

Evangelia Adamou

# Replicating Spanish *estar* in Mexican Romani

**Abstract:** Based on first-hand data, this paper shows how Mexican Romani-Spanish bilinguals have replicated several uses of the Spanish *estar* in Romani, using the nearly obsolete 3d person subject clitic pronouns. The Romani subject clitics (*lo*, *la*, *le*) have become associated with the Spanish copula *estar* ‘to be’ in affirmative present tense clauses, thus restricting the uses of the native Romani copula. Moreover, the subject clitics have replicated the uses of *estar* in locative predications and in constructions with participles. These findings contribute to the general discussion over the complexification or simplification of languages in contact settings. It is argued that although the replication of Spanish *estar* has rendered Romani more complex, the resulting convergence may be considered as an overall simplification for the bilingual speakers (Matras 2009).

**Keywords:** language contact, copula, complexification, Spanish, Romani

---

**Evangelia Adamou:** Centre National de recherche scientifique (CNRS), Langues et civilisations à tradition orale (LACITO), UMR 7107, 7 rue Guy Moquet, 94801 Villejuif Cedex, France.  
E-mail: adamou@vjf.cnrs.fr

## 1 Introduction

This paper shows how Romani speakers living in Mexico over the past century have replicated Spanish *estar*, using the nearly obsolete subject clitic pronouns in a process of exaptation (Lass 1990). It is argued that under the influence of Spanish, the Romani subject clitics<sup>1</sup> developed a range of innovative uses: in 3rd person affirmative and present tense attributive predications (1a), locative predications (1b), and participle constructions (1c):

(1) a. *o raklo =lo felis*  
DEF.M boy 3SG.M happy

SPANISH: El niño está feliz.

‘The boy is happy.’

(Answer to the question: ‘Who is happy in this picture, the boy or the girl?’)

---

<sup>1</sup> Clitics are marked by the symbol =.

- b. *o zukel pafa eskamin =lo*  
 DEF.M dog beside chair 3SG.M  
 SPANISH: El perro está sentado al lado de una silla.  
 ‘The dog is next to a chair.’
- c. *mulo=lo*  
 dead-3SG.M  
 SPANISH: Está muerto.  
 ‘He’s dead.’

Replication of Spanish *estar* has added variation in contexts where the Romani copula was previously the dominant option and has thus rendered Romani grammar more complex. However, the resulting convergence between the grammatical systems in contact can be viewed as a simplification for the bilingual Romani-Spanish speakers (Matras 2009).

In this article I will begin by presenting some background information on Romani in Mexico (Section 2). I will then present the various uses of Romani clitics (Section 3). In Section 4, I will discuss the replication process and its effects on Romani.

## 2 Romani in Mexico

### 2.1 Historical and sociological background of Romani communities in Mexico

Today, Roma are settled in several Mexican cities, mainly in Mexico DF, but also in Tuxtla Gutierrez (State of Chiapas), Veracruz, Guadalajara, Puebla, as well as in the outskirts of the city of Oaxaca. Despite the significant presence of several thousand people, very little work exists on Mexican Roma and Romani, with the notable exception of Pickett (1962), Pickett and Gonzalez (1964) and a preliminary anthropological study by Ripka (2007). The data presented in this paper were collected in 2010, 2011 and 2012, in the Zapotec city Tlacolula de Matamoros in the State of Oaxaca (see map 1 in Figure 1). The Romani community of Tlacolula was traditionally itinerant and had more than 200 members in the 1950s. Nowadays, the Tlacolula community members are mostly scattered over several Mexican cities and only 30 of them still live in the area of Oaxaca.

Upon arrival in Mexico, the group is said to have worked in itinerant entertainment activities, e.g., traveling circuses and cinemas, some of the most common Roma activities in Mexico (Pickett 1962). The men also worked in coffee and



Fig. 1: The state of Oaxaca, Mexico

horses, and, more recently, cars. The women have always worked as fortune tellers, whether one hundred years ago or today. Group members are closely connected and intermarry with other Roma living in Mexico; marriages with outsiders are also frequent.

Mexican Roma are Catholic but some communities belong to Pentecostal churches (Ripka 2007). Roma women in their forties have usually had no formal education and are illiterate; the younger generation has attended primary school, often within the INEA system (Instituto Nacional para la Educación de los Adultos), which offers a more flexible structure for itinerant families.

Men, women, and sometimes children are bilingual in Romani and Mexican Spanish, but a strong tendency to shift to Spanish is observed, especially in families made up of both Roma and Mexicans. The most common language name is *ɣomanes* (adv.) or *ɣomaj tʃip* ‘Romani language’, characterized by the fricative [ɣ]. The group’s self-identification name is *ɣom* and among other denominations one finds *ungaros* (in Mexican Romani *ungarezuria*). The dialectological study, which will be presented in Section 2.2, clearly shows that the current Oaxacan variety has its roots in a Transylvanian dialect. The ancestors of the Tlacolula Roma probably began their migration from Transylvania towards Mexico in the middle of the nineteenth century, at a moment of major political changes in the area. At the time, Transylvania was part of Austria-Hungary, but it has since become part of Romania. According to oral traditions, the presence of Romani speakers in Mexico dates as far back as 1910, to the time of the Mexican Revolution; however, Pickett’s consultant, Dora Montes, indicates the presence of Romani communities in

Mexico even before 1860 (Pickett 1962: 8). The ancestors of the small Oaxacan community are said to have arrived via Honduras, Guatemala and Chiapas. Unfortunately, the Mexican census of 1910<sup>2</sup> does not identify the Romani communities as such and according to Ripka (2007: 12), no specific reference to Hungarian gypsies is made in the Archivo General de la Nación (unlike some references to *gitanos* from Greece who arrived between 1890–1950). Indication as to the country of origin, e.g., Austria-Hungary, does not constitute solid proof as Hungarian migration of non-Romani origin also took place during the same period (Szente-Varga 2009).

## 2.2 Dialect classification

Romani is an Indo-Aryan language spoken throughout Europe, in the Americas, and in Australia. The earliest migrants, who probably belonged to service-providing castes (Matras 2002), arrived from India during the period of the Byzantine Empire, around the tenth century. Romani was considerably influenced by Greek during this period. At the end of the Byzantine era, several groups migrated toward western and northern Europe and new contact languages were added.

The first dialect classifications were carried out by Miklosich (1872–1880), followed by the very influential work of Gilliat-Smith (1915). More recently, several studies have furthered and modernized Romani dialectology, e.g., Bakker and Matras (1997), Elšík (2000), Boretzky and Iglá (2004). In contemporary Romani linguistics, dialectal categorization based on the geographic spread of linguistic features has become dominant (Matras 2005, 2010). Matras (2005, 2010) identifies a cluster of isoglosses dividing the dialects of western and northern Europe from those of southeastern Europe. He further distinguishes four dialectal groups: Northern (subdivided into Northwest and Northeast), Central, Vlax, and Balkan Romani (Matras 2010: 55–56).

The Romani variety spoken in Oaxaca shares several features with the southeastern dialects of Europe and more specifically the Vlax group (Adamou 2011). According to Matras' classification, the Vlax group is dominant both in the Romanian-Moldavian geographic zone and in neighboring Serbia, Croatia and Hungary. It is also spoken in several areas farther afield, due to migrations from Transylvania and Banat starting in the nineteenth century (Matras 2010: 47, 55).

---

<sup>2</sup> Censo de Población de los Estados Unidos Mexicanos de 1910, available at <http://www.inegi.org.mx/>

Below are some features encountered in Oaxaca Romani that allow us to classify it as a Vlax variety.

- Oaxaca Romani has no initial jotation in words such as *anxo* ‘egg’ and *axo* ‘flour’, contrary to the northern branch of Romani, which shows “a preference for prothetic jotation in selected words, among them *jaro* ‘egg’” (Matras 2010: 49).
- The Romani speakers of Oaxaca do not truncate the initial syllable in words such as *akana* ‘now’, unlike speakers of the northern Romani dialects. Matras argues that syllable truncation in the northern branch was triggered by a shift to word-initial stress due to contact with German (Matras 2010: 49).
- Oaxaca Romani shows strengthening of initial segments through *a*-prothesis in words such as *anau* ‘name’ and *abijau* ‘wedding’, as observed in the Romani dialects of the area between Ukraine and northern Bulgaria and Serbia (Matras 2010: 50).
- The historical cluster \**ɲd* (Matras 2010: 83) is realized in Oaxaca Romani as *nɛ* in several words such as *manxo* ‘bread’, *anxo* ‘egg’, *kanxo* ‘thorn’, as opposed to the simplified realization *r*, which dominates in the northern branch (*maro* ‘bread’, *jaro* ‘egg’).
- Comparing the copula stem also identifies Oaxaca Romani as a southeastern dialect through the use of the *s*-stem (as opposed to *h*- in the northern dialects and individual dialects in Transylvania and in Greece, [Matras 2010: 44]); *me sim* ‘I am’, *tu san* ‘you are’, *vo si, voj si* ‘he is, she is’, *ame sam* ‘we are’, *tume san* ‘you are’, *von si* ‘they are’.
- The second plural form of the past perfective in *-an*, as in *tume xasardian* ‘you lost’, is a feature characteristic of the northwestern dialects and the central zone as opposed to the *-e* form which dominates in the northeastern dialects, and the *-en* form in the southernmost dialects (Matras 2010: 51).
- Moreover, the third singular past of intransitive verbs still shows gender distinction in Oaxaca Romani (e.g., *mulo/muli* ‘he/she died’), whereas this distinction has been lost in the northern dialects.
- The reduplicated demonstrative stem *k\_d*, in *kodja* ‘that’, *kodola* ‘those’, is common to contemporary dialects of Hungary and Slovakia according to Matras (2010: 53), while the stem *k\_k*, also attested in Oaxaca Romani e.g., in *kako* ‘this’, is common to dialects of northern Bulgaria and Romania.
- Several words of Romanian origin are still in use in Oaxaca Romani: *kopatfi* ‘tree’, *pruna* ‘plum’, *maira* and *maria* ‘sea’, as well as the modal ‘must’ of Slavic origin borrowed via Romanian as in *trobul te zav* ‘I have to go’, *trobusedjamas te azukeras* ‘we had to wait’.
- A few words of Slavic origin are also attested: *zeleno* ‘green’, *briame* ‘time’ < *vreme*, *lendala* ‘mirror’.

- A number of words of Greek origin, borrowed in the Early Romani period, are still in use in Oaxaca Romani: *eskamin* ‘chair, bench’, *foro* ‘city’, *luluži* or *luludži* ‘flower’, *papo* ‘grandfather’; numerals in Romani known to be borrowed from Greek such as *jefta*, *oxto*, *enja*, *trianda* ‘seven, eight, nine, thirty’; and verbs such as *ɣramol* ‘he/she writes’, *del angali* ‘he/she embraces’, *xolja-vol* ‘he/she’s upset’. Moreover, we note the loan verb marker of Greek origin *-isar* which is characteristic of Romani as spoken in Romania and Moldavia.
- Some lexical features shown by this dialect are the semantic shift from *naj* ‘nail’ to ‘finger’ (as opposed to *anguft* in the northern dialects), the use of *jilo* ‘heart’ (*gi* being dominant in the north), *koɹ* ‘neck’ (*men* being used in the north). These are features that distinguish this dialect from the varieties spoken in the southern Balkans where there is variation in the use of the items of the northern and the southern branches of Romani (Matras 2010: 50).

Based on Matras’ dialect classification (2005, 2010), it is therefore possible to categorize this dialect among the northern Vlax dialects, and to locate its origins in the area of Transylvania, in today’s Romania.

### 2.3 Contact with Spanish

Oaxaca Romani shows several borrowings from Spanish, and codeswitching is frequent.

- *Lexicon*. For some areas of the vocabulary, either Spanish or inherited materials are used, e.g., for place adverbials: *infrente* ‘in front’ in variation with the inherited form *anglal* ‘in front’. For others, the Spanish material is consistently used: this is the case for temporal expressions (days, months and seasons) and time adverbials, e.g., *despues* ‘after’. Several Spanish adjectives are used when speaking Romani: *felis* ‘happy’, *alegre* ‘happy’, *estupido* ‘stupid’, *mal* ‘bad’, *triste* ‘sad’, *agradable* ‘friendly’, *asul* ‘blue’. Interestingly, numerals are consistently Romani, including the numerals borrowed from Greek during the Byzantine era.
- *Morphology*. Some speakers integrate Spanish elements in the Romani morphology, i.e., *liono* ‘lion.M.SG’, while others do not, *lion* ‘lion’. In general, the Spanish lexicon takes Romani morphology: *kriatura* ‘child’; *kriaturenge* ‘children.DAT’. Masculine nouns from Spanish bear the Romanian plural *-uri* < *-uria* as is common for Vlax dialects: i.e., *amigo* ‘friend’ – *amiguria* ‘friends’; *nonito* ‘child’ – *nonituria* ‘children’. Loan verbs may be accompanied by a loan verb marker of Greek origin *-isar*, typical of Vlax dialects: *sufrisardia*

‘(he) suffered’ < Spanish *sufrir*. Adjective morphology sometimes seems to copy the Spanish morphology, as in *dragostoso* ‘dear’ which resembles the Spanish *amistoso* ‘friendly’. Despite this accidental formal resemblance, Romani *dragostoso* derives from Romanian *drăgăstos* and bears the Romani morphology for masculine singular, *-o*.

- *Phonology*. Oaxaca Romani has replaced the inherited voiced labial /v/ by the Spanish bilabial approximant /β/, as can be seen in the inherited part of its vocabulary: *βudar* ‘door’, *βas* ‘hand’.
- *Syntax*. In Mexican Romani, the Spanish adversative marker *pero* ‘but’ is used consistently, as is the committative *kon* ‘with’. There is variation between the inherited Romani complementizer *kaj* and the complementizer *ke* which could either be a borrowing from Spanish or from Romanian, as it is also attested in other Vlax Romani dialects. Although *ke* was most likely borrowed from the former contact language, Romanian, the formal and functional similarity to the actual contact language, Spanish, facilitates its use<sup>3</sup>.

A short excerpt from a narrative on the community’s origins is presented in (2). It illustrates some of the above mentioned borrowings, namely the adversative *pero* as well as the complementizer *ke*. Note that the indefinite *uni* ‘some’ is a borrowing from Romanian and not an adaptation of the Spanish *unos*:

- (2) a. *so zanas si ke tfi sam yaze*  
 what know.1PL be.3SG that NEG be.1PL non-Roma  
 SPANISH: Lo que sabemos es que no somos Gadze (non-Roma)  
 ‘What we know is that we are not Gadze (non-Roma)’
- b. *ke sam but xom*  
 that be.1PL many Roma  
 SPANISH: que somos muchos . . . Roma  
 ‘that we are many . . . Roma. [. . .]’
- c. *pero katka trais uni befas katse*  
 but here live.1PL some stay.1PL here  
 SPANISH: pero aquí vivimos unos estamos aquí  
 ‘But we live here, some of us are here.’  
 (Narrative)

<sup>3</sup> In other language contact settings, the Romanian *ke* was replaced by material from the new contact language.

## 2.4 Data and methodology

The present study is based on data elicited during several fieldwork trips (carried out in 2010, 2011 and 2012) using various questionnaires and visual stimuli. Four female speakers responded to the questionnaires, three of them aged 29–31, and one aged 44.

The Romani Morphosyntax questionnaire (henceforth RMS), designed by Matras and Elšík is based on the translation of 1061 entries comprising lexicon, sentences and verb paradigms. I submitted the questionnaire, with translations from Spanish to Romani, to two female speakers. The answers as well as the recordings of the eldest of the speakers are available on the RMS database (<http://romani.humanities.manchester.ac.uk/index.html>). The RMS questionnaire has the advantage of facilitating cross-dialectal comparison and greatly enhanced the present study.

The Questionnaire on Information Structure (henceforth QUIS) by Skopeteas et al. (2006) includes several tasks based on visual stimuli (see Figure 2). All four speakers responded to selected tasks from this questionnaire to test topicality and spatial relations (Given/new information in locative expressions; Broad/narrow focus in noun phrases; Tasks involving topicality, i.e., action, implicational topic, contrastive topic).

In order to explore the topic of attributive constructions, I used the questionnaire on the Spanish distinction between *ser* and *estar*, designed by Geeslin and Guijarro-Fuentes (2008). The four Romani speakers were first asked to choose among the two Spanish copulas in several pre-constructed sentences which followed a paragraph-long context. For the needs of the present study, they were then asked to translate the clauses into Romani. This allowed me to compare Spanish and Romani attributive constructions in identical contexts.



Fig. 2: Stimuli by Skopeteas et al. (2006)



The Pear Story video, designed by Chafe (1975) and described in Chafe (1980), was also used to elicit semi-spontaneous narratives which provide comparable material on different linguistic aspects. The five minute video presents a short story that takes place in a rural setting. The consultants watch the video once in order to become familiar with the story. The film is then shown a second time and the consultants are invited to comment on the actions as they happen. Lastly, the consultants are asked to provide a short summary of the film from memory.

The elicited material is supplemented by 20 minute life story narratives in Romani and 20 minute descriptions of family members, in both Spanish and Romani.

The data discussed in the paper are referenced by the elicitation task's name (*RMS* for Romani Morphosyntax Questionnaire, *QUIS* for Questionnaire on Information Structure, *Pear Story* for the material elicited through the video clip, and *Preference Task* for the questionnaire on copula choice), followed by the number of the question or task when possible. When no such marking is provided then the data has either been elicited in working sessions or is part of a narrative. Examples from published sources are followed by the appropriate references.

### 3 Romani subject clitic pronouns

In Oaxaca Romani, the 3rd person clitic pronouns attested are *lo*, *la*, *le* (see Table 1)<sup>4</sup>. The feminine form *la*, although rare, can also be found in other Northern Vlach dialects (i.e., Swedish Taikon Kalderaš, Serbian Kalderaš, Yugoslavian Lovari, see Elšík [2000]). Both feminine forms, *li* and *la*, are reconstructed for Early Romani (Elšík and Matras 2006: 316).

**Table 1:** 3rd person clitic pronouns in Mexican Romani

Singular	Masculine	<i>lo</i>
	Feminine	<i>la</i>
Plural		<i>le</i>

<sup>4</sup> Subject clitics (*lo*, *la*, *le*) shouldn't be confused with the etymologically related and formally similar 3rd person oblique pronouns (*le*, *la*, *le*). In Oaxaca Romani, the oblique pronouns are used as direct objects with full lexical verbs and combine with the copula 'to be' to form the possessive construction.

Romani subject clitics are what Anderson names “phonological clitics” in that their phonological form “lacks prosodic structure at the level of the (Prosodic) Word.”<sup>5</sup> (Anderson 2005: 19). Their position in the clause is not fixed, and they encliticize to a wide variety of morphemes.

Interestingly, in Oaxaca Romani subject clitics partake in some innovative uses not attested in other Romani dialects (Section 3.1), namely in attributive predication (Section 3.2), in locative predication (Section 3.3), and in constructions with participles (Section 3.4).

### 3.1 Historical development and dialect distribution

Subject clitics are an archaism in Romani (Matras 2002: 102). Reconstruction of Proto-Romani – i.e., pre-European Romani – indicates a set of demonstratives with a stem in *l-*, which later developed first into full 3rd person pronouns and then into clitics M.SG *\*-ta* > *\*-lo*; F.SG *\*-ti* > *\*-li*; PL *\*-te* > *\*-le* (Matras 2000: 111). A new set of demonstratives developed in Early Romani – i.e., Romani spoken after the arrival of the Roma in Asia Minor and the Balkans – as emphatic 3rd person pronouns (M.SG *\*ov*, F.SG *\*oj*, PL *\*on*), then as the default anaphora attested today in most Romani dialects (Matras 2000: 111). The development of this second set into the default anaphora is believed to have led to the loss of the *l-* set in most present day dialects:

- Clitics in *l-* have disappeared from the Northeastern and Northern Central Romani dialects (Matras 2002: 102).
- In the Vlax Romani dialects, subject clitics are now encountered only in nonverbal predication with presentatives *eta-lo* ‘there he is!’, place deictics *kate-lo* ‘here he is!’, and interrogatives *kaj-lo* ‘where is he?’ (Elšík and Matras 2006: 316).
- In Northwestern dialects, such as Finnish Romani, Welsh and English Romani, Slovene/Croatian Romani and partially in Caló (Spain) and in Romungro (Slovakia), subject clitics are used in existential predication with a copula: *si-lo* ‘he is’ (Matras 2002: 102).
- Only a few dialects have maintained the use of subject clitics with full lexical verbs (Elšík and Matras 2006: 213). In Austrian Lovari and Klenovec Romungro (Slovakia) subject clitics only co-occur with intransitive verbs. In Sinti and in Roman (Austria), subject clitics co-occur with transitive and intransitive verbs without any tense-aspect-mood restrictions. Below are two examples from the Sinti varieties spoken in France (3a) and Germany (3b):

---

<sup>5</sup> This expression allows the inclusion of clitics that can be stressed on some occasions within the larger prosodic word they form with their host (see Anderson 2005: 20, 39).

## (3) a. Sinti-Manuš

*o biboldo dikas lo ku kova*  
 the Jew looked.3SG CL at that

‘The Jew looked at that.’

(Valet 1991: 130 cited and glossed in Matras 1999: 150)

b. Hamlin Sinti<sup>6</sup>

*vajaso pal mende rodehs men lo*  
 came.3SG.CL after us.LOC searched us.ACC CL

‘He came after us, he was looking for us.’

(Holzinger 1993, cited and glossed in Matras 1999: 151)

## 3.2 Attributive predication

### 3.2.1 Distribution of the Romani clitics and copula

In affirmative, present tense attributive predication, Oaxaca Romani shows consistent variation between two structures. On the one hand [*si* ‘is, are’ + predicative adjective] exemplified in (4a), and on the other [predicative adjective + 3rd person clitic pronoun] exemplified in (4b). Contrary to what is attested in other Romani dialects (see Section 3.1), in Oaxaca Romani it is no longer possible to combine the clitics and the copula. Such constructions are considered ungrammatical by the speakers and are not attested in the current corpus (see 4c).

(4) a. *le fave tuxa bibiake si barbale*

DEF.PL children POSS.1SG aunt.DAT be.3PL rich

‘My aunt’s children are rich.’

(RMS-866)

b. *o yadzɔ tsulo =lo*

DEF.M man fat 3SG.M

‘The non-Gypsy man is fat.’

(Pear Story)

c. *\*o yadzɔ si =lo tsulo*

DEF.M man be.3SG 3SG.M fat

‘The non-Gypsy man is fat.’

<sup>6</sup> Note that in Hamlin Sinti the root has been reduced in the verbs, *lo* > *o*; see *vajaso* in Example (3b).

**Table 2:** Main variables for the distribution of *ser* and *estar* depending on the adjective's properties, in Silva-Corvalán (1986: 590) following Falk (1979) and Navas Ruiz (1963)

<i>ser</i>	<i>estar</i>
Inherent or essential	Accidental or circumstantial
Imperfective	Perfective
Permanent	Temporary
Defining, abstract and independent of immediate experience	Dependent on concrete and/or immediate experience
Not susceptible to change	Susceptible to change
Presented within a class frame of reference	Presented within an individual frame of reference

The variation between the Romani clitics and copula brings to mind the well-known variation between the Spanish copulas *ser* and *estar* (among others Navas Ruiz [1963]; Falk [1979]; Vañó-Cerdá [1982]; Hengeveld [1986]; Clements [1988]; de Jonge [1993]; Maienborn [2005]). Several variables have been examined in the literature in order to account for the distinction between *ser* and *estar*. Some of the most well established studies indicate that *ser* is used for essential, defining, permanent properties, independent of the referent's immediate situation; whereas *estar* is used with accidental, circumstantial, temporary properties which are dependent on the referent's immediate experience (see Table 2). Other variables are adjective class – some adjectives being constrained as to copula choice – and animacy – inanimate subjects being more often used with *ser* (Clements 1988).

The distribution of the Oaxaca Romani clitics and copula with attributive adjectives seems to parallel the main variables discussed for the Spanish copula distinction. Similar to Spanish *estar*, the Romani clitics are used with temporary properties (Examples 5 and 6), susceptible to change, and dependent on the immediate situation and subjective evaluation (Examples 7, 8 and 9) (for the importance of subjective evaluation in Mexican Spanish see García-Markina [2011: 44]).

## (5) Age

*miŋo      fau    tsinoŋo    =lo*  
 POSS.1SG   son    small      3SG.M  
 'My son is small.'  
 (RMS-546)

## (6) Mental state

*o            yaɖʒo    =lo      felis      kaj      naʃel*  
 DEF.M    man      3SG.M    happy    that    run.3SG

‘The non-Gypsy man is happy to run.’

(QUIS-S1-75, Answer to the question ‘Who is happy to run? The man or the woman?’)

(7) Description/evaluation

*o xabe kaj tson ande kako khero desa lafo*  
 DEF.M food that put.3PL in DEM.PROX house very good  
 =*lo desa ando foro*  
 3SG.M very in.DEF.M town

‘The food they serve in this house is the best in town.’

(RMS-829)

(8) Condition and description/evaluation

*o eskamin nevo =lo pero o tsan desa purano*  
 DEF.M chair new 3SG.M but DEF.M bed very old

‘The chair is new, but the bed is very old.’

(RMS-612)

(9) Condition and description/evaluation

*le gada abiveske desa kutfi aj desa fukar =le*  
 DEF.PL clothes wedding.DAT very expensive and very nice 3PL

‘Bridal clothes are very beautiful and very expensive.’

(RMS-587)

The parameters that govern the use of the Romani copula in 3rd person clauses are similar though not entirely parallel to those which have been discussed in the literature for the Spanish copula *ser*. Like Spanish *ser*, the Romani copula is used with essential properties not susceptible to change and independent of the immediate setting, as can be seen in the following examples ((10) in response to a specific picture, (11) for a close kinship relation and (12) for the description of a material). Unlike Spanish where, according to Clements (1988), animacy is relevant for choosing between the two copulas, the Romani copula is used with both animate (10, 11) and inanimate subjects (12):

(10) Color

*o zukel naj asul o zukel si galbeno*  
 DEF.M dog NEG.be.3SG blue DEF.M dog be.3SG yellow

‘The dog isn’t blue, the dog is yellow.’

(QUIS-Extra questions added to the “Animal game” task in order to elicit corrective focus)

## (11) Kinship

*kodja si tʃi fej*  
 DEM.DIST be.3SG POSS.2SG daughter  
 ‘Is that your daughter?’  
 (RMS-558)

## (12) Description

*kodja prama si sastruno*  
 DEM.DIST thing be.3SG metal  
 ‘That thing over there is made of metal.’  
 (RMS-673)

### 3.2.2 Preference task

In order to systematically test the equivalence between Spanish [*estar* + predicative adjective] and Romani [predicative adjective + 3rd person clitic pronoun], I used the contextualized preference task designed for the study of the Spanish *ser* and *estar* distinction by Geeslin and Guijarro-Fuentes (2008). Four Romani-Spanish bilinguals were presented with 28 clauses introduced by a paragraph-long context. For each clause, the consultants were asked to choose between the copulas *ser* and *estar*. Unlike for the original questionnaire, the Mexican Roma participants were also immediately asked to translate the target clauses into Romani.

The Romani copula was used for the negative clauses and all 1st and 2nd person clauses (11 clauses out of 28). See in (13) an example for 2nd singular:

(13) *tʃe phuri san*  
 INTJ old be.2SG  
 ‘How old you are!’  
 (Preference task, clause 9)

In the 3rd person affirmative causes, when Spanish *estar* was selected in the task, all speakers consistently translated it into Romani using the clitic pronouns (with two exceptions).<sup>7</sup> Below are some examples illustrating this equivalence:

---

<sup>7</sup> Item 4 was complicated for the consultants who found the context and the statement contradictory and this is believed to have influenced the result. In item 7, ‘I didn’t like the owner of the apartment, he is unpleasant’, no match could be found because in Romani there was neither a copula nor a clitic. This could be due to the structure of the Spanish clause which was “freely” translated in Romani by all consultants.

[Paula says thank you and asks if their friend Pablo will be joining them for dinner. Paula wants to talk to him about their math class. Raúl says that Pablo isn't coming and Paula wants to know why:

Paula: Why isn't Pablo coming?

Raúl: Because I didn't call him and now *he is mad*.]

- (14) *vo xuljariko =lo*  
 3SG.M.NOM angry 3SG.M  
 SPANISH: . . . está enojado.  
 ' . . . he is mad.'  
 (Preference task, clause 2)

[Paula and Raúl leave the apartment and go to a local restaurant. They eat there frequently and the people who work there are always very nice. This time, Raúl has ordered something new on the menu and Paula is curious about what Raúl thinks of the food:

Paula: Raúl, do you like your food?

Raúl: Yes, *dinner is good*.]

- (15) *lafo =lo o xabe*  
 good 3SG.M DEF.M food  
 SPANISH: La cena está buena.  
 'The dinner is good.'  
 (Preference task, clause 5)

[This comment seems funny to Paula. She didn't know that Raúl had such traditional ideas. She shows him her fingernails, which she just painted.

Paula: Well then, what do you think of my fingernails?

Raúl: *Your fingernails are blue too!*]

- (16) *bi le unzi asules =le*  
 also DEF.PL nails blue.PL 3PL  
 SPANISH: ¡Las uñas están azules también!  
 'The fingernails are blue too!'  
 (Preference task, clause 15)

The equivalence between *ser* and the Romani copula is less straightforward. Spanish *ser* was translated by the Romani copula *si* by all four speakers in only one case, presented in (17):

[Paula and Raúl also get a chance to catch up on the events of the day. Raúl mentions that his sister called and said she might like to come visit and stay at the apartment for a while. Paula didn't know Raúl had a sister so she asks how old she is:

Paula: How old is your sister?

Raúl: *Like us, she is young.*]

- (17) *sar amende bi voj si terni*  
 like 1PL.LOC also 3SG.F.NOM be.3SG young  
 SPANISH: Como nosotros, ella es bastante joven.  
 'Like us, she is young too.'  
 (Preference task, clause 17)

In Example (18), the Romani copula was used by all four speakers but the Spanish copula *ser* was selected by three of the four speakers.

[Since Raúl is curious about Paula's boyfriend, they discuss the possibility of a wedding. Paula is not sure about marrying her boyfriend because she is Catholic and her boyfriend is Protestant. She knows Raúl had a friend who married a man with a different religion and she asks what happened with her:

Paula: What religion does your friend practice now?

Raúl: *Now she is Catholic, too.*]

- (18) *akana si katolika bi voj*  
 now be.3SG Catholic also 3SG.F.NOM  
 SPANISH: Ahora ella es católica también.  
 'Now she is Catholic too.'  
 (Preference task, item 21)

There were discrepancies between the Romani and the Spanish answers in two cases. In the clause 'our apartment is very big', all four speakers preferred the Spanish *ser* copula but used in Romani the 3rd person clitic pronoun (item 6 in the questionnaire). The preference for the Romani clitic to be used for describing the size of a house is confirmed by examples from the RMS questionnaire presented in (19) and (20):

- (19) *lako kher maj baro =lo katar mura kher*  
 POSS.3SG.F house more big 3SG.M from POSS.1SG house  
 'Her house is bigger than my house.'  
 (RMS-614)



**Table 3:** Answers of the Mexican Romani-Spanish bilinguals to the preference task, compared to the copulas allowed in these contexts (Geeslin and Guijarro-Fuentes 2008: 379)

Copulas allowed in Spanish	<i>ser</i>			<i>estar</i>	<i>both</i>	
Spanish copulas selected	<i>ser</i>	<i>estar</i>	<i>both</i>	<i>estar</i>	<i>estar</i>	<i>ser</i>
Number of clauses in the task	2	2	2	3	10	9

(20) *lako kher maj baro =lo ando gau*  
 POSS.3SG.F house more big 3SG.M in.DEF.M village  
 ‘Her house is the biggest in the village.’  
 (RMS-615)

The second case where the Spanish and the Romani material did not show the expected equivalence concerns the clause ‘the tomatoes *are* unripe (lit. green)’ (clause 23). Although all four consultants selected the Spanish *ser*, the translations in Romani were divided (two speakers used the copula and two the clitic pronoun).

The Spanish copula distribution in the Roma answers is in keeping with the studies on Mexican Spanish, which document the expansion of *estar* in contexts previously occupied by *ser* (see Gutiérrez [1994]; de Jonge [1993]; Cortés-Torres [2004]; García-Markina [2011]). As summarized in Table 3, *estar* was selected in four contexts where only *ser* is allowed (in two contexts unanimously, in the other two contexts only partly). Moreover, in 10 cases where both Spanish copulas were allowed, the Mexican Romani bilinguals showed a preference for *estar*. Interestingly, *estar* was preferred in the 3rd person affirmative clauses (10/12), whereas *ser* was preferred in the negative and the non-3rd person clauses (7/9). A tentative analysis of these results could be that the Romani-Spanish bilinguals are integrating the Romani person and clause type restrictions in their selection of the Spanish copula.

To conclude, the results of the preference task confirm the equivalence between the Spanish *estar* and the Romani clitics (see Appendix). The equivalence between *ser* and the Romani copula is more complex and requires further research. This pilot study indicates that the predicate type, the frame of reference, and the adjective class don’t seem to be determining the selection of the Spanish or the Romani material (see Appendix). Immediate experience to the referent, susceptibility to change and inanimate subjects are factors that seem to favor the use of *estar* and of the Romani clitic but do not offer a one-to-one correspondence.

### 3.2.3 Restrictions in the uses of the 3rd person Romani clitics

Unlike Spanish *estar*, the uses of the Romani subject clitic pronouns in nonverbal attributive predications are restricted to the 3rd person (singular and plural), in affirmative and present tense clauses. Interestingly, these restrictions are commonly found in attributive predications in the world's languages. For instance, the restricted use of the Romani clitics in 3rd person predication is in keeping with the fact that zero copulas are always licensed for 3rd person pronouns but never for 1st person pronouns if the 3rd person is not already allowed (Eriksen [2005: 27] following Stassen [1997: 100, 317]). This observation could be applied to the Romani example if one considers that the verbal copula is usually first omitted in the 3rd person, and maintained longer in the 1st and 2nd persons. Moreover, the restriction applying to Romani clitics in present attributive clauses echoes the fact that in many languages zero copulas are favored for time stable situations (present tense), whereas verbal copulas are needed for temporal or aspectual specifications (Stassen 1997). Lastly, restrictions applying to Romani clitics in affirmative attributive clauses are a reminder that attributive strategies are frequently related to clause type. For instance it is common crosslinguistically for declarative, negative, and interrogative attributive clauses to use different means, such as specialized verbal copulas, for each clause type or for the roles to be distributed between a verbal and a zero copula (Adamou and Costaouec 2010: 77).

### 3.3 Locative predication

Similar to most Vlax Romani dialects, the Romani variety documented in Oaxaca uses subject clitic pronouns in nonverbal predication with place deictics (21a and 21b), and location interrogatives (21c):

- (21) a. *kotse* =*lo*  
 there 3SG.M  
 'There he is!'
- b. *katka* =*le von*  
 here 3PL 3PL.NOM  
 'Here they are!'  
 (RMS-540)
- c. *kaj* =*lo/la vo/voj ando zes*  
 where 3SG.M/F 3SG.M/F.NOM in.DEF.M day  
 'Where is he/she during the day?'  
 (RMS-526, 530)

For the 1st and 2nd person clauses the verbal copula is used, see (22a) and (22b):

- (22) a. *katka sam*  
 here be.1PL  
 ‘Here we are!’
- b. *kaj san tume*  
 where be.2PL 2PL.NOM  
 ‘Where are you<sub>pl</sub>?’  
 (RMS-537)

Unlike the European Vlax dialects, Oaxaca Romani also uses the 3rd person clitic pronouns in nonverbal locative predication where the locatum (the referent whose spatial location is expressed) is situated with respect to the relatum (the referent with respect to which the locatum is located in space). The clitics are used with both animate (23a) and inanimate (23b) subjects, and can either be modified by a definite article or not:

- (23) a. *o/jek liono =lo angla baɣ*  
 DEF.M/one lion 3SG.M in.front stone  
 ‘The/a lion is in front of a stone.’  
 (QUIS-S1-1)
- b. *o/jek kopatfi =lo pala rakloɣo*  
 DEF.M/one tree 3SG.M behind child  
 ‘The/a tree is behind a child.’  
 (QUIS-S1-69)

Subject clitics always agree with the subject in gender and number:

- (24) a. *e ɣazi =la pafa e barda*  
 DEF.F woman 3SG.F beside DEF.F fence  
 ‘The non-Gypsy woman is next to the fence.’  
 (QUIS-S1-170)
- b. *o ɣaɣo =lo angla e barda*  
 DEF.M man 3SG.M in.front DEF.F fence  
 ‘The non-Gypsy man is in front of the fence.’  
 (QUIS-S1-169)

Word order is relatively flexible. Although the clitic generally attaches to the subject, it can also attach to any other element in the clause, such as the relatum (25a), or the adverbial phrase (25b):

- (25) a. *e*      *ɣaʒi*      *paʃa*      *e*      *piramide*      =*la*  
 DEF.F woman beside DEF.F pyramid 3SG.F  
 ‘The non-Gypsy woman is next to the pyramid.’  
 (QUIS-S1-201)
- b. *kako*      *kher*      *maj*      *paʃe*      =*lo*      *katar*      *okolavera*  
 DEM.PROX house more close 3SG.M from other  
 ‘This house is much closer than that one (over there).’  
 (RMS-524)

The 3rd person clitic pronoun is used with a variety of adpositions, both inherited and borrowed, as shown in the following examples:

- (26) a. *o*      *ɣaʒo*      =*lo*      *infrente*      *katar*      *o*      *kher*  
 DEF.M man 3SG.M in.front from DEF.M house  
 ‘The non-Gypsy man is in front of the house.’  
 (QUIS-S1-125)
- b. *o*      *bar*      *intʃal*      *e*      *βulitsa*      =*lo*  
 DEF.M bar opposite DEF.F street 3SG.M  
 ‘The bar is across the street.’  
 (RMS-679)

While non-verbal predication with deictics is common cross-linguistically (Hengeveld 1992), the Romani use of place adpositions is intriguing. It is not easy to determine the grammatical role of the clitics in these constructions; although they have probably evolved from clitic doubling in topical constructions, topicality is no longer a relevant parameter in Mexican Romani. The Romani locative expressions cited above translate into Spanish using the verb *estar*, originally meaning ‘to stand’, from the Latin *stāre*, which has maintained its locative uses. The parallel between the Romani clitics and the Spanish verb *estar* in locative predication becomes even more obvious when the locative adposition is extended to non-spatial functions as can be seen in Example (27):

- (27) *e*      *istoria*      *opral*      *pe*      *muʒo*      *amigo*      =*la*  
 DEF.F story upon on POSS.1SG friend 3SG.F  
 SPANISH: La historia está encima de mi amigo.  
 ‘The story is about my friend.’  
 (RMS-664)

According to Hengeveld (1986: 397) *ser* is used with event terms, contrary to *estar*, which is not compatible with them. This is not the case however for Mexican Roma Spanish and corresponds to the Romani structure with the clitics:

- (28) *e fjestá=la ando sokalo*  
 DEF.F party-3SG.F in.DEF.M square  
 SPANISH: La fiesta está en el zócalo.  
 ‘The party is in the central square.’

Another possibility for describing spatial relations in Mexican Romani is to use presentational constructions with the inherited copula *si* ‘is’. This construction is translated into Spanish by the intransitive presentational verb *haber* ‘to have’, *hay* in the 3rd person singular. It would seem that in Romani the presentational clause with the copula *si* is preferred for low topical participants (29a), but this is not a constraint (29b):

- (29) a. *si but kopatfa*  
 be.3SG many trees  
 SPANISH: Hay muchos árboles.  
 ‘There are many trees.’  
 (Pear Story)
- b. *jek/o lion si pafa ek baɣ*  
 one/DEF.M lion be.3SG beside one stone  
 SPANISH: Hay un león al lado de una piedra.  
 ‘There is one/the lion next to one stone.’  
 (QUIS-S1-1/150)

To conclude, Mexican Romani uses two constructions for spatial descriptions: locative predication with the 3rd person clitic, which parallels the Spanish *estar*, and a presentative structure with the 3rd person copula *si*, which parallels the Spanish verb *hay*.

### 3.4 Participle constructions

Similar to the Vlax Romani dialects of Europe, the subject clitics in Oaxaca Romani are not used as anaphoric pronouns in doubling constructions with a full lexical verb (30), whether transitive or intransitive:

- (30) *\*o yaɣo kerel =lo exersisjo*  
 DEF.M man make.3SG 3SG.M exercise  
 ‘The non-Gypsy man exercises.’

Unlike European Vlax Romani, subject clitics are used in Oaxaca Romani with past participles of transitive and intransitive verbs. The construction [participle + 3rd person clitic pronoun] indicates a resulting state, similar to the Spanish constructions with *estar* (see Hengeveld [1986: 411–414] on restrictions in Spanish).

- (31) *o yaʒo desa pilo=lo*  
 DEF.M man very drunk.3SG.M-3SG.M  
 ‘The non-Gypsy man got very drunk.’  
 (RMS-553)

This type of construction with Romani clitics and past participles is not possible if there are temporal specifications (with past tense in Example 32), as this triggers the presence of the copula:

- (32) *desa pilo sas kaj tʃi arakjelas pesko*  
 very drunk.3SG.M was.3SG that NEG found.3SG REFL.POSS  
*propio kher*  
 own house  
 ‘He was so drunk, that he couldn’t find his own house!’  
 (RMS-554)

Furthermore, some innovative Romani participle forms parallel the Spanish participles. For example, instead of the Romani participle *befto* ‘seated’ one finds the form *befado*, which is probably a copy of the morphology of the Spanish participle *sentado*. Note that the Romani clitics attach either to the participle, as in (33a), or to the noun, in (33b):

- (33) a. *jek nonita befadi =la*  
 one girl seated.3SG.F 3SG.F  
 ‘One girl is seated.’  
 (QUIS-S1-52)  
 b. *e raktioxi =la befadi*  
 DEF.F girl 3SG.F seated.3SG.F  
 ‘The girl is seated.’  
 (QUIS-S1-52)

**Table 4:** Summary: Convergence between Mexican Romani and Spanish

	Spanish Model language	Mexican Romani Replica language	Construction	
1a.	<i>El mercado está en el pueblo.</i>	<i>o platso ando gav lo</i>	Spatial	Locative predication
1b.	<i>Hay una niña afuera.</i>	<i>si ek faoai avai</i>		Presentative
2a.	<i>El hombre está feliz.</i>	<i>o ya3o lo felis</i>	Attributive	Immediate experience
2b.	<i>El hombre es feliz.</i>	<i>o ya3o si felis</i>		Independent of immediate experience
3.	<i>Está sentado.</i>	<i>befado lo</i>	Participle	Resulting state

### 3.5 Summary

Mexican Romani has converged with Spanish as far as the uses of *estar* are concerned (illustrated in Table 4 with examples elicited through the RMS and the QUIS questionnaires). Romani clitics and Spanish *estar* are both used in locative and attributive predication, and both co-occur with the past participles of transitive and intransitive verbs expressing a resulting state.

## 4 Convergence between Mexican Romani and Spanish

In Section 4.1, I will present a convergence scenario to account for the innovative uses of Oaxaca Romani clitics. In Section 4.2, I will provide a discussion on the result of the Romani-Spanish convergence in terms of simplification or complexification.

### 4.1 Replication of Spanish *estar* in Mexican Romani

The convergence observed between Mexican Romani and Spanish in attributive and locative predication as well as in participle use, took place through a grammatical replication process; that is through the transfer not of the forms, but of the functions found in the model language (Heine and Kuteva [2005] building on

Weinreich [1953]). The directionality of the replication process is clearly from Spanish (L2) to Romani (L1), since similar variation in attributive, locative and participle constructions is not attested in any European Romani dialects and can be considered as an innovation of the Mexican Romani variety.

While the convergence between the two grammatical systems in contact is obvious, the replication process that led to this convergence is not easy to reconstruct. Typically, in a replication process, if two languages in contact both have a unit for the same function  $x$ , then function  $z$  of the unit which exists only in the model language (ML), may be replicated in the replica language (RL). The shared function  $x$  is what Matras and Sakel (2007) call the “pivot” feature (also in Matras 2009: 240–243).

The main problem in reconstructing the pivot matching process in Mexican Romani lies in our ignorance of the state of Romani at the early times of contact with Spanish. While in present-day Oaxaca Romani clitics and copulas are in complementary distribution, it is likely that when the Romani speakers arrived in Mexico, subject clitics were still used at least in copular predication, e.g., *si-lo* ‘he is’ (Matras 2002: 102). Co-occurrence of clitics and copulas is probably what allowed the split observed in the current state of the dialect for attributive constructions: on the one hand the Romani copula on its own, and on the other hand the Romani clitic on its own.

Through contact with Spanish, the need to replicate the Spanish copula distinction was probably felt by the Romani-Spanish bilinguals. The clitics, which although in decline were already used in attributive constructions, would have appeared suited to the task. This seems to be an illustration of the process of exaptation which has been suggested by Lass (1990) to account for the use of marginal grammatical material with a different function. According to Lass, when a form becomes marginal in a language it can be lost, it can be kept as marginal junk, or it can be reused for something else (exaptation). If subject clitics were lost or marginalized in most Romani dialects of Europe, it seems that in Mexican Romani language contact allowed them to be “recycled.”<sup>8</sup>

Beside the co-occurrence of the clitics in the attributive predication, a similarity in the functions shared by the Romani clitics and the Spanish copula *estar* may have facilitated the replication process. This pivotal feature could be that both structures were used for anchoring predication in a specific topic situation. On one hand, the Spanish copula *estar* is generally found in pragmatic contexts related to the speaker’s concrete and immediate experience. Maienborn (2005)

---

<sup>8</sup> A similar contact-induced exaptation has been suggested by Matras (1995) for Romani active participles.



further develops the “temporary vs. permanent” distinction and offers a very interesting analysis of the Spanish copulas based on pragmatics. She suggests that *estar* is the copula with which the speaker restricts his or her claim to a particular topic situation, as opposed to *ser* with which the speaker remains neutral as to the specificity of the topic situation. Maienborn explains that at the pragmatics level, by using *estar*, the speaker may restrict the predication if there are alternatives to the topic situation at a temporal, spatial or epistemic level. On the other hand, Romani subject clitic pronouns are claimed to be used in Sinti dialects for topical subjects (Holzinger 1993) even though Matras (1999) suggests this to be a kind of side-effect of syntactic specialization in obligatory VS clauses, as a result from convergence with German. Although no direct equivalence can be drawn between Mexican Vlach and Sinti (where the clitics are part of a more complex setting involving word order and a third type of reduced clitic forms according to Matras [1999]), we have evidence for a possible path to the clitics loss through topicality. More significantly, we can rely on the fact that unstressed pronominal morphemes are known to be the preferred topics across the languages (Lambrecht 1994: 165). Topical subjects are defined by Lambrecht as given, known agents which are [+active], [+accessible], and [+anchored] (Lambrecht 1994: 165). Lambrecht gives an example, shown in (33), to illustrate the importance of anchoring. According to Lambrecht, despite the fact that the referent ‘boy’ is indefinite and the predicate stative, it is limited by ‘in my class’ and anchored in the speaker’s universe:

- (33) *A boy in my class is real tall.*  
(Lambrecht 1994: 167)

The Romani clitics, which might have been used for topical subjects in attributive predications, replicated the uses of Spanish *estar* in predications which were, in Maienborn’s terms, anchored in a specific topic situation. In parallel, new, unknown participants, which did not trigger the use of subject clitics in Romani and were expressed by the Romani copula alone, may have been associated with the Spanish copula *ser*, preferred in contexts independent of the speaker’s immediate experience, or, in Maienborn’s terms, in cases where the speaker is neutral as to the specificity of the topic situation.

Replication at the level of attributive predication was certainly facilitated through two distinct replication processes that took place at roughly the same time, concerning namely locative predications and participle uses.

Figure 3 illustrates the replication process of the copula distinction (in bold) from the Model language (ML), Spanish, in the Replica language (RL), Mexican Romani. It illustrates that formerly in Romani, there was variation in copular

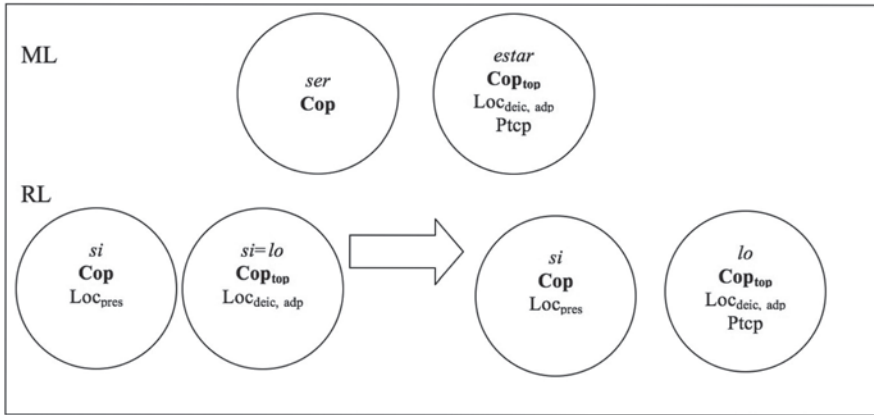


Fig. 3: Convergence between Mexican Romani and Mexican Spanish

predication between non-topical (Cop) and topical structures (Cop<sub>top</sub>). In Spanish and in today's Romani the distinction is between predication where the topic situation is not specified (Cop) and predication anchored in a specific topic situation (Cop<sub>top</sub>), in Maienborn's terms (2005). The other shared functions which played a facilitating role are also indicated, i.e., use with place deictics (Loc<sub>deic</sub>), locative adpositions (Loc<sub>adp</sub>) and presentatives (Loc<sub>pres</sub>) and co-occurrence with participles (Ptcp).

If at the moment of contact with Spanish the whole clitic system was restricted to locative predication with place deictics, presentatives and interrogatives, as is the case in most Vlax Romani dialects spoken in Europe, then the replication process of locative constructions would necessarily have been at the heart of the entire convergence process. However, such a convergence scenario lacks a convincing pivotal feature (Matras and Sakel 2007) for replication at the attributive predication level and for the uses with past participles.

## 4.2 Discussion: Complexification or simplification?

Several studies have been devoted to the effects of language contact on how Spanish copulas *estar* and *ser* are used. The expansion of *estar* at the expense of *ser* in American Spanish within the last three generations has led to claims that "language contact favors processes of linguistic simplification" (Silva-Corvalán [1986: 588], also Silva-Corvalán [1994]). This hypothesis spawned several studies on correlations between the Spanish copula distinctions and language contact.

Gutiérrez (1994) investigates the distribution of *ser* and *estar* in monolingual Mexican communities in Michoacán and observes a similar tendency towards the generalization of *estar*. Further work on Mexican Spanish by de Jonge (1993), Cortés-Torres (2004), and García-Markina (2011) confirms the expansion of *estar* in contexts previously dominated by *ser*. Testing Silva-Corvalán's hypothesis, Geeslin and Guijarro-Fuentes (2008) examine the variation between *estar* and *ser* in large-scale samples from monolingual and bilingual communities of the Iberian Peninsula, but no direct correlation could be established between bilingualism and the generalization of *estar*. A similar study on Puerto Rican communities (Ortiz-López 2000) shows that monolinguals tend to use *estar* more innovatively than bilinguals. Education, and therefore adherence to prescriptive norms, is an explanatory factor for these results given that bilingual communities in Puerto Rico also have higher education levels.

The Mexican Romani data contribute to the debate on the effects of language contact on the distinction between the Spanish copulas by taking a different tack, namely by examining the effects that this distinction may have on unrelated contact languages. Contrary to Silva-Corvalán's findings for the simplification of the Spanish copula distinction in American-Spanish bilingual speech, the case of Mexican Romani-Spanish bilinguals offers data supporting the complexification of the L1 (Romani) under influence from the L2 (Spanish) by replication of the L2 copula distinction. However, following Matras (2009), it can be argued that convergence between the grammatical systems in contact simplifies matters for bilinguals in that it leads to the reduction of the cognitive load. According to Matras: "In a situation of intense bilingualism coupled with loyalty toward language maintenance, it seems advantageous for bilingual speakers to be able to syncretise the mental planning operations applied while interacting in each of the languages" (Matras 2009: 238).

The process which leads up to this convergence may be either the addition or the loss of morphosyntactic complexity. Whether the process will be one of loss or addition seems to depend on the type of contact settings, as convincingly argued by Trudgill (2009: 101). As he puts it, the age of the speakers at the moment of language acquisition is a crucial parameter: "Simplification in language contact does not result from nonnative language learning as such, but from post-critical threshold nonnative language learning" (Trudgill 2010: 313). High-contact settings involving adult speakers may lead to simplification, while stable long-term contact settings involving child bilingualism seem to favor complexification (Trudgill 2009). Romani Mexican-Spanish bilinguals are clearly part of the second contact setting, where complexification of the L1 system at the level of attributive predication took place under the influence of the more widely spoken and prestigious L2, Spanish.

## 5 Conclusion

Oaxaca Romani presents a range of innovative uses for the subject clitic pronouns in a process of exaptation (Lass 1990). The use of the Romani subject clitics in attributive and locative predications, as well as in constructions with participles, appears to be replicating the uses of the Spanish *estar*. The expansion of the copula *estar* in Mexican Spanish is a well-documented tendency (see Gutiérrez [1994]; de Jonge [1993]; Cortés-Torres [2004]; García-Markina [2011]), and the present study shows it is a particularly salient feature for the Mexican Romani-Spanish bilinguals. Rather than simply borrowing *estar*, Mexican Romani speakers replicate its uses by recycling the native material in a highly creative way, confirming the view of bilingual speakers as active “language builders” (Hagège 1993). Although the replication process adds complexity in the replica language, the overall convergence of the grammatical systems in contact simplifies matters for bilinguals since it leads to reduction of the cognitive load (Matras 2009).

**Acknowledgments:** This study was presented at the 10th International Conference on Romani linguistics (University of Barcelona, 2012) and benefited from the remarks of its participants. I’m also grateful to the reviewers and the editorial team of *Linguistics* for their constructive suggestions. Moreover, I wish to thank Yaron Matras, Denis Costaouec, Jean-Léo Léonard and Michael Swanton for their support and encouragement of my research on Romani in Mexico, and Yekaterina García-Markina for sharing useful bibliographical references. Many thanks to the language consultants of the Romani community of Oaxaca.

## References

- Adamou, Evangelia. 2011. A preliminary study of Romani in Mexico. Paper presented at the Gypsy Lore Society meeting 2011, Graz, 1–3 September.
- Adamou, Evangelia & Denis Costaouec. 2010. Connective constructions in the world’s languages. *La linguistique* 46(1). 43–80.
- Anderson, R. Stephen. 2005. *Aspects of the theory of clitics*. Oxford: Oxford University Press.
- Bakker, Peter & Yaron Matras. 1997. Introduction. In Yaron Matras, Peter Bakker & Hristo Kyuchukov (eds.), *The typology and dialectology of Romani*, vii–xxx. Amsterdam & Philadelphia: John Benjamins.
- Boretzky, Norbert & Birgit Igla. 2004. *Kommentierter Dialektatlas des Romani*. Wiesbaden: Harrassowitz.
- Chafe, L. Wallace. 1975. *The pear story film*. University of California.
- Chafe, L. Wallace. 1980. *The pear stories: Cognitive, cultural and linguistic aspects of narrative production*. Norwood, NJ: Ablex.

- Clements, J. Clancy. 1988. The semantics and pragmatics of the Spanish <copula + adjective> construction. *Linguistics* 26. 779–822.
- Cortés-Torres, Mayra. 2004. ¿Ser o estar? La variación lingüística y social de *estar* más adjetivo en el español de Cuernavaca, México. *Hispania* 87. 788–795.
- Elšík, Victor. 2000. Dialect variation in Romani personal pronouns. In Victor Elšík & Yaron Matras (eds.), *Grammatical relations in Romani: The noun phrase*, 65–94. Amsterdam & Philadelphia: John Benjamins.
- Elšík, Victor & Yaron Matras. 2006. *Markedness and language change*. Berlin & New York: Mouton de Gruyter.
- Eriksen, K. Pål. 2005. *On the typology and the semantics of non-verbal predication*. Oslo: University of Oslo dissertation.
- Falk, Johan. 1979. *Ser y estar con atributos adjetivales*. Uppsala: Almqvist & Wiksell.
- García-Markina, Yekaterina. 2011. *Ser and estar in evolution: Two stages of the extension of estar in Mexican Spanish*. In David Hornsby (ed.), *Interfaces in language 2*, 35–52. Cambridge: Cambridge Scholars Publishing.
- Geeslin, L. Kimberly & Pedro Guijarro-Fuentes. 2008. Variation in contemporary Spanish: linguistic predictors of *estar* in four cases of language contact. *Bilingualism: Language and cognition* 11(3). 365–380.
- Gilliat-Smith, Bernhard. 1915. A report on the Gypsy tribes of North East Bulgaria. *Journal of the Gypsy Lore society* 9. 1–54.
- Gutiérrez, J. Manuel. 1994. ‘Ser’ y ‘estar’ en el habla de Michoacán, México. México, D.F.: UNAM.
- Hagège, Claude. 1993. *The language builder*. Amsterdam & Philadelphia: John Benjamins.
- Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.
- Hengeveld, Kees. 1986. Copular verbs in a functional grammar of Spanish. *Linguistics* 24. 393–420.
- Hengeveld, Kees. 1992. *Non-verbal predication: Theory, typology, diachrony*. Berlin & New York: Mouton de Gruyter.
- Holzinger, Daniel. 1993. *Das Romanes: Grammatik und Diskursanalyse der Sinte*. Innsbruck: Verlag des Instituts für Sprachwissenschaft der Universität Innsbruck.
- de Jonge, Bob. 1993. (Dis)continuity in language change: ‘Ser’ and ‘estar’ + age in Latin American Spanish. In Frank Drijkoningen & Kees Hengeveld (eds.), *Linguistics in the Netherlands*, 69–80. Amsterdam & Philadelphia: John Benjamins.
- Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press.
- Lass, Roger. 1990. How to do things with junk: exaptation in language evolution. *Journal of linguistics* 26. 79–102.
- Maienborn, Claudia. 2005. A discourse-based account of Spanish *ser/estar*. *Linguistics* 43(1). 155–180.
- Matras, Yaron. 1995. Verb evidentials and their discourse function in Vlach Romani narratives. In Yaron Matras (ed.), *Romani in contact*, 95–123. Amsterdam & Philadelphia: John Benjamins.
- Matras, Yaron. 1999. Subject clitics in Sinti. *Acta Linguistica Hungarica* 46. 147–168.
- Matras, Yaron. 2000. The structural and functional composition of Romani demonstratives. In Victor Elšík & Yaron Matras (eds.), *Grammatical relations in Romani: The noun phrase*, 95–122. Amsterdam & Philadelphia: John Benjamins.

- Matras, Yaron. 2002. *Romani: A linguistic introduction*. Cambridge: Cambridge University Press.
- Matras, Yaron. 2005. The classification of Romani dialects. A geographic-historical perspective. In Dieter Halwachs & Barbara Schrammel (eds.), *General and applied Romani linguistics*, 7–26. Munich: Lincom Europa.
- Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.
- Matras, Yaron. 2010. *Romani in Britain: The afterlife of a language*. Edinburgh: Edinburgh University Press.
- Matras, Yaron & Jeanette Sakel. 2007. Investigating the mechanisms of pattern replication in language convergence. *Studies in language* 31(4). 829–865.
- Miklosich, Franz. 1872–1880. *Über die Mundarten und Wanderungen der Zigeuner Europas*, vol. 3. Vienna: Carl Gerold's Sohn.
- Navas Ruiz, Ricardo. 1963. *Ser y estar: estudio sobre el sistema atributivo del español*. Salamanca: Acta Salmanticensia.
- Ortiz-López, Luis. 2000. La extensión de *estar* en contextos de *ser* en el español de Puerto Rico: ¿Evolución interna o contacto de lenguas? *Boletín de la academia puertorriqueña de la lengua española*. 98–118.
- Pickett, W. David. 1962. *Prolegomena to the study of gypsies of Mexico*. Ithaca, NY: Syracuse University MA thesis.
- Pickett, W. David & Rafael Gonzalez. 1964. A Mexican gypsy legend. *Journal of the Gypsy Lore society* 43. 3–11.
- Ripka, Štěpán. 2007. *Informe de avances de investigación: gitanos pentecostales en México*. Prague: Charles University MA thesis.
- Silva-Corvalán, Carmen. 1986. Bilingualism and language change: The extension of *estar* in Los Angeles Spanish. *Language* 62. 587–608.
- Silva-Corvalán, Carmen. 1994. *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon Press.
- Skopeteas, Stavros et al. 2006. Questionnaire on information structure. Audiovisuelles Zentrum der Universität Potsdam und sd:k Satz Druck GmbH Potsdam.
- Stassen, Leo. 1997. *Intransitive predication*. Oxford: Clarendon Press.
- Szente-Varga, Mónika. 2009. Inmigrantes húngaros en México y la formación de una colonia húngara. *Dimensión antropológica* 43. <http://www.dimensionantropologica.inah.gob.mx/?p=2328> (accessed 6 July 2012).
- Trudgill, Peter. 2009. Sociolinguistic typology and complexification. In Geoffrey Sampson, David Gil & Peter Trudgill (eds.), *Language complexity as an evolving variable*, 98–109. Oxford: Oxford University Press.
- Trudgill, Peter. 2010. Contact and sociolinguistic typology. In Raymond Hickey (ed.) *The handbook of language contact*, 299–319. Malden, MA & Chichester: Wiley-Blackwell.
- Valet, Joseph. 1991. Grammar of the Manush as it is spoken in the Auvergne. In Peter Bakker & Marcel Cortiade, *In the margin of Romani*, 106–131. Amsterdam: Institute for General Linguistics.
- Vañó-Cerdá, Antonio. 1982. *Ser y estar + adjetivos: un estudio sincrónico y diacrónico*. Germany: Gunter Narr.
- Weinreich, Uriel. 1953. *Languages in contact*. The Hague: Mouton.

## Appendix

Answers of the Mexican Romani-Spanish bilinguals to the 3rd person affirmative clauses of the preference task (Geeslin and Guijarro-Fuentes 2008: 379)

Answers for Spanish	<i>estar</i>			<i>ser</i>		<i>both</i>	
Answers for Romani	<i>clitic</i>	<i>copula</i>	<i>clitic</i>	<i>both</i>	<i>copula</i>		
Number of clauses in the task	11		1	2	1	1	
Predicate type (Number of clauses)	stage (7/11)	individual (4/11)	stage (1/1)	individual (2/2)	stage (1/1)	stage (1/1)	
Referent (Number of clauses)	individual (7/11)	class (4/11)	individual (1/1)	class (2/2)	individual (1/1)	individual (1/1)	
Change (Number of clauses)	yes (9/11)	no (2/11)	yes (1/1)	no (1/2)	yes (1/2)	yes (1/1)	no (1/1)
Experience with the referent (Number of clauses)	immediate (8/11)	ongoing (3/11)	ongoing (1/1)	ongoing (2/2)	immediate (1/1)	ongoing (1/1)	
Animacy (Number of clauses)	no (8/11)	yes (3/11)	yes (1/1)	no (1/2)	yes (1/2)	yes (1/1)	yes (1/1)
Adjectives	mad, good, bad, blue, pretty, sick, beautiful, cold, sad		happy	big, young		unripe	Catholic

## Abbreviations

Glosses follow the Leipzig glossing rules. More specific abbreviations are presented below.

CL	clitic
INTJ	interjection
QUIS	Questionnaire on Information Structure
RMS	Romani Morphosyntax Questionnaire

