Linguistica sine finibus
Estudis dedicats a Montserrat Batllori Dillet

Elisabeth Gibert-Sotelo, Isabel Pujol Payet,
Assumpció Rost Bagudanch, Teresa de Jesús Tro Morató
(eds.)

LINGUISTICA SINE FINIBUS

ESTUDIS DEDICATS A MONTSERRAT BATLLORI DILLET

Girona, 2023

ÍNDEX

TERCERA PART. Història de la llengua

QUARTA PART. Estructura argumental: teoria i aplicacions

CINQUENA PART. Anàlisi de corpus

CLITIC CLIMBING IN MODAL CONSTRUCTIONS IN ALGHERESE CATALAN

IOANNA SITARIDOU
University of Cambridge and Queens' College
is269@cam.ac.uk

TRISTAN LEE
Independent Scholar
tristan.lee001@gmail.com

Abstract

This work represents the first analysis of clitic climbing in Algherese, the variety of Catalan spoken in the city of Alghero, Sardinia. On the basis of novel and existing data, we claim that clitic climbing in modal constructions is obligatory in Algherese, as opposed to most other varieties of Catalan in which it is optional (with the exception of some varieties of Balearic Catalan, in which clitic climbing is also obligatory). However, optionality is emerging in the grammars of some speakers due to language contact with Standard Italian, a variety in which clitic climbing is optional. We show this by controlling for two variables that predict for crosslinguistic interference from Italian, one being language dominance and the other frequency of use. Both variables are important factors of crosslinguistic interference in research on L1 attrition.

Resum

Aquest treball és la primera anàlisi de la pujada de clítics en alguerès, variant del català parlada a la ciutat de l'Alguer, Sardenya. Partint de dades noves i ja existents, mostrem que la pujada de clítics en les construccions modals és obligatòria en alguerès, contràriament a la majoria de parlars catalans, en els quals és opcional (a part d'algunes varietats de català balear, en les quals també és obligatòria). Tanmateix, trobem opcionalitat en les gramàtiques d'alguns parlants degut al contacte lingüístic amb l'italià estàndard, una llengua en la qual la pujada de clítics és opcional. Per a demostrar-ho, analitzem dues variables associades a l'atrició de la L1, que tenen com a objectiu predir la interferència interlingüística de l'italià en l'alguerès: d'una banda, controlem el domini lingüístic dels parlants i, de l'altra, la freqüència d'ús de la llengua.

To Montse, whom I have known for 20 years during which I have had the gift of her friendship, guidance, sabiduría and kindness. I owe her more than I could possibly put in words, and my student Tristan and I offer this piece of new research as a small token of gratitude for her amazing contributions to the diachronic generative study of Catalan, Spanish and Romance.

Ioanna Sitaridou

1. INTRODUCTION

This work constitutes the first analysis of clitic climbing in Algherese, the variety of Catalan spoken in the city of Alghero, Sardinia. The morphosyntax of Algherese is understudied and most existing literature on Algherese focusses on other aspects of the language, such as phonology (e.g., Ballone 2016), lexicon (e.g., Bosch 2012; Corbera 1993, 2000), and its sociolinguistic situation both historically (e.g., Bosch 2002; Chessa 2008) and contemporaneously (e.g., Argenter 2008; Ballone 2017; Caria 2006; Perea 2010). Works that purport to describe the grammar of Algherese are limited and none address clitic climbing directly; such works include two grammars published in the early 20th century (Pais 1970 [1906]; Palomba 2001 [c.1907]), the prescriptive standard promulgated by the *Institut d'Estudis Catalans* (*El català de l'Alguer: un model d'àmbit restringit*, 2003), a historical grammar of Catalan with a focus on Algherese in particular (Blasco Ferrer 1984) and an approximation of colloquial Algherese (Bosch 1998). To our knowledge, there is just one dedicated analysis of the Algherese clitic system by Bosch and Scala (1999) who provide a morphophonological analysis.

Consequently, the central claim in this study regarding the position of complement clitics in modal constructions in Algherese is entirely novel and based on the analysis of both novel and existing data of spoken Algherese. The claim is outlined below:

- • Clitic climbing in modal constructions is obligatory in Algherese, but optionality has emerged in the grammar of some speakers due to language contact with Italian.

The current work takes the following structure. In §2, we give a brief overview of the phenomenon of clitic climbing across Romance with special reference to the typology established by Cinque (2006). In §3, we outline the methodologies that underpin our quantitative analysis of spontaneous speech data from two sources: in-person interviews conducted with 15 Algherese speakers in May 2022 (Lee 2022); and four corpora of spoken Algherese that are available online. The results are presented in §4 and discussed in §5 in the context of language contact between Algherese and both Sardinian and Italian. We conclude in §6 and identify areas for future research.

2. CLITIC CLIMBING ACROSS ROMANCE

Clitic climbing (CC) in Romance can be defined as a syntactic configuration where a clitic that is selected as an argument by a lower verb attaches to a higher verb (Rivas 1977: in Sitaridou et al. 2015). The semantic class of the main verb is an important factor that restricts the availability of CC across Romance, and CC occurs typically only with auxiliaries, modals, volitionals, aspectuals and causatives (Roberts 2016; Sitaridou et al. 2015). CC is ungrammatical when verbs that do not belong to the semantic classes listed above take a non-finite verb as a complement. For example, compare the grammaticality of CC in the volitional construction in (1) to the ungrammaticality of CC in (2), where the main verb *detestare* does not belong to any of the aforementioned semantic classes. In the latter case, the clitic must appear attached to the lexical verb that selects it, as in (3).

(1) *Ci   vorrei   andare   con   Maria*
    loc=   want.COND.1SG   go.INF   with   Maria
    «I would like to go with Maria»
                                    Italian (Cardinaletti and Shlonsky 2004: 522)

(2) *Ci* *detesterei* *andare* *con* *Maria*
loc= hate.COND.1SG go.INF with Maria
«I would hate to go there with Maria»
<div align="right">Italian (Cardinaletti and Shlonsky, 2004: 521)</div>

(3) *Detesterei* *andarci* *con* *Maria*   Italian
hate.COND.1SG go.INF=LOC with Maria
«I would hate to go there with Maria»

Since Kayne (1975), CC in causative constructions is considered a different syntactic phenomenon to that in auxiliary, modal, volitional and aspectual constructions. Additionally, auxiliary verbs have been shown to behave differently from modals, aspectuals and volitionals and are often considered separately (e.g., Gavarró and Laca 2008; Jones 1993). Modals, aspectuals and volitionals go by the conventional term «restructuring verbs» in the spirit of Rizzi (1982).

The distribution of CC in restructuring constructions across Romance is linked to the availability of null subjects. CC is available in null-subject languages such as Catalan and Spanish but unavailable in non-null-subject languages such as French (Kayne 1989). Moreover, CC in such constructions is obligatory in some Romance varieties, such as Sardinian (Jones 1993; Mensching 2017), but optional in others, such as Spanish (Sitaridou et al. 2015) and most varieties of Catalan (with the exception of some Balearic varieties, in which CC is obligatory) (Gavarró and Laca 2008). Lastly, in some Romance varieties such as Piedmontese (Parry 1994; Tortora 2014), Chilean Spanish (Sitaridou et al. 2015) and certain Central Catalan varieties (Bonet 2008), there are instances of a phenomenon known as clitic repetition, where a copy (in the theoretically neutral sense) of a clitic with the same referent appears in both the climbed and unclimbed position simultaneously. Cinque (2006, in Sitaridou et al. 2015: 98) summarises these empirical observations and makes the following typology of the availability of CC in restructuring constructions across Romance:

Table 1. Typology of clitic climbing across Romance

	Clitic climbing with restructuring verbs	Exemplary varieties
Type 1	Optional	Standard Italian, Spanish, most varieties of Catalan
Type 2	Obligatory	Sardinian, some varieties of Balearic Catalan
Type 3	Heavily restricted or disallowed entirely	French
Type 4	Clitic repetition	Piedmontese, Chilean Spanish, certain varieties of Central Catalan

Adapted from Cinque (2006) in Sitaridou et al. (2015: 98)

Given the typological variation between different varieties of Catalan, the purpose of the current work is to determine the typological status of Algherese with respect to CC, which has so far not received study, and in doing so to examine the contact hypothesis as stated in (i).

3. METHODOLOGY

The basis of the current work is a quantitative analysis of spontaneous speech data from two sources: in-person interviews with 15 Algherese speakers that took place in May 2022 (Lee 2022); and four corpora of spoken Algherese that are accessible online, listed below. Corpora 1 and 3 are available online in their entirety, while Corpora 2 and 4 appear as a fragment of the complete corpora.

- [**Corpus 1**] Armangué, Joan and Luca Scala. 1997. *L'Alguer. Primera campanya de gravacions de literatura popular de tradició oral.* http://prosodia.upf.edu/coalgueres/ca/corpus/adt.html

- [**Corpus 2**] Ballone, Francesc. 2000-2008. *Corpus Oral de l'Alguerès.* http://prosodia.upf.edu/coalgueres/ca/corpus/ballone.html

- [**Corpus 3**] Bosch i Rodoreda, Andreu, and Susanna Sanna. 1996. *Històries de l'Alguer, entre la marina i la campanya.* http://prosodia. upf.edu/coalgueres/ca/corpus/bosch.html

- [**Corpus 4**] Corbera, Jaume, and Enrico Chessa. 2009. *Conversa amb…Corpus Oral de l'Alguerès*. http://prosodia.upf.edu/coalgueres/ca/corpus/corbera.html

With regards to the linguistic feature under study, we restrict the scope of this work to the position of clitics in modal and volitional constructions. While modal and volitional verbs are commonly grouped together with aspectual verbs in terms of their syntactic behaviour (as discussed in §2), Gavarró and Laca (2008) observe that these three semantic classes of verbs do not behave uniformly in Catalan varieties. With this in mind, we focus on modals and volitionals –the latter class being represented by just one verb, volguer «to want», in our sample– and exclude aspectuals. For the sake of brevity and precision, we group modals and volitionals under the term «modal verbs» in the current work.

3.1. Interviews

Interviews were conducted with a total of 15 speakers, of which 9 were female and 6 were male. The average age of the group was 58.9, the eldest participant being 80 and the youngest 25. The only selection criteria for potential participants was the ability to speak Algherese «fluently», [1] and the interviews were conducted in Algherese. All speakers reported fluency in Italian in addition to Algherese. None of the participants reported any formal learning in linguistics.

Due to universal Algherese/Italian bilingualism among the speakers, we encounter what D'Alessandro et al. (2021) coin as the «baseline challenge», where the language variety in question does not have a sizeable population of monolingual speakers or a widely disseminated standard to act as a control for crosslinguistic interference (CI) from other varieties.[2] However, it is widely acknowledged in the literature on bilingualism that not all bilinguals behave the same with respect to CI. Thus, we control

---

1 Due to time constraints, Lee was not able to measure fluency by means of utterance length or lexical diversity beforehand albeit this can be done retrospectively.

2 While there exists the regional standard for Algherese that is promulgated by the Institut d'Estudis Catalans (2003), it is not disseminated widely in Alghero and only 8.1% of speakers report being able to read Algherese (Ballone, 2017: 9).

for two variables— language dominance and frequency of use— that influence crosslinguistic interference and divide the speakers in our interviews into groups. The comparison between these groups will allow a finer-grained analysis of our sample. To establish the criteria for the groupings when we control for each variable, we rely on sociolinguistic data collected with a questionnaire that formed part of the interviews.

3.1.1. Language dominance

Language dominance has been shown to be an important factor for CI in L1 attrition (see Polinsky 2018). In this study, we operationalised language dominance as relative proficiency, where a comparison is drawn between a bilingual's languages; one is considered «stronger» and one (or more) is considered «weaker». While some researchers attempt to quantify the proficiency of each of a bilingual's languages independently through testing lexical diversity, sentence length and other factors (e.g., Unsworth et al., 2018), an alternative is to employ a qualitative approach. For example, Pérez-Leroux et al. (2011) assess the relative proficiency of the bilingual children in their study by asking the children's parents to give a rating for the children's fluency in either language. In a similar vein, we invited the participants in our study to give a qualitative judgement of their own relative proficiency in Algherese and Italian (and any other languages they spoke) with the question: *With which language do you express yourself best?*

Based on the speakers' response, they were divided into three groups: Algherese dominant, balanced, and Italian dominant. See Table 2 below for the groupings:

Table 2. Groupings based on responses to Q9 of the sociolinguistic questionnaire: «With which language do you express yourself best?»

Dominant language	Participants
Algherese	S2, S3, S5, S8, S9, S13
Balanced	S4, S10
Italian	S1, S6, S7, S11, S12, S14, S15

With regards to predictions, the speakers who reported dominance in Italian are expected to show relatively higher levels of CI from Italian than

the other two groups, while those who reported dominance in Algherese are expected to show the opposite.

3.1.2. Frequency of use

The frequency at which a speaker uses and is exposed to a language is difficult to operationalise as a variable but warranted inclusion in the study because it is considered an important factor of CI, especially in studies on attrition (Grosjean 2010; Paradis 2004; Schmid and Köpke 2017). To quantify the frequency with which a speaker used their languages, we relied on the participants' responses to the following questions.

- *Which language(s) do/did you use with:*

    a) *Your parents?*

    b) *Your siblings?*

    b) *Your grandparents?*

- *Which language(s) do you use:*

    a) *With your friends?*

    b) *With your partner?*

    c) *With your children?*

    d) *At work?*

To quantify the responses and standardise the criteria that were used to assign participants to groups, we assigned a numerical value to each possible response using a 5-point Likert scale, outlined below:

Table 3. Scoring criteria for responses to Q11 and Q12 of the sociolinguistic questionnaire using a 5-point Likert scale

Likert-scale score	
-2	Italian exclusively
-1	Italian mostly
0	Both languages equally
+1	Algherese mostly
+2	Algherese exclusively

With the scores, we calculated the average for each participant: a positive score meant that they used Algherese more and vice versa. In so doing, we made a binary variable that indicated which language a participant used most. The average results for each participant and their groupings appear below:

Table 4. Grouping of each participant in a binary variable based on frequency of use of Italian and Algherese

Grouping	Speakers (Score)
Algherese	S2 (2.00), S3 (0.50), S4 (0.57), S5 (0.29), S8 (0.57), S9 (0.60), S13 (1.00)
Italian	S1 (-1.75), S6 (-1.57), S10 (-0.33), S11 (-0.17), S12 (-1.50), S14 (-1.60), S15 (-1.80)

The predictions to be made for these groups are as follows: the participants who reported using Italian more in daily life should show more CI from Italian, while those who reported using Algherese more frequently should show less.

3.2. Online corpora

The speakers in the online corpora lack detailed demographic and sociolinguistic profiles. Only the age and gender of the speaker is provided in each corpus, except for Corpus 4, in which the creators generalise that all their speakers were over the age of 70 at the time of interviewing in 2008.

Despite the lack of information on the speakers in the corpora, there are a number of reasons to still consider the data in the analysis. Given that each corpus aims to document the language and oral traditions of Alghero, it is assumed that the corpus creators made a qualitative assessment that the speakers' language was «representative» enough to exemplify Algherese. While this assumption provides enough motivation to consider the data from the online corpora, it does not constitute strong enough justification to include the corpus speakers in any of the groupings that we make for the participants of my interviews to account for CI. Thus, the data from the online corpora will be presented separately in the results section.

3.3. Data collection and analysis

Spontaneous speech data is commonly taken to be representative of linguistic proficiency. Examples of clitics in modal constructions were transcribed using the standard for Algherese (Institut d'Estudis Catalans 2003). Pauses and false starts were not represented in the transcriptions. The online corpora present transcripts along with the audio files, and examples from these transcripts were double checked with the corresponding audio file before inclusion in the study.

All examples of clitics in modal constructions in the sample were amalgamated in a database in Microsoft Excel and exported to statistical-analysis software, Stata, to visualise and analyse the data. The quantity of clitics in the climbed position were compared to that of clitics in the unclimbed position. Chi-square tests are performed to determine whether the difference in behaviour between the groups of speakers is statistically significant when controlling for language dominance and frequency of use.

4. RESULTS

The total length of spontaneous speech data analysed in this study totals to 15 hours and 29 minutes: 2 hours and 11 minutes from the online corpora and 13 hours and 18 minutes from the interviews. The recording of the interview with one participant, S5, was corrupted, but we were able to salvage the data collected from the sociolinguistic questionnaire and make two observations of clitics in modal constructions. The total number of tokens in the entire sample totals 83, and consult the Table 5 for the distribution of tokens per source.

When analysing the examples of clitics in restructuring constructions, there were three syntactic patterns which problematised conclusions as to whether clitic climbing had occurred or whether clitic climbing was even an available option. We outline the three patterns below and explore the challenges that each group poses before making a decision whether to include or exclude the examples in the analysis.

Table 5. Distribution of examples of clitics in modal constructions per source

Source	Number of examples	Percentage (%)
Corpus 1	14	16.9
Corpus 2	9	10.8
Corpus 3	10	12.0
Corpus 4	5	6.0
Interviews	45	54.3
Total	**83**	**100**

4.1. Clitic repetition

In the sample of clitics in restructuring constructions, there were five examples in which the same clitic with the same referent appears twice in the oration, once in the climbed position and once in the unclimbed position. See examples (4) to (8) below:[3]

(4) *Qui   les   teniva   de   menjar-les,   aqueixes*
    who   3PL.FEM.ACC= have.impf.3sg of   eat.inf=3PL.FEM.ACC those
    dues   peres?
    two   pear
    «Those two pears, who must have eaten them?»

    (CAO, M, 1927, free speech) (Corpus 3)

(5) *Tu, li   ha   dit,   los   tens   de*
    you 3SG.DAT= have.3sg said   3PL.MASC.ACC= have.2sg   of
    portar-los   en giro del   país   en carrossa,   a
    take.inf=3PL.MASC.ACC   in tour   of.the country in train.carriage to
    mos   fills!
    my   children
    «You, she said to him, 'you have to take my children around the country by train!'»

    (CXG, F, 1939, free speech) (Corpus 1)

---

3 When presenting examples from our interviews, we give the following information: (<speaker code>, <gender>, <birthyear>, <context of the example>) (<corpus>).

(6) *Ell   ha   deixat   una   planta,   que   so   jo,   i   ara*
    he   have.3sg left   a   plant   that be.1sg   I   and   now
    la   tinc   d'amarar-la   cada   dia
    3sg.fem.acc=have.1sg of water.inf=3sg.fem.acc every day
    «He left behind a plant, which is me, and now I have to water it every day»
                                              (S3, F, 1953, free speech) (Interviews)

(7) *Si   tinc   de   tornar   a   aquella   cançó,   la*
    if   have.1sg   of   return.inf   to   that   song   3sg.fem.acc=
    tenc   de   llegir-la
    have.1sg   of   read.inf=3sg.fem.acc
    «If I have to go back to that song, I have to read it»
                                              (S3, F, 1953, free speech) (Interviews)

(8) *Ha   vist   un   bocí   de   pa   i   se'l*
    have.3sg   seen   a   piece   of   bread   and   3sg.refl=3sg.masc.acc=
    vol   menjar-lo
    want.3sg   eat.inf=3sg.masc.acc
    «He has seen a piece of bread and he wants to eat it»
                                              (S2, M, 1945, free speech) (Interviews)

These examples are similar to patterns of clitic placement in Type 4 languages in the typology given by Cinque (2006) (see Table 1 in §2). In these cases, the clitic that appears in the unclimbed position is interpreted as a trace or copy (see Chomsky 1995) that the clitic in the climbed position has left behind after undergoing movement (e.g., Villa-García 2019). Given that movement (i.e., clitic climbing) has occurred, these examples were grouped with the examples of climbed clitics.

4.2. Coordinated infinitives in modal constructions

In the following three examples, a clitic appears attached to the second of two coordinated infinitives that are complements of the same restructuring verb, see (9) to (11).

(9) *Tot   tenim   de deixar   i   de   mos   estimar   com*
    all   have.1pl   of leave.inf   and   of   1pl.refl=   love.inf   like
    a   germans
    to   brothers
    «We all have to leave and love each other like brothers»
                                              (CAF, F, 1904, folksong) (Corpus 1)

(10) *Fra   Jose ne   vol   l'escuma   i   lo   viure*
brother   Jose PART=   wants   the.foam   and   3SG.MASC.ACC=   live.INF
del   tresor
of.the   treasure
«Brother Jose wants the best of the best and to live in luxury»

<div align="right">(CAF, F, 1904, folksong) (Corpus 1)</div>

(11) *No   podívem   agafar   al   poblet   del   costat   una*
NEG   can.IMPF.1PL   take.INF   to.the   town.DIM   of=the   side   a
cançó   i   fer-la   nostra
song   and   make.INF=3SG.FEM.ACC   ours.FEM
«We couldn't take a song from the neighbouring village and make it our own»

<div align="right">(S11, M, 1974, free speech) (Interviews)</div>

In movement-based accounts of CC, the movement of the clitic from the lexical verb is subject to locality constraints, and in the examples above, the movement of the clitic is blocked by the Coordinate Structure Constraint, according to which a conjunct, or any element contained therein, may not be moved out of the conjunct (Ross 1967). Consequently, the clitic on the second of coordinated infinitival complements in (9) to (11) must remain in situ and clitic climbing is not an option. Given that the focus of this study is on the optionality of clitic climbing, we choose to exclude these examples since they would skew the data towards a preference for clitics to appear in the unclimbed position.

4.3. *Volguer* and non-coreferential infinitival subjects

The following two examples of unclimbed clitics with the volitional verb *volguer* come from the same speaker.

(12) *Aqueix   home   vol   a   hi   posar-li   d'aqueixos*
that   man   want.3SG   to   3SG.DAT=   put.INF=3SG.DAT   of=those
plantons   de   pera   diumenge
saplings   of   pear   Sunday
«That man wants to get given some of those pear saplings on Sunday»

<div align="right">(CAO, M, 1927, free speech) (Corpus 3)</div>

(13) *Si   vol   a   els hi   posar-lis-hi,*
if   want.3SG   to   3SG.DAT=3PL.ASC.ACC=   put.INF=3SG.DAT=3PL.ASC.ACC
lis hi   posem
3SG.DAT=3PL.MASC.ACC=   put.1PL
«If he wants to get given them, we'll give him them»

<div align="right">(CAO, M, 1927, free speech) (Corpus 3)</div>

These examples differ from the other examples of *volguer* to want for two reasons: firstly, the underlying subject of the infinitival complement is non-coreferential with the subject of the main verb; and secondly, the infinitival complements are introduced with a complementiser *a*. These sentences are ungrammatical in other varieties of Catalan, in which *que*-clauses are the only complementation strategy in instances where the verb *voler* «to want» takes a sentential complement with a non-coreferential subject (Gavarró and Laca 2008). Compare (14), where the infinitival complement of *voler* is coreferential with the main verb, to ungrammatical (15) and grammatical (16), where the subject is not.

(14) *No   ho   vull   trencar*   Catalan
     NEG   3.NEUT=   want.1SG   break.INF
     «I don't want to break it»

(15) **No   vull   a   trencar-ho   tu*   Catalan
     NEG   want.1SG   to   break.INF=3.NEUT   you
     «I don't want you to break it»

(16) *No   vull   que ho   trenquis*   Catalan
     NEG   want.1SG   that 3.NEUT=   break.SUBJ.2SG
     «I don't want you to break it»

Notably, the pattern represented by (12), (13) and (15) is an available alternative to *que*-complementation in Sardinian, where an infinitival complement whose subject is not coreferential is always introduced by either *a* or *de* (Jones 1993: 268). Compare (17), an infinitival complement of the verb *kérrere* «to want» with a coreferential subject, to (18), where the subject of the infinitival complement is non-coreferential:

(17) *Non lu   keljo   secare*   Sardinian
     NEG 3.NEUT= want.1SG   break.INF
     «I don't want to break it»

(18) *Non keljo   a   lu   secare   tue*   Sardinian
     NEG want.3.SG to   3.NEUT= break.INF you
     «I don't want you to break it»

In (18), Jones (1993) analyses the verb *kérrere* as behaving as a lexical verb which introduces an infinitival complement in a biclausal structure, evidence of which lies in the appearance of a complementiser, *a* or *de*, and the unavailability of CC. Moreover, Sardinian, and the Logudorese

variety in particular, has been identified as a very important contact language for Algherese (Bosch 2002; among others). The structural similarity between the Sardinian construction and the Algherese examples and the likelihood that the Algherese examples are the result of a syntactic calque of the Sardinian construction make it likely that clitic climbing is not an available option in examples (12) and (13). Thus, examples (12) and (13) are excluded from the analysis for the same reasons as the examples in §4.2.

4.4. Results with outliers excluded

Following the exclusion of the examples in §4.2 and §4.3, the total number of tokens falls to 78. The distribution of CC in restructuring constructions is visualised in Table 6.

Table 6. Distribution of climbed versus unclimbed clitics in the sample with outliers removed

Clitic position	Frequency	Percentage (%)
Climbed	72	92.3
Unclimbed	6	7.7
Total	78	100

The 6 examples of unclimbed clitics come from 5 speakers (S1, S6, S10, S11, and S14), which means that we found no examples of unclimbed clitics in modal constructions in the online corpora and all 34 clitics in modal constructions from the corpora appeared in the climbed position. Below are the examples of unclimbed clitics in the sample, (19) to (24).

(19) *Vols   anar   -hi?*
    want.2SG   go.INF   =LOC
    «Do you want to go there?»

                                            (S1, M, 1967, free speech) (Interviews)

(20) *Té   de   se   l'inserir-la   com*
    have.3SG   of   REFL=   3SG.FEM.ACC=insert.INF=3SG.FEM.ACC   as
    a   llengua   paral·lela   a   l'italià
    to   language   parallel   to   the=Italian
    «They have to insert it as a parallel language to Italian»

                                            (S6, F, 1950, free speech) (Interviews)

(21) *Envia   un   correu,   o   fes   com   vols,   tens*
  send.IMP.2SG an   email   or   do.IMP.2sg   as   want.2SG   have.2SG
  de   contractar-lo
  of   contact.INF=3SG.MASC.ACC
  «Send an email or do what you like, you have to contact him»
<div align="right">(S10, M, 1950, free speech) (Interviews)</div>

(22) *Lego,   ve   lo   funcionari   del   municipi   del   temps*
  then   come.3SG   the   public.servant of.the   town.hall of.the   time
  i   no   sap   llegir-lo
  and   NEG   know.3SG   read.INF=3SG.MASC.ACC
  «Then the public servant from the town hall comes and he can't read it»
<div align="right">(S10, M, 1950, free speech) (Interviews)</div>

(23) *Si   entopes   una   persona   que   empra   la   paraula   cansat,*
  if   meet.2SG   a   persona   that   use.3SG   the   word   tired
  tens   de la   fotografiar-la
  have.2SG   of 3SG.FEM.ACC=   photograph.INF=3SG.FEM.ACC
  «If you meet someone who uses the word "cansat", you have to photograph
  them»
<div align="right">(S11, M, 1974, free speech) (Interviews)</div>

(24) *Jo   no   sep   llegir-lo   i   no   sep*
  I   NEG   know.1SG   read.INF=3SG.MASC.ACC   and   NEG   know.1SG
  escriure   l'alguerés
  write.INF   the.Algherese
  «I don't know how to read it and I don't know how to write Algherese»
<div align="right">(S14, F, 1975, free speech) (Interviews)</div>

4.5. Controlling for crosslinguistic interference

In this subsection, we implement the measures that aimed to control for crosslinguistic interference from Italian outlined in §3.1 and analyse the sample identified in §4.4. As stated in §3.2, the examples from the online corpora are excluded from this part of the analysis due to the lack of sociolinguistic information on the speakers.

4.5.1. Relative proficiency

The distribution of clitics in restructuring constructions by the groups based on the relative proficiencies of the participants' languages are shown in Table 7 below.

Table 7. Distribution of climbed versus unclimbed clitics in modal constructions by dominant language

Dominant language	Climbed	Unclimbed	Total
Algherese	22	0	22
Balanced	3	2	5
Italian	13	4	17
Total	38	6	44

$\chi^2(2)$ = 7.8374, p=0.02 (p<0.05)

4.5.2. Frequency of use

The distribution of clitics in modal constructions by the groups based on the speakers' most frequently used language are shown in Table 8 below.

Table 8. Distribution of climbed versus unclimbed clitics in modal constructions by most frequently used language

Most frequent language	Climbed	Unclimbed	Total
Algherese	22	0	22
Italian	16	6	22
Total	38	6	44

$\chi^2(2)$ = 6.9474, p=0.008 (p<0.01)

5. DISCUSSION

The results show that every speaker in the study demonstrated a very strong preference for clitics to appear in the climbed position in modal constructions in spontaneous speech data. Across the whole sample, 92.3% of clitics appeared in the climbed position while just 7.7% of clitics appeared in the unclimbed position. These findings, while novel, suggest that Algherese is a Type 1 language (see Table 1, §2) like most other varieties of Catalan, in which the climbed position is also preferred in colloquial spoken registers (Gavarró and Laca 2008). Additionally, the pattern in modern Algherese is similar to that found in Old Catalan texts from the 14th century, in which CC in modal constructions is optional but strongly preferred (Batllori et al. 2005; cf. Fischer 2000, 2002; Lee 2022).

Given that Catalan was introduced to Alghero following the conquest of the city at the hands of the Crown of Aragon in 1354, it seems that the typological classification of the Catalan variety in Alghero as a Type 1 language has remained diachronically stable.

However, a closer look at the results suggests that Algherese is instead a Type 2 language. Firstly, all 34 examples taken from the online corpora showed clitics in the climbed position without exception. Secondly, only 6 examples of unclimbed clitics were found in the data, all of which were produced in our interviews by speakers that were predicted to show crosslinguistic interference from Italian, a Type 1 language (see Table 1, §2). The differences in behaviour between the groups when controlling for crosslinguistic interference from Italian reached statistical significance. Given that those speakers whose Italian was dominant or who used Italian more often in daily life showed optionality while those whose Algherese was dominant or who used Algherese more often in daily life did not, it is likely that the optionality in CC has emerged in their grammars due to language contact.

The claim that Algherese is a Type 2 language is also plausible from a sociohistorical perspective due to intensive language contact between Algherese and Sardinian, another Type 2 language (see Table 1, §2). Researchers have identified the influence of Sardinian, and the Logudorese variety in particular, in all aspects of Algherese, such as vocabulary (see Bosch 1997), phonology (see Corbera 2003: 323), and morphosyntax (see Bosch 1998: 151), among others. To explain this pervasive influence, Bosch (2002: 27) proposes the Sardinian Substrate Hypothesis, which postulates that Algherese diverges from other varieties of Catalan due in large part to the mass acquisition of Catalan as a second language (L2) by first-language (L1) Sardinian speakers from the 15th century until the mid-20th century, but particularly during the 16th and 17th centuries. The mass acquisition of Algherese as an L2 by L1 Sardinian speakers would entail the transfer of linguistic patterns from Sardinian, such as the obligation for CC in modal constructions.

6. CONCLUSION

The data and analysis in this study show that clitic climbing in modal constructions in Algherese is obligatory, which makes Algherese a Type 2 variety under the typology of clitic climbing proposed for the Romance languages by Cinque (2006). However, optionality in clitic climbing in modal constructions has emerged in the grammars of some speakers due to language contact with Italian. All speakers in the online corpora produced clitics in the climbed position without exception. Only 6 examples of unclimbed clitics were found in the data from our interviews, all of which were produced by speakers who were predicted to show higher levels of crosslinguistic interference from Italian based on two variables: language dominance and frequency of use.

The typological classification as a Type 2 language distinguishes Algherese from most other varieties of Catalan, which are Type 1 (with the exception of some varieties of Balearic Catalan which are Type 2 as well). Intensive language contact between Algherese and Sardinian, another Type 2 language, in the past (i.e., the Sardinian Substrate Hypothesis) lends plausibility to the claim as well as partially explains the divergence between Algherese and most varieties of Catalan. Thus, the central claim of this work is borne out empirically and plausible socio-historically.

However, further research is required. Most glaringly, the claim that language contact with Sardinian since the 14th century caused Algherese to undergo a diachronic change from a Type 1 to a Type 2 language requires a diachronic analysis of CC in Sardinian, and the Logudorese variety in particular. Such a study has not yet been published to our knowledge. Secondly, a mechanism for the transfer of obligation or optionality of CC is required to determine the plausibility of the claim that language contact is the cause of the diachronic change of Algherese from a Type 1 language in the 14th century to a Type 2 language currently. This, in itself, presupposes a detailed synchronic analysis of optionality in CC (see Sitaridou 2022 for a claim that CC optionality cannot necessarily be attributed to a single internal (syntactic) or sociolinguistic variable), and CC more generally, something that is yet to crystallise in the literature (for a review, see Sitaridou et al. 2015). Thus, the study of Algherese morphosyntax both in synchrony and diachrony, something that has been understudied in the literature to

date, has potential to inform our understanding of not only clitic climbing but morphosyntax more generally.

REFERENCES

Argenter, Joan A. 2008. L'Alguer (Alghero), a Catalan linguistic enclave in Sardinia. *International Journal of The Sociology of Language* 193/194: 205-217. https://doi.org/10.1515/IJSL.2008.056

Ballone, Francesc. 2016. Sobre la correcta interpretació dels fonemes vocàlics de l'alguerès: un estudi acústic. *Caplletra. Revista Internacional de Filologia* 60: 55-81. https://doi.org/10.7203/caplletra.60.7925

Ballone, Francesc. 2017. *Els Usos Lingüístics a l'Alguer, 2015*. Biblioteca Tècnica de Política Lingüística. Generalitat de Catalunya, Direcció General de Política Lingüística.

Batllori, Montserrat, Narcís Iglésias and Ana Maria Martins. 2005. Sintaxi dels clítics pronominals en català medieval. *Caplletra* 38: 137-177.

Blasco Ferrer, Eduardo. 1984. *Grammatica Storica del Catalano e dei suoi dialetti con speciale riguardo all'Algherese*. Tübingen: Gunter Narr Verlag.

Bonet, Eulàlia. 2008. Cliticització. In Joan Solà, Maria-Rosa Lloret, Joan Mascaró, and Manuel Pérez Saldanya (dirs.), *Gramàtica del català contemporani*, 4th ed., 1: 933-89. Barcelona: Editorial Empúries.

Bosch i Rodoreda, Andreu. 1997. L'alguerés i les seves relacions amb el sard. *Serra d'Or* 451-452: 45-46.

Bosch i Rodoreda, Andreu. 1998. Aproximació a l'alguerès col·loquial. *Revista de l'Alguer* 9 (9): 137-163.

Bosch i Rodoreda, Andreu. 2002. *El català de l'Alguer*. Barcelona: Publicacions de l'Abadia de Montserrat.

Bosch i Rodoreda, Andreu. 2012. *El lèxic alguerès de l'agricultura i la ramaderia entre els segles XVII i XVIII*. Biblioteca Filològica 68. Barcelona: Institut d'Estudis Catalans.

Bosch i Rodoreda, Andreu and Luca Scala. 1999. Variació fonètica i cliticització pronominal en alguerès: una primera aproximació intradialectal. *Caplletra* 26: 107-130.

Cardinaletti, Anna and Ur Shlonsky. 2004. Clitic Positions and Restructuring in Italian. *Linguistic Inquiry* 35 (4): 519-57. https://doi.org/10.1162/0024389042350523

Caria, Rafael. 2006. El català a l'Alguer: apunts per a un llibre blanc. *Revista de Llengua i Dret* 46: 29-102.

Chessa, Enrico. 2008. Dialect contact, Koinéisation, and new dialect formation on the North-western Coast of Sardinia. In Alexandre Nikolaev and Jussi Niemi (eds.), *Two or More Languages: Proceedings from the 9th Nordic Conference on Bilingualism, August 10-11, 2006, Joensuu*, 27-36. Joensuu: University of Joensuu.

Chomsky, Noam. 1995. *The Minimalist Program / Noam Chomsky.* Cambridge, Mass: MIT Press.

Cinque, Guglielmo. 2006. «Restructuring» and Functional Structure. In Guglielmo Cinque (ed.), *Restructuring and Functional Heads*, 11-64. The Cartography of Syntactic Structures, v. 4. Oxford / New York: Oxford University Press.

Corbera, Jaume. 1993. La influència de l'italià i dels parlars sards i itàlics en les denominacions alguereses dels animals marins. In Antoni Ferrando Francés, Lluís Meseguer and Rafael Alemany Ferrer (coords.), *Actes del novè Col·loqui Internacional de llengua i Literatura Catalanes, Alacant-Elx*, 9-14 setembre de 1991, II, 513-528. Barcelona: Publicacions de l'Abadia de Montserrat.

Corbera, Jaume. 2000. *Caracterització del lèxic alguerès: contribució al coneixement del lèxic alguerès modern*. Palma: Universitat de les Illes Balears.

Corbera, Jaume. 2003. L'alguérais: un dialecte de frontière du catalan. *Revue de Linguistique Romane* 67 (267-268): 321-41.

D'Alessandro, Roberta, David Natvig and Michael T. Putnam. 2021. Addressing Challenges in Formal Research on Moribund Heritage

Languages: A Path Forward. *Frontiers in Psychology* 12 (July). https://doi.org/10.3389/fpsyg.2021.700126

Fischer, Susann. 2000. Obligatory Clitic Climbing in Old Catalan. In Hans-Martin Gärtner (ed.), *Two Papers on Clitics*, 63-77. Linguistics in Potsdam 9. Potsdam: Univ.-Bibliothek, Publikationsstelle.

Fischer, Susan. 2002. *The Catalan Clitic System: A Diachronic Perspective on its Syntax and Phonology*. Berlin / New York: Mouton de Gruyter.

Gavarró, Anna and Brenda Laca. 2008. Les perífrasis temporals, aspectuals i modals. In Joan Solà, Maria-Rosa Lloret, Joan Mascaró and Manuel Pérez Saldanya (dirs.), *Gramàtica del català contemporani*, 4th ed., 3: 2663-2728. Barcelona: Editorial Empúries.

Grosjean, François. 2010. *Bilingual: Life and Reality*. Cambridge, Mass: Harvard University Press.

Institut d'Estudis Catalans. 2003. *El català de l'alguer: un model d'àmbit restringit*. Biblioteca filològica, XLVIII. Barcelona: Institut d'Estudis Catalans.

Jones, Michael Allan. 1993. *Sardinian Syntax*. Romance Linguistics. London / New York: Routledge.

Kayne, Richard S. 1975. *French Syntax: the Transformational Cycle.* Cambridge, Mass.: MIT Press.

Kayne, Richard S. 1989. Null Subjects and Clitic Climbing. In Osvaldo Jaeggli and Kenneth J. Safir (eds.), *The Null Subject Parameter*, 239-261. Studies in Natural Language and Linguistic Theory, v. 15. Dordrecht / Boston: Kluwer Academic Publishers.

Lee, Tristan. 2022. Clitic climbing in Algherese Catalan: Empirical findings and analysis. MPhil Thesis. University of Cambridge.

Mensching, Guido. 2017. Morfosintassi: sincronia. In Eduardo Blasco Ferrer, Peter Koch, and Daniela Marzo (eds.), *Manuale di linguistica sarda*, 376-96. Manuals of Romance linguistics 15. Berlin / Boston: De Gruyter.

Pais i Melis, Joan. 1970. *Gramàtica algueresa: Transcripció, introducció i notes de Pasqual Scanu*. Barcelona: Barcino.

Palomba, Giovanni. 2001. *Grammatica del dialetto algherese odierno*. L'Alguer: Obra Cultural de l'Alguer.

Paradis, Michel. 2004. *Neurolinguistic Theory of Bilingualism.* Amsterdam: John Benjamins Publishing Company.

Parry, M. 1994. Posizione dei clitici complemento nelle costruzioni verbali perifrastiche del piemontese. In Gianrenzo P. Clivio and Censin Pich (eds.), *At Dël VIII Rëscontr Antërnassional Dë Studi an Sla Lenga e La Literatura Piemontèisa*, 247-259. Alba: Famija Albèisa.

Perea, Maria-Pilar. 2010. The Dialect of Alghero: Continuity and Change. In Robert McColl Millar (ed.), *Marginal Dialects: Scotland, Ireland and Beyond*, 1: 131-49. Aberdeen: University of Aberdeen.

Polinsky, Maria. 2018. *Heritage Languages and Their Speakers*. 1st ed. Cambridge University Press. https://doi.org/10.1017/9781107252349

Rivas, A. M. 1977. *A Theory of Clitics*. Ph.D. thesis, Massachusets Institute of Technology.

Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Studies in Generative Grammar, v. 11. Dordrecht, Holland / Cinnaminson, N.J., U.S.A: Foris Publications.

Roberts, Ian. 2016. Object Clitics. In Adam Ledgeway and Martin Maiden (eds.), *The Oxford Guide to the Romance Languages*, 786-801. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199677108.003.0048

Ross, John. 1967. *Constraints on Variables in Syntax*. Ph.D. thesis, Massachusets Institute of Technology.

Schmid, Monika S. and Barbara Köpke. 2017. The relevance of first language attrition to theories of bilingual development. *Linguistic Approaches to Bilingualism* 7 (6): 637-667. https://doi.org/10.1075/lab.17058.sch

Sitaridou, Ioanna. 2022. La subida de los clíticos en rioplatense: mecanismos universales, variación y opcionalidad. En Ángela Di Tullio, Enrique Pato (eds.), *Universales vernáculos en la gramática del español*, 1-35. Madrid: Iberoamericana Vervuert.

Sitaridou, Ioanna, Helen Whimpanny and Laura Ayres. 2015. Variation and Optionality in Clitic Climbing in Argentinean Spanish. *Isogloss. A Journal on Variation of Romance and Iberian Languages* 1 (2): 247-291. https://doi.org/10.5565/rev/isogloss.21

Tortora, Christina. 2014. Patterns of Variation and Diachronic Change in Piedmontese Object Clitic Syntax. In Paola Benincà, Adam Ledgeway, and Nigel Vincent (eds.), *Diachrony and Dialects: Grammatical Change in the Dialects of Italy*, First edition, 218-240. Oxford Linguistics 8. Oxford: Oxford University Press.

Villa-García, Julio. 2019. Clitic climbing (or lack thereof) and the Copy Theory of Movement. *Glossa: A Journal of General Linguistics* 4 (1): 84. https://doi.org/10.5334/gjgl.965