado
 CL.3SG have-PAST.1SG blend-PART

‘For example, I blended the fresh juice.’ (LSCV1)

LSCV (Mayer Dataset 2006)

Within unmarked animate DPs, we looked at human DPs, usually ranked highest in the animacy hierarchy, and we also found evidence of lack of DOM with human DPs as illustrated in (34) and (35) for CLD constructions and (36) and (37) for CLLD and CLRD respectively.

Lack of DOM with human DP in CLD

- (34) *Y lo agarr-a Ø el niño*
 and CL.3.M.SG grab-PRES.3SG DOM DET.M.SG boy-M.SG
 ‘And he grabbed the boy.’ (H16, 98)

Huánuco Quechua-Spanish (Sánchez Dataset 2005)

- (35) *El tortug-a le agarr-a Ø el niño*
 DET.M.SG turtle-F.SG CL.3SG grab-PRES.3SG DOM DET.M.SG boy-M.SG
en su pantalón
 PREP POSS trouser

‘The turtle grabs the boy by his trousers.’ (PO, 43)

Asháninka-Spanish (Mayer Dataset 2016)

Lack of DOM with human DP in CLLD

- (36) *Ella yo le llev-é a Iquitos*
 PP.3.F.SG PP.1SG CL.3SG take-PERF.1SG LOC Iquitos
 ‘The other one, I took her to Iquitos.’ (LSCV1)

LSCV (Mayer Dataset 2006)

Lack of DOM with human DP in CLRD

(37) *Cuando ella la cort-aban a la barriga*
 when PP.F.SG CL.F.SG cut-IMP.F.3PL DOM DET.F.SG tummy.F.SG
 ‘When they cut her tummy.’ (LSCV2)

LSCV (Mayer Dataset 2006)

In terms of DOM as a function of animacy and language, the linear model found animacy to be a predictor of DOM ($\beta = 4.05$, $SE = 0.39$, $z = 10.26$, $p < 0.01$) and Huánuco Quechua was also found to be a predictor of DOM ($\beta = 2.43$, $SE = 0.36$, $z = 6.62$, $p < 0.01$).

6.3.3 DOM and thematic roles

The data in Figure 5 show a preference for PATIENTS (64%) over THEMES (36%) overall. However, the Huánuco Quechua group defies this trend by a clear preference for THEMES (66%) over PATIENTS (34%). The Shipibo, LSCV and Asháninka groups show the opposite pattern with the Shipibo group leading the pattern with a preference for PATIENTS (76.7%) over THEMES (23.2%), followed by the Asháninka groups strongly preferring PATIENTS (71%) over THEMES (29%). The LSCV group follows this distribution closely with similar percentages for PATIENTS (61.4%) over THEMES (38.6%). Independence of preference was clearly established by a chi-square test $\chi^2(6, N = 300) = 12.5$, $p < .005$.

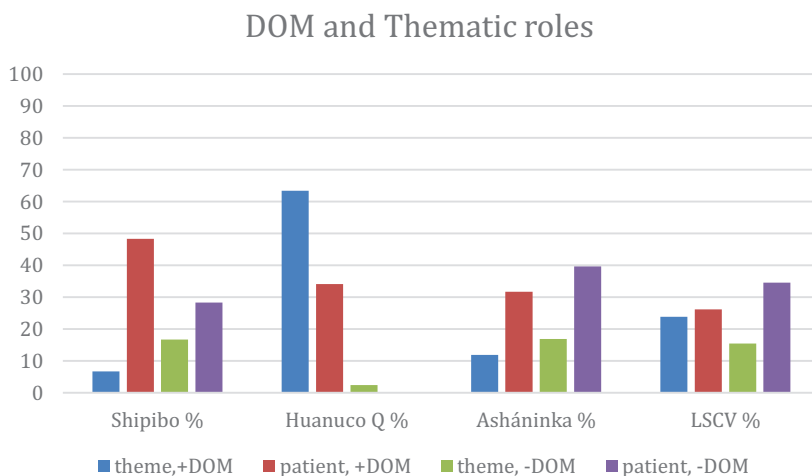


Figure 5: Distribution of thematic roles across groups.

When looking at the distribution of DOM marking according to thematic role, we find that the Huánuco Quechua group has a unique pattern in which the frequency of direct objects with theme roles marked with DOM is the highest (63.4%) followed by those with patient roles with DOM (34.1%) and by unmarked direct objects with a theme role (2.4%). Unlike the Huánuco Quechua group, the Shipibo group shows a higher frequency of DOM marking with PATIENTS (48.3%), than THEMES (6.7%). The Asháninka group shows a similar pattern of preference for DOM with PATIENTS (31.7%) over THEMES (11.9%). Unlike the other groups, the LSCV individuals do not exhibit a clear preference for either and they only show a slightly higher frequency of DOM with PATIENTS (26.2%) than THEMES (23.8 %). The following examples illustrate the most frequent cases of DOM with THEME and PATIENT arguments in each group.

DOM with theme argument in CLD

- (38) *Y el tortuga también le*
 and DET.M.SG turtle.F.SG also CL.3SG
mir-a a-l niñ-o
 look-PRES.3SG DOM-DET.M.SG boy-M.SG
 ‘And the turtle also looked at the boy.’ (H16)
 Huánuco Quechua-Spanish (Sánchez Dataset 2005)

DOM with patient argument in CLD

- (39) *Le muerde a-l lor-ito*
 CL.3SG bite-PRES.3SG DOM-DET.M.SG parrot-DIM.M.SG
 ‘(He) bites the little parrot.’
 Shipibo-Spanish (Sánchez Dataset 2002)

DOM with theme argument in CLD

- (40) *Est-á mir-ándo-le a-l turtuga*
 AUX-PRES.3SG look-GER-CL.3SG DOM-DET.M.SG turtle.F.SG
 ‘He is looking at the turtle.’(PO43)
 Asháninka-Spanish (Mayer Dataset 2016)

DOM with theme argument in CLD

- (41) *Tien-e que est-ar viéndo-la a su*
 has-PRES.3SG that be-INF look.after-CL.3.F.SG DOM POSS
hij-it-a
 daughter-DIM-F.SG
 ‘She has to keep looking after her little daughter.’ (LSCV1)
 LSCV (Mayer Dataset 2006)

For the coding of information structure, the Huánuco Quechua group shows no preference for either of the roles (for 7 CLLD 3 THEMES and 4 PATIENTS and 5 CLRD 2 THEMES and 3 PATIENTS) while the Shipibo group only marks PATIENTS in 2 instances of CLLD and even distribution of both thematic roles for 4 instances of CLRD. The last bilingual group, Asháninka-Spanish differs from both previous groups in terms of high numbers for CLLD and a pronounced preference for PATIENTS (21/23) to mark those with no clear preference for CLRD (3 THEMES, 2 PATIENTS) in (44).

DOM with theme argument in CLRD

- (42) *Y aquí el niño-to le está mir-ando*
 and here DET.M.SG boy-DIM.M.SG CL.3SG is.PRES.3SG look-PART
molesto al sap-ito
 annoyed DOM-DET.M.SG toad-DIM.M.SG
 ‘And here the boy is looking annoyed at the little toad.’ (H11)
 Huánuco Quechua (Sánchez Dataset 2005)

DOM with patient argument in CLLD

- (43) *Al, el perro eh le pate-a*
 DOM-DET.M.SG DET.M.SG dog-M.SG EH CL.3SG kick-PRES.3SG
 ‘The, the dog, ehem, he kicked him.’
 Shipibo (Sánchez Dataset 2002)

DOM with patient argument in CLLD

- (44) *Al otro sap-ito le hab-ían dej-ado*
 DOM-DET.M.SG other toad-DIM CL.3SG PERF-3PL leave-PART
porque era muy gruñón.
 because is.IMPF.3SG very grumpy
 ‘The other toad, they left him behind he was very grumpy.’ (PO 30)
 Asháninka Spanish (Mayer Dataset 2016)

The monolingual group shows a clear preference for PATIENTS in both information structures with 26:10 in CLLD (45) and less so in CLRD with 17:11 in (46).

DOM with patient argument in CLLD

- (45) *A la yuca sí Ø pel-o*
 DOM DET.F.SG yuca.F.SG yes peel-PRES.1SG
 ‘The yuca, yes, I peel it.’ (LSCV1)

Mayer (Dataset 2016)

DOM with patient argument in CLRD

- (46) *Est-á loc-o la abuel-a porque no lo*
 is-PRES.3SG crazy-M.SG DET.F.SG grandmother-F.SG because not CL.3.M.SG
oblig-ó así a su hij-o cuando Eva nac-ió
 oblige-PERF.3SG thus DOM POSS son-M.SG when Eva born-PERF.3SG
 ‘She is crazy, the grandmother, because she did not oblige her son this
 way when Eva was born.’ (LSCV2)

Mayer (Dataset 2016)

Huánuco Quechua was also a predictor for DOM, as a function of thematic role and language ($\beta = 0.94$, $SE = 0.33$, $z = 2.86$, $p = 0.00$), but no interactions were found. The monolingual group shows a more extended spread of features in terms of frequency.

The group shows a common strategy with DOM occurring most preferably with definite and animate objects. This is followed by a pronounced drop in frequency for unmarked patient, definite, and inanimate objects. The following bars show an array of combinations including DOM with patient and theme objects, as well as unmarked themes. Although with very low frequencies, these data also show animates without DOM and inanimates with DOM which reflect the differences between the three LSCV speakers. While LSCV3, the Lima-born youngest speaker only marks animate and definite objects, LSCV2 exhibits DOM with inanimate objects to a lesser degree than LSCV1. Lack of DOM with animates occurs in both to the same degree.

7 Discussion

The results we have presented above show that all factors under study play a role in determining DOM in the bilingual contact varieties under study and to some extent in the monolingual contact data. However, their effect differs across groups. Typological differences between the languages with which Spanish is in contact also play an important role as they do result in different frequency patterns. We interpret these results as the outcome of a complex interaction between contact with typologically distinct languages for the bilingual groups and language-specific ecological factors for all groups but especially for the monolingual individual data that exhibit the greater level of variability.

In response to our first research question, where we asked a) whether differences found in Spanish DOM patterns in information structures depended on typological characteristics of the contact language, and b) whether monolingual

varieties showed differences as well, our hypothesis was born out. We find that DOM patterns in clitic doubling and clitic dislocated structures in contact varieties are affected by differences in the configuration of the morphology of the contact language, as shown by the emerging Differential Object Marking systems across the bilingual and the monolingual group in Table 4.

Table 4: Emerging Differential Object Marking systems.

	+DOM	-DOM
Huánuco Quechua-Spanish	DEF > ANIM > THEME > PATIENT 73.1% > 71.2% > 63.4% > 34.1%	DEF = ANIM > THEME > PATIENT 25% = 25% > 2.4% > 0%
Shipibo-Spanish	PATIENT > DEF > ANIM > THEME 48.3% > 35.8% > 28.7% > 6.7%	PATIENT > DEF > ANIM > THEME 28.3% > 24.7% > 23.5% > 16.7%
Asháninka-Spanish	DEF > ANIM > PATIENT > THEME 41.6% > 38.6% > 31.7% > 11.9%	DEF > ANIM > PATIENT > THEME > 49.5% > 42.6% > 39.6% > 16.8%
LSCV-monolinguals	DEF > ANIM > PATIENT > THEME 61.4% > 56.1% > 26.2% > 23.8%	PATIENT > DEF > THEME > ANIM 34.5% > 29.5% > 15.5% > 5.3%

As hypothesized, the Quechua group data exhibits not only a higher frequency of DOM when compared to the other groups, but also a very different scale with very low frequency of lack of DOM with the semantic features definite and animate, and the thematic role *THEME*. DOM in Huánuco Quechua is linked to the semantic features definite and animate followed by the thematic roles with a prominent result for *THEME*. This result can be attributed to the similarities Huánuco Quechua and Spanish share through a common nominative/accusative alignment system and the availability of an accusative case marker. By not having to focus on differences in case marking Huánuco Quechua speakers can focus on definiteness and animacy. We take this to indicate that morphological differences between the preposition *-a* in Spanish and the Quechua case suffix *-ta* have hardly any effect on the production of DOM. Huánuco Quechua-Spanish bilinguals seem to be able to successfully map the relevant features onto a DOM marker. These results question the assumption that morphology is the locus of second language acquisition difficulties as proposed by the Bottleneck Hypothesis. The data show that differences in morphological patterns across languages do not seem to be a barrier for the acquisition of DOM marking.

Our expectation that the Shipibo and the Asháninka groups will exhibit lower frequencies of DOM due to differences in alignment between these languages and Spanish are born out as well. Of the two Amazonian languages, we see a higher frequency of DOM with definite and animate features in the Asháninka group followed closely by the thematic role *PATIENT* over *THEME*. The fact that

Asháninka is a nominative/accusative language albeit with fluid/split transitivity and the availability of gender specific bound morphemes as argument markers may be factors that support the acquisition of DOM in the relevant clitic structures. However, unlike the Huánuco Quechua group, the Asháninka group also exhibits lack of DOM with definite, animate and patient DPs (49.5%, 42.6%, and 39.6% respectively) and to a lesser frequency with theme DPs (16.8%). We take this to suggest that the fact that it is a language with fluid/split transitivity may also result in greater difficulty in the acquisition of DOM despite some sensitivity to the semantic features we have studied.

Shipibo-Spanish shows a different distribution from Asháninka-Spanish for marked and unmarked objects. The distribution of DOM follows a clear preference of the thematic role *PATIENT* followed by the definite semantic feature over animate and *THEME*, the latter only to a small extent. Interestingly, the same scale is found when looking at lack of DOM but with lower frequencies. A possible cause for a marking hierarchy that ranks the thematic role *PATIENT* higher than all other features in Shipibo could be the fact that Shipibo speakers are beginning to pay attention to *PATIENTS* as objects as a way to overcome the fact that in Shipibo, an ergative language, the subject of a transitive sentence receives case marking while the object remains unmarked. Unlike the Huánuco Quechua and the Asháninka groups who can focus their attention directly on definiteness and animacy, Shipibo speakers must first focus on the thematic role *PATIENT* as a way of identifying the direct object as the recipient of case marking.

Monolingual Spanish (LSCV) shares a similar distribution to the Asháninka-Spanish group in terms of marked objects with higher percentages in each category for LSCV and it shares a similar distribution with the Shipibo Spanish group in terms of unmarked objects although with a reversal in the ranking of *THEME* and animacy at the end of the scale. In LSCV the features definite and animate (61.5% and 56.1% respectively) play a stronger role in marking over the thematic roles (26.2% and 23.8% respectively).

For the second research question asking for the effect of animacy, definiteness and thematic role on the presence or absence of DOM in bilingual Spanish clitic doubling and dislocated structures in contact bilingual and monolingual varieties, our results show that all factors have an effect although their frequency varies in each of the groups.

Our hypothesis that animacy has an important role in the acquisition of Spanish was partly confirmed as there was a main effect for animacy and it is a predictor of DOM, but its relevance in terms of frequency is mostly found in the Huánuco Quechua group. Despite its presence in the input, animacy is the third most frequent feature in DPs with DOM among the Shipibo group and the second ranked feature among the Asháninka and the LSCV individuals. The defi-

nite feature plays a role in DOM as an effect for definiteness was found. It was also ranked highest in the Asháninka and the Huánuco Quechua groups as well as among the LSCV individuals and second in the Shipibo group. We would like to propose that the results for two of the bilingual groups with higher and similar frequencies for definiteness and animacy (Huánuco Quechua and Asháninka) point to an increase marking information structure in conjunction with a two-dimensional scale of animacy and definiteness (Aissen 2003) in a given discourse situation. This does not seem to be the case for Shipibo Spanish speakers given the higher frequency of DOM with patient DPs.

Our expectation for thematic role to play a more salient role in Asháninka Spanish given the sensitivity to thematic role in the morphology of Asháninka was met in terms of frequency. The Asháninka and Shipibo groups lead the preferences for PATIENT over THEME (71.2% and 63.6% respectively) followed by lower frequencies of preferences by the LSCV and Huánuco Quechua groups (38.3% and 23.6% respectively). The higher frequency of patient roles over themes (Dowty 1991) in the former groups can be related to sensitivity to thematic roles in the case of Asháninka and to the higher frequency of production of patient direct objects as well as the need to identify the direct object as a bearer of marking in the Shipibo group.

Finally, the extension of DOM to indefinites among Shipibo speakers seems to reflect the fact that there are no definite determiners in this language and this group seems to be more inclined to focus on the thematic role PATIENT as the semantic feature driving DOM rather than on definiteness. The extension of DOM to inanimates among the Asháninka speakers could be an effect of the low frequency of inanimates in their data. There is some extension of DOM to indefinites in the data of individuals of the LSCV group. These kinds of extensions of DOM can be linked to the fact that LSCV speakers are located on a continuum dependent on variable access to formal instruction.

In relation to the ecological factors, we propose that the scalar DOM systems (Table 4) found across the bilingual groups and the monolingual individuals can be linked to a feature pool with a reduced subset of features shaped by each group individually in accordance with their communicative needs (Mufwene 2001; 2002). In the case of the bilingual groups, the typology of the first languages contributes significantly to the ability to determine the relevance of semantic features as triggers of DOM in clitic doubling constructions.

The monolingual group exhibits clear intragroup differences with a scalar spread from LSCV1 most removed from the prestigious Lima norm, LSCV2 closer to the norm and LSCV3 – born and raised in Lima representing the prestigious norm. The variability found among the three monolingual Spanish speakers demonstrates clearly the choices individual speakers make dependent on the

availability and mix of features they are exposed to in the pool. It also shows that variability naturally exists within communities and reflects individual and group-specific communicative needs (Matras 2010).

For the bilingual group, the feature pool would be shaped by language ecological factors such as contact with language particular alignment systems, input and stage of language acquisition, and social networks. In bilingual systems, functional features from typologically different languages compete against each other and features shared by both languages may converge towards sets of features combined from both languages into a single matrix (Sánchez 2003). Our results as described above support our hypothesis that contact between typologically different languages influences the outcome for DOM in bilingual Spanish. While contact with alignment systems closer to Spanish, as in the case of Huánuco Quechua, strongly support and facilitate the production of DOM based on a two-dimensional scale of the semantic features definiteness and animacy, contact with ergative/absolutive alignment in Shipibo yields a very different system in terms of ranking of features. Asháninka, being closer to Spanish in terms of nominative-accusative alignment but exhibiting differences in marking transitivity, shows less sensitivity to the semantic features and more to transitivity, basing DOM primarily on the thematic role *PATIENT*.

Apart from these typological differences, the feature pool is particularly susceptible, and constrained by access to formal instruction in either or both languages as well as input from their social networks. As Mufwene (2001, 52) argues, in bilingual community languages, the selection from the feature pool works in the same ways as for idiolects. The feature pool represents a space where languages or dialects – varieties of Spanish in the case of the monolingual group – coexist, are acquired and used individually by speakers with differences across individuals that may not necessarily be reflected in the community language.

8 Concluding remarks

Our study provides preliminary evidence of typological effects in the development of DOM in clitic doubling structures. It shows that similarity in nominative-accusative alignment in Huánuco Quechua and Spanish favours a higher frequency of production of DOM in Spanish among the Huánuco Quechua speakers than among the Shipibo group and the Asháninka group. We advance the idea that this different distribution could be the result of Shipibo being a language with predominance of ergative-absolutive alignment and Asháninka one with nominative-accusative alignments but also split or fluid transitivity. In addition to case alignment, seman-

tic features such as definiteness and animacy as well as thematic role are present in all groups as determinants of DOM. Definiteness and animacy are more frequent in the Huánuco Quechua and Asháninka groups. We attribute their lower frequency in the Shipibo group to the fact that Shipibo speakers face the challenge of first identifying the constituent with a thematic role PATIENT as potentially subject to special morphological marking. Our data point in the direction that in contact situations typological similarities play a large role but the development of sensitivity to definiteness, animacy and thematic roles is not blocked even when there are typological differences. In the case of the LSCV individuals, we found evidence of higher levels of lack of DOM than among the Huánuco Quechua group, a distribution that seems to be better accounted for by the differences in the ecological factors that determine difference at the idiolectal level.

Bibliography

- Aissen, Judith, *Differential Object Marking. Iconicity vs. economy*, *Natural Language and Linguistic Theory* 21:3 (2003), 435–483.
- Alsina, Alex, *The role of argument structure in grammar. Evidence from Romance*, Stanford, Center for the Study of Language and Information, 1996.
- Blake, Barry J., *Case*, Cambridge, Cambridge University Press, 2001.
- Bossong, Georg, *Differential Object Marking in Romance and beyond*, in: Wanner, Dieter/ Kibbee, Douglas A. (edd.), *New analyses in Romance linguistics. Selected papers from the Linguistic Symposium on Romance Languages XVIII, Urbana-Champaign, April 7–9, 1988*, Amsterdam/Philadelphia, John Benjamins, 1991, 143–170.
- Bossong, Georg, *Nominal and/or verbal marking of central actants*, in: Fiorentino, Giuliana (ed.), *Romance Objects. Transitivity in Romance languages*, Berlin/New York, de Gruyter, 2003, 17–48.
- Bresnan, Joan/Aissen, Judith, *Optimality and functionality. Objections and refutations*, *Natural Language and Linguistic Theory* 20 (2002), 81–95.
- Company Company, Concepción, *Multiple dative-marking grammaticalization. Spanish as a special kind of primary object language*, *Studies in Language* 25:1 (2001), 1–47.
- Company Company, Concepción, *Transitivity and grammaticalization of object. The diachronic struggle of direct and indirect object in Spanish*, in: Fiorentino, Giuliana (ed.), *Romance objects. Transitivity in Romance languages*, Berlin/New York, de Gruyter, 2003, 217–260.
- Cuza, Alejandro/Pérez-Leroux, Ana Teresa /Sánchez, Liliana, *The role of semantic transfer in clitic drop among simultaneous and sequential Chinese-Spanish bilinguals*, *Studies in Second Language Acquisition* 35:1 (2013), 93–125.
- Dalrymple, Mary/Nikolaeva, Irina, *Objects and information structure*, Cambridge, Cambridge University Press, 2011.
- Dowty, David, *Thematic proto-roles and argument selection*, *Language* 67:3 (1991), 547–619.
- Givón, Talmy, *Topic, pronoun, and grammatical agreement*, in: Li, Charles N. (ed.), *Subject and topic*, New York, Academic Press, 1976, 149–188.

- Givón, Talmy (ed.), *Grammatical relations. A functionalist perspective*, Amsterdam/Philadelphia, John Benjamins, 1997.
- Guijarro-Fuentes, Pedro, *Feature composition in Differential Object Marking*, *EUROSLA Yearbook* 11:1 (2011), 138–164.
- Guijarro-Fuentes, Pedro, *The acquisition of interpretable features in L2 Spanish. Personal “a”*, *Bilingualism* 15:4 (2012), 701–720.
- Iemmolo, Giorgio. *Towards a typological study of Differential Object Marking and Differential Object Indexation*, Pavia, University of Pavia dissertation, 2011.
- Jensen, Isabel N./Slabakova, Roumanya/Westergaard, Marit, *The Bottleneck Hypothesis in second language acquisition. A study of L1 Norwegian speakers’ knowledge of syntax and morphology in L2 English*, in: LaMendola, Maria/Scott, Jennifer (edd.), *Proceedings of the 41st annual Boston University Conference on Language Development*, Somerville, Cascadilla Press, 2017, 333–346.
- Laca, Brenda, *El objeto directo. La marcación preposicional*, in: Company Company, Concepción (ed.), *Sintaxis histórica de la lengua española. Primera parte: La frase verbal*, vol. 1, México DF, UNAM/Fondo de Cultura Económica, 2006, 423–475.
- Lapesa, Rafael, *Estudios de morfosintaxis histórica del español*, Madrid, Gredos, 2000.
- Lardiere, Donna, *Dissociating syntax from morphology in a divergent L2 end-state grammar*, *Second Language Research* 14:4 (1998), 359–375.
- Lardiere, Donna, *On morphological competence*, in: Dekydtspotter, Laurent/Sprouse, Rex A./Liljestrang, Audrey (edd.), *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference*, 2005, 78–192.
- Leonetti, Manuel, *Specificity in clitic doubling and in Differential Object Marking*, *Probus* 20:1 (2008), 33–66.
- Loriot, James/Lauriault, Erwin/Day, Dwight, *Diccionario shipibo-castellano*, Lima, Instituto Lingüístico de Verano, 1993.
- Matras, Yaron, *Contact, convergence, and typology*, in: Hickey, Raymond (ed.), *The handbook of language contact*, Oxford, Wiley-Blackwell, 2010, 66–85.
- Mayer, Elisabeth, *Clitics on the move. From dependent marking to split marking*, in: Butt, Miriam/Holloway King, Tracy (edd.), *Proceedings of LFG ’08 Conference*, Stanford, CSLI Publications, 2008, 352–437, <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/13/lfg08.html> [last access: 13.09.2021].
- Mayer, Elisabeth, *Spanish clitics on the move. Variation in time and space*, 14 vol., Berlin/Boston, de Gruyter, 2017.
- Mayer, Elisabeth/Sánchez, Liliana, *Object agreement marking and information structure along the Quechua-Spanish contact continuum*, *Revista Española de Lingüística Aplicada* 29:2 (2016), 544–581.
- Mayer, Elisabeth/Sánchez, Liliana, *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 (2017), 883–896.
- Mayer, Elisabeth/Sánchez, Liliana, *Feature selection in clitic expression in two bilingual Amazonian Spanish varieties*, in: Zariquiey, Roberto, et al. (edd.), *Spanish diversity in the Amazon. Dialect and language contact perspectives*, Leiden/Boston, Brill, in press.
- Mayer, Mercer/Mayer, Marianna, *One frog too many*, New York, Puffin, 1992.

- Melis, Chantal, *El objeto directo personal en “El Cantar de Mio Cid”*. *Estudio sintáctico-pragmático*, in: Pensado, Carmen (ed.), *El complemento directo preposicional*, Madrid, Visor, 1995, 133–164.
- Mihas, Elena, *A grammar of Alto Perené (Arawak)*, Berlin/Boston, de Gruyter, 2015.
- Montrul, Silvina, *The acquisition of Spanish. Morphosyntactic development in monolingual and bilingual L1 acquisition and adult L2 acquisition*, Amsterdam/Philadelphia, John Benjamins, 2004.
- Mufwene, Salikoko, *The ecology of language evolution*, Cambridge, Cambridge University Press, 2001.
- Mufwene, Salikoko, *Competition and selection in language evolution*, *Selection* 3:1 (2002), 45–56.
- Næss, Åshild, *What markedness marks. The markedness problem with direct objects*, *Lingua* 114:9 (2004), 1186–1212.
- Nichols, Johanna, *Head-marking and dependent-marking grammar*, *Language* 62:1 (1986), 56–119.
- Payne, Judith K./Payne, David L., *The pragmatics of split intransitivity in Ashéninka*, *Revista Latinoamericana de Estudios Etnolingüísticos* 10 (2005), 37–56.
- Putnam, Michael T./Sánchez, Liliana, *What’s so incomplete about incomplete acquisition? A prolegomenon to modeling heritage language grammars*, *Linguistic Approaches to Bilingualism* 3:4 (2013), 478–508.
- Rodríguez-Mondoñedo, Miguel, *The acquisition of Differential Object Marking in Spanish*, *Probus* 20:1 (2008), 111–145.
- Sánchez, Liliana, *Quechua-Spanish bilingualism. Interference and convergence in functional categories*, Amsterdam/Philadelphia, John Benjamins, 2003.
- Sánchez, Liliana/Mayer, Elisabeth, *Clitics and argument marking in Shipibo Spanish and Asháninka-Spanish bilingual speech*, in: Fafulas, Stephen (ed.), *Amazonian Spanish. Language contact and evolution*, Amsterdam/Philadelphia, John Benjamins, 2020, 127–154.
- Sinnemäki, Kaius, *A typological perspective on Differential Object Marking*, *Linguistics* 52:2 (2014), 281–313.
- Slabakova, Roumyana, *Meaning in the second language*, 34 vol., Berlin/New York, de Gruyter, 2008.
- Suñer, Margarita, *Dialectal variation and clitic-doubled direct object*, in: Kirschner, Carl/DeCesaris, Janet A. (edd.), *Studies in Romance linguistics. Selected proceedings from the XVII Linguistic Symposium on Romance Languages*, Amsterdam/Philadelphia, John Benjamins, 1989, 377–395.
- Torrego, Esther, *The dependencies of objects*, Cambridge, Mass., MIT Press, 1998.
- Valenzuela, Pilar M., *Ergativity in Shipibo-Konibo, a Panoan language of the Ucayali*, in: Gildea, Spike/Queixalós, Francesc (edd.), *Ergativity in Amazonia*, Amsterdam/Philadelphia, John Benjamins, 2010, 65–96.
- Weber, David J., *Una gramática del quechua del Huallaga (Huánuco)*, Lima, Instituto Lingüístico de Verano, 1996.
- Witzlack-Makarevich, Alena/Seržant, Ilja A., *Differential Argument Marking. Patterns of variation*, in: Seržant, Ilja A./Witzlack-Makarevich, Alena (edd.), *Diachrony of Differential Argument Marking*, Berlin, Language Science Press, 2018, 1–40.
- Zdrojewski, Pablo/Sánchez, Liliana, *Variation in accusative clitic doubling across three Spanish dialects. Morphological variation in Spanish*, *Lingua* 151 (2014), 162–176.

