

**The syntax of subject pronouns
in heritage languages**
Innovation and complexification

Published by
LOT
Binnengasthuisstraat 9
1012 ZA Amsterdam
The Netherlands

phone: +31 20 525 2461

e-mail: lot@uva.nl
<http://www.lotschool.nl>

ISBN: 978-94-6093-407-0
DOI: <https://dx.medra.org/10.48273/LOT0622>
NUR: 616

Copyright © 2022: Alberto Frasson. All rights reserved.

**The syntax of subject pronouns
in heritage languages**
Innovation and complexification

**De syntaxis van persoonlijke voornaamwoorden in erfgoedtaalen:
verandering en complexificatie**
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

woensdag 15 juni 2022 des ochtends te 10.15 uur

door

Alberto Frasson

geboren op 11 januari 1988
te Treviso, Italië

Promotoren:

Prof. dr. Roberta D'Alessandro Universiteit Utrecht
Prof. dr. Martin Everaert Universiteit Utrecht

Beoordelingscommissie

Prof. dr. Enoch Aboh Universiteit van Amsterdam
Prof. dr. Norbert Corver Universiteit Utrecht
Prof. dr. Marjo van Koppen Universiteit Utrecht
Dr. Diego Pescarini CNRS, Université Côte d'Azur
Prof. dr. Michelle Sheehan Newcastle University

This research has received funding by the European Research Council under the Horizon 2020 research and innovation programme (grant agreement CoG 681959_MicroContact).

*And if I fly, or if I fall
Least I can say, I gave it all*

RuPaul Andre Charles

Table of contents

Acknowledgments	<i>xi</i>
List of abbreviations	<i>xiii</i>
Chapter 1. Introduction	1
1.1. Heritage languages and contact	1
1.2. How to analyse contact-induced change in Italo-Romance varieties	2
1.2.1. <i>Feature-reassembly and the bottleneck hypothesis</i>	7
1.2.2. <i>Subject pronouns: simplification, stability and complexification</i>	8
1.3. Information structure and syntax: two approaches	11
1.3.1. <i>The cartographic approach to the left periphery</i>	12
1.3.2. <i>Towards a feature-based approach</i>	15
1.3.3. <i>Beyond C: discourse features in v and D</i>	17
1.4. Research questions and structure of the dissertation	18
1.4.1. <i>Research questions</i>	18
1.4.2. <i>Chapter outline</i>	19
Chapter 2. Referentiality and the internal structure of subject pronouns	23
2.1. Introduction	23
2.2. Null and overt subjects: an overview	25
2.3. The internal structure of pronouns	29
2.3.1. <i>Introducing the [R] feature</i>	32
2.4. The interpretation of null and overt subjects: topic, linking and logophoricity	34
2.4.1. <i>Frascarelli (2007): The A-topic and the Topic Criterion</i>	35
2.4.2. <i>Sigurðsson (2011, 2014): Context Scanning</i>	38
2.4.3. <i>The Logophoric Centre and the [R] feature: accessibility and salience</i>	40
2.4.4. <i>The assignment of [R]</i>	45
2.5. The derivation of null and overt subjects	47
2.5.1. <i>Summary</i>	51
2.6. Subject realisation in heritage languages	52

viii The syntax of subject pronouns in heritage languages

2.6.1. <i>The interpretation of null subjects in heritage Italo-Romance</i>	53
2.6.2. <i>Heritage Italo-Romance varieties as consistent null subject languages</i>	54
2.6.3. <i>Topic shift and switch reference: the role of [uR]</i>	71
2.7. Conclusion	74
Chapter 3. The internal structure and the status of subject clitics	77
3.1. Introduction	77
3.2. Subject clitics and [R]	79
3.2.1. <i>Friulian and Venetan subject clitics</i>	80
3.3. Subject clitics, pragmatics and discourse	86
3.3.1. <i>The restriction on doubling</i>	86
3.3.2. <i>The Venetan discourse-related clitic</i>	87
3.3.3. <i>The Štivorian context-related expletive clitic</i>	89
3.3.4. <i>Subject clitics in Brazilian Venetan</i>	91
3.4. The interpretation of subject clitics in heritage Friulian and Venetan	94
3.4.1. <i>Presentation of the study: tasks and methodology</i>	94
3.4.2. <i>Results of the study</i>	98
3.5. Subject clitics as pronouns	108
3.5.1. <i>Saliency and discourse</i>	109
3.5.2. <i>The syntax of subject clitics</i>	114
3.5.3. <i>Why subject clitics do not get deleted at PF: adjacency to the verb</i>	116
3.5.4. <i>Summary</i>	125
3.6. Open questions	125
3.6.1. <i>Second person</i>	126
3.6.2. <i>Interrogative inversion</i>	128
3.7. Conclusion	131
Chapter 4. The position of the antecedent of pronominal, clitic and null subjects	133
4.1. Introduction	133
4.2. Pronoun interpretation and the position of the antecedent	135
4.2.1. <i>Previous studies on pronoun interpretation</i>	135
4.2.2. <i>The [uR] feature and antecedent selection</i>	140
4.2.3. <i>Summary</i>	144

4.3. The study	144
4.3.1. <i>Research question and hypotheses</i>	144
4.3.2. <i>The questionnaire</i>	148
4.3.3. <i>Results</i>	151
4.4. Antecedent selection and [uR]	163
4.4.1. <i>Antecedent selection properties of overt and null subjects</i>	164
4.4.2. <i>The interpretation of subject clitics in subordinate clauses</i>	167
4.4.3. <i>A note on the notion of “preference”</i>	171
4.4.4. <i>Summary</i>	173
4.5. The role of the contact languages	173
4.5.1. <i>The PAH in contact languages</i>	174
4.5.2. <i>Feature-reassembly in contact: a possible solution</i>	177
4.6. Conclusion	178
Chapter 5. Decomposing Topic: the Saliency Structure Hypothesis	181
5.1. Introduction	181
5.2. Challenging the PAH: the role of discourse structure	182
5.2.1. <i>The Saliency Structure Hypothesis</i>	185
5.2.2. <i>A further application of [uR]</i>	188
5.2.3. <i>Summary</i>	190
5.3. The study	191
5.3.1. <i>Research questions and hypotheses</i>	191
5.3.2. <i>The questionnaire</i>	192
5.3.3. <i>Results</i>	196
5.4. The saliency structure and [uR]: an extension of the model	207
5.4.1. <i>The syntactic encoding of [uR] and saliency structure</i>	208
5.4.2. <i>Subject clitics and non-argument antecedents</i>	211
5.5. Further consequences for the analysis of contact-induced change	216
5.5.1. <i>Innovation and complexification in heritage languages</i>	217
5.6. Conclusion	218
Chapter 6. Conclusion	221
6.1. Summary of the thesis	221

x	The syntax of subject pronouns in heritage languages	
	6.1.1. <i>Chapter overview</i>	221
	6.2. Suggestions for future research	227
	6.2.1. <i>Beyond subjects: the salience of object pronouns</i>	227
	6.2.2. <i>Subject resumption in relative clauses</i>	229
	6.2.3. <i>The problem of intra-speaker variation</i>	230
	6.3. General conclusion	232
	References	235
	Appendix A - Informants	249
	Appendix B – Tasks and questionnaires	257
	Summary	279
	Samenvatting in het Nederlands	285
	Curriculum vitae	291

Acknowledgments

This dissertation was written entirely during the Covid-19 pandemic, mostly from my apartment in Utrecht and occasionally from Hannover and Treviso, when travelling restrictions allowed me to visit my beloved ones. I want to thank all the people who supported me professionally and personally during this hard, lonely but extremely productive period of my life.

First of all, my supervisor and promotor Roberta D'Alessandro and my co-supervisor Martin Everaert for the inspiring conversations on syntax and especially for encouraging me to fight for my ideas. I owe a special acknowledgment to two more people, who helped me becoming a better linguist and a better scientist: Jan Casalicchio, my daily supervisor, for teaching me everything I know about the syntax of subject clitics and for showing me how to carry out scrupulous syntactic analyses; Brechje van Osch, for awakening in me the curiosity for data analysis and statistics. Finally, I would like to thank Luigi Andriani and Manuela Pinto, who joined the Microcontact Team in the second part of my PhD, for the very useful suggestions and discussions.

Throughout the years, I had the pleasure of discussing my ideas with brilliant linguists: Enoch Aboh, Artemis Alexiadou, Theresa Biberauer, Norbert Corver, Caterina Donati, Luis López, Angel Jiménez-Fernández, Marjo van Koppen, Maria Carme Parafita Couto, Diego Pescarini, Francisco Ordóñez, Jason Rothman, Andrés Saab, Michelle Sheehan, Juan Uriagereka. Thank you for your valuable feedback and discussion.

I would like to thank all the people with whom I shared the past four years at UiL-OTS. My fellow PhD Candidates from the Microcontact project, Luana Sorgini and Silvia Terenghi; my paranymphs, Sonya Ros and Mariano González; my office mates, Imke Kruitwagen, Nina Sangers and Iris Selten; all the colleagues at UiL-OTS, in particular: Myrthe Bergstra, Ngoc Doan, Emma Everaert, Rachida Ganga, Na Hu, Bambang Kartono, Anika van der Klis, Martijn van der Klis, Chou Mo, Sonya

xii The syntax of subject pronouns in heritage languages

Nikiforova, Giada Palmieri, Tijn Schmitz, Florentine Sterk, Joanna Wall, Jan Winkowski, Yuan Xie.

A special acknowledgment goes to the people who made my research possible: my informants and all the people that helped me in the data collection. In particular: Eduardo, Richard, Marcos and the Venetan Association of Rio Grande do Sul in Brazil; Ernestina, Silvina and the Friulian Society of Buenos Aires in Argentina. Obrigado! Gracias! Grazie! Graciis!

I would like to thank my first and most important Venetan informant: my grandfather Luigi ‘Gigeto’. You left us some months before the end of my PhD, but I know you would have been proud of my job.

The most important acknowledgment goes however to my beloved family: my parents Antonella and Enzo and my sister Arianna. Being away from you was hard, but you have never let me feel alone. I could never thank you enough for making me the person I am, for always being there for me when I needed you and for always helping me finding the right way. Vi voglio due mondi di bene.

Finally, the person I love the most, Željko. Meeting you was the best thing that could ever happen to me, I would not have been able to get to the end of this PhD without you by my side, without your strength and support. I cannot find the words to explain how much I admire you and love you. Or maybe I can: тамо где си ти, тамо станује ми душа и срце моје је тамо где си ти. Волим те.

List of abbreviations

1	1 st person
2	2 nd person
3	3 rd person
A	aboutness
Adv	adverb
Aux	auxiliary
C	complementiser
Cl	clitic
CLn	C/edge linker
Cont	topic continuation
D	determiner
Dat	dative
Dem	demonstrative
DOM	differential object marking
Fam	familiar topic
Fin	finite
Foc	focus
Imp	imperative
Inf	infinitive
Infl	inflection
LC	logophoric centre
LI	lexical item
Loc	locative

xiv The syntax of subject pronouns in heritage languages

N	noun
Neg	negation
Num	number
OCI	object clitic
PAH	Position of the Antecedent Hypothesis
Pl	plural
Pron	pronoun
Prt	past participle
R	referential
Refl	reflexive
Sg	singular
SCI	subject clitic
SL	speech location
ST	speech time
T	tense
Top	topic
V	verb
v	light verb
ΛA	logophoric agent
ΛP	logophoric patient

Chapter 1. Introduction

1.1. Heritage languages and contact

The aim of this study is to investigate syntactic change in situations of language contact. The languages included in the study qualify as heritage languages and are spoken in Italy and by communities of descendants of Italian immigrants in Argentina, Belgium, Brazil and Canada. Rothman (2009: 156) defines a heritage variety as a language that is not a dominant language of the larger society, spoken mainly at home in a familiar context.

All heritage speakers are either sequential or simultaneous bilinguals (Polinsky 2018: 2) and the dominant language of the larger national society may affect their heritage language in different ways and at different levels. Therefore, according to Rothman (2009: 156), heritage speakers have some command of their native heritage language acquired naturalistically, but their competence differs from that of native monolinguals as a consequence of language contact.

The idea that heritage competence in the native language differs from monolingual competence often led linguists to conclude that heritage speakers are unbalanced bilinguals, since they use the heritage language only in very restricted situations and circumstances. Their weaker language is in fact their native language, while the stronger language is the dominant language of the society (Polinsky 2018: 4).

Studies on heritage languages in contact generally compare the variety spoken by heritage speakers to that of baseline speakers and homeland speakers. Baseline speakers correspond to the first-generation immigrants, who constitute the primary (and often only) input for heritage speakers; for instance, the language spoken by first-generation Russian immigrants in the United States constitutes the baseline system for the variety of Russian spoken by their children (or grandchildren), born

2 The syntax of subject pronouns in heritage languages

and raised in the United States. Homeland speakers are monolingual speakers of the language in the country of origin: unlike heritage speakers, homeland speakers live in an environment in which their language is the dominant one in the society and is therefore used at all levels; for instance, homeland Russian is the variety of the language spoken by monolingual speakers in Russia.

The methodology in heritage linguistics generally requires heritage speakers to be compared to baseline and homeland speakers to identify a change in their language.

However, as shown by D'Alessandro et al. (2021), these speakers are not always available; for example, it is possible that there are no first-generation baseline speakers left in the community; it is also possible that there is no larger monolingual community of speakers in the country of origin: for instance, the language may have a minority status in its native settings as well.

In these cases, any attempt to compare heritage speakers to baseline and homeland speakers would be flawed, referring to first-generation speakers that are not the real baseline for heritage speakers, or to homeland speakers that are also dominant in a different language. The cases presented in this dissertation represent precisely this type of language contact situation. It will be shown that it is still possible to account for contact-induced-change in such varieties, but the issues of how contact with another language drives change needs to be tackled from a different perspective.

1.2. How to analyse contact-induced change in Italo-Romance varieties

All the languages included in this study are regional Romance varieties originally spoken in the Italian Peninsula. This area presents considerable linguistic variation, in that a large number of local languages are spoken alongside the national language, Italian. These varieties are generally referred to as 'Italian dialects', a definition that may lead to think that they represent local varieties of Italian. However, these varieties developed from Latin parallel to Italian; they are therefore better defined as autonomous Romance varieties that are related to Italian in that they all evolved from

Latin. Pellegrini (1977) identified five main groups of Romance varieties spoken in Italy: northern varieties (including Gallo-Italic varieties spoken in north-western Italy and Venetan varieties spoken in north-eastern Italy), Friulian varieties, Tuscan varieties, central-southern varieties (divided into median, upper-southern and extreme southern varieties) and Sardinian varieties. More recently, Maiden and Parry (1997) proposed a division of regional languages of Italy in four groups: northern, central, upper southern and extreme southern languages. This dissertation focusses on seven varieties: two northern languages (Friulian, Venetan), three upper southern languages (Eastern Abruzzese, Neapolitan, Northern Calabrian) and two extreme southern languages (Southern Calabrian, Sicilian). The distribution of these languages in Italy is shown in Figure 1.

4 The syntax of subject pronouns in heritage languages

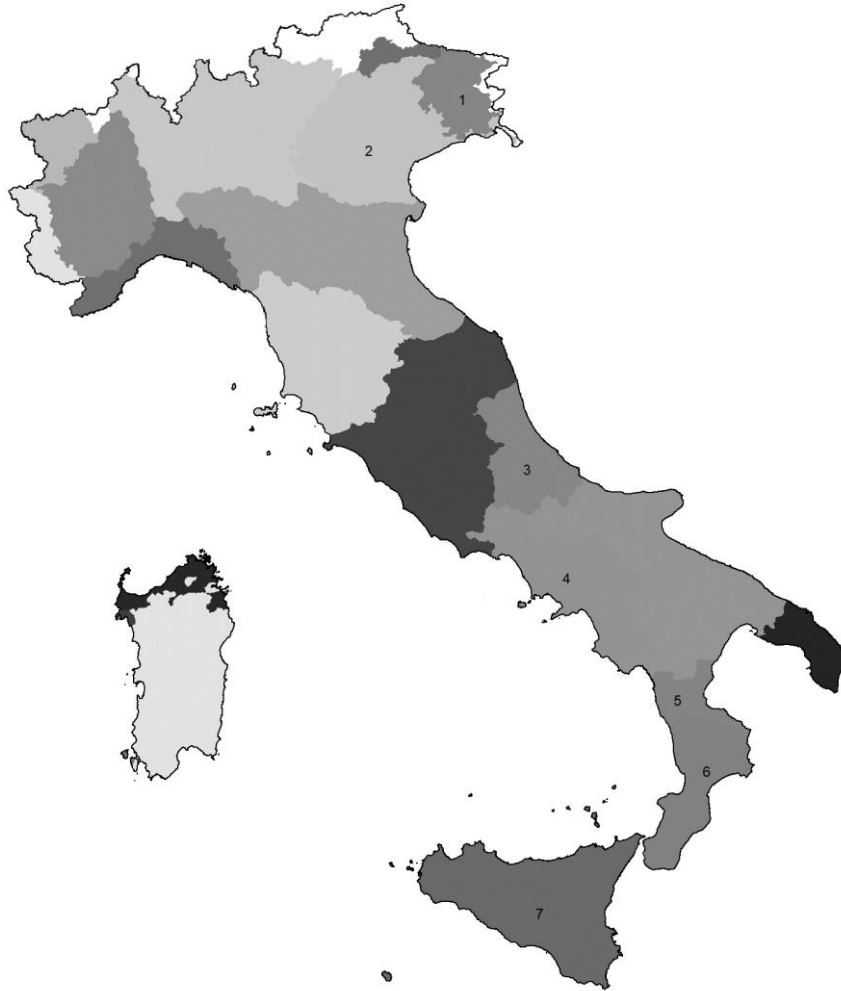


Figure 1. Simplified map of the languages spoken in Italy. The varieties considered in the study are Friulian (1), Venetan (2), Eastern Abruzzese (3), Neapolitan (4), Northern Calabrian (5), Southern Calabrian (6), Sicilian (7). Other Romance varieties of Italy are not numbered and marked in grey. Non-Romance varieties of Italy are marked in white. Consider that microvariation in Italy is much more fine-grained than what is shown by this map and several minor local varieties could not be represented.

These languages have different statuses in Italy. The only language with an official recognition as a minority language among them is Friulian. This fact implies that it is

also the only language with an accepted standard variety, although microvariation in spoken Friulian is considerable. Sicilian and Venetan have a partial regional recognition, while the remaining languages do not have any recognition. Some languages (Venetan, Neapolitan, Sicilian) have a long literary tradition as they used to be official languages in the various republics, principalities and kingdoms that existed in pre-unitary Italy.

All speakers of these varieties nowadays are bilingual, with the regional language being used mainly in informal contexts (at home, with friends) and standard Italian being used in formal or official contexts. The situation in which the two linguistic codes are systematically employed in distinct contexts, is defined as diglossia (Ferguson 1959): one variety is used in formal and public occasions while the other variety is used in everyday life. With respect to the Italian reality, the first variety is represented by standard Italian, while the second variety corresponds to the local Romance language.

However, Berruto (1987) suggested that the Italian linguistic scenario does not perfectly coincide with this definition and should be rather defined in terms of *dilalia*, intended as the case in which the division of domains of use between the two linguistic systems is less clear than that found in cases of diglossia. *Dilalia* differs substantially from diglossia because there are uses and domains in which both Italian and the local variety can be used alternately or jointly.

This situation, described by De Mauro (1970), has been changing rapidly in the past few decades. As shown in Loporcaro (2009), the use of these languages is nowadays mainly limited to the familiar context, most frequently in small towns or villages. Crucially, many younger speakers have only a passive knowledge of their regional language, some of them can understand it but cannot speak it, while others can no longer understand it, being therefore monolingual speakers of Italian. In this respect, regional languages of Italy can be defined as heritage languages in Rothman's (2009) terms, as they are not the dominant languages of the larger language community and they are spoken mainly at home. Besides, speakers of these languages are nowadays generally dominant in Italian, while their native language has a more limited use. In other words, the use of these languages in Italy is more similar to that

6 The syntax of subject pronouns in heritage languages

of heritage languages than that of regional or minority languages. The linguistic situation of Italy is extremely complex and an exhaustive discussion of the status of Italo-Romance varieties is beyond the goal of this dissertation. I refer to Telmon (1992), Grassi et al. (2006) and Tamburelli (2010, 2014) for a complete overview of the Italian linguistic situation.

Aside from the Italian scenario, this study focusses also on regional languages of Italy as spoken by heritage speakers in other countries: Argentina, Belgium, Brazil and Canada. For the sake of clarity, the term “heritage language” will be reserved for varieties spoken outside Italy (e.g. Heritage Calabrian, Heritage Venetan); varieties of the languages spoken in Italy will be referred to as Italian varieties (e.g. Italian Calabrian, Italian Venetan). Specific heritage varieties will be referred to as “Argentinian varieties” (e.g. Argentinian Neapolitan), “Brazilian varieties” (e.g. Brazilian Calabrian), “Quebec varieties” (e.g. Quebec Venetan) and “Belgian varieties” (e.g. Belgian Sicilian).

Heritage speakers of these languages are the descendants of Italian immigrants that maintained, to a certain level, their heritage languages. Some of these languages are still quite widely spoken among communities of Italian descendants; it is the case of Venetan and Calabrian in Brazil or Friulian in Argentina and Brazil. This fact may depend on a variety of extralinguistic factors: the prestige inside the community in the case of Brazilian Venetan, the size of the community in the case of Argentinian Friulian and Brazilian Calabrian or the prolonged isolation from the larger society in the case of Brazilian Friulian. Other varieties qualify as moribund (D’Alessandro *et al.* 2021), in that there are only few speakers left and the language is no longer transferred to younger generations. This is the case of Neapolitan in Argentina or Sicilian in Brazil. There are only few elderly speakers left of these languages; besides, Brazilian Sicilian and Argentinian Neapolitan are very often

strongly influenced by the dominant language, respectively Brazilian Portuguese and Argentinian Spanish¹.

In sum, the study of these Romance varieties is challenged by a number of factors. First of all, there are few speakers left of these varieties outside Italy. Besides, all the varieties, including those spoken in their native setting in Italy, qualify as heritage languages: there is no monolingual homeland variety that can be used as a comparison. Finally, there is no baseline for comparison in communities of descendants of immigrants and any attempt to compare different generations would not provide reliable results: most of the times it is no longer possible to find speakers belonging to the generation that migrated and transmitted the language to current speakers of the language. It is therefore impossible to analyse heritage varieties of the languages in a continuum from the monolingual homeland variety and the baseline variety: the different varieties of the languages under investigation will be analysed separately and compared to each other, without making any prediction on the starting point of the changes.

1.2.1. Feature-reassembly and the bottleneck hypothesis

This study approaches the analysis of heritage languages in relation to their homeland and baseline varieties building on two hypotheses originally proposed for second language acquisition: the Feature-reassembly Hypothesis (Lardiere 2008) and the Bottleneck Hypothesis (Slabakova 2014).

The effects of contact on different varieties of the languages considered in this study are analysed starting from the assumption that all of the varieties underwent language contact and this contact affected the structures of the languages to some extent. Given the impossibility of considering change in the varieties under analysis with respect to a baseline and a homeland variety of the same language, they will all

¹ Specifically for the Argentinian case, the situation evolved rapidly in the past few decades. The situation described in Saltarelli (1977), for instance, was very different from the one we find today.

8 The syntax of subject pronouns in heritage languages

be treated as separate languages that are the result of a process in which contact functions as a bottleneck (Slabakova 2014) or as a blender. Contact breaks down the link between syntactic features and specific lexical items of a language; such links have to be eventually re-established, but features may be reassembled in lexical items in slightly different ways than before contact (Lardiere 2008).

This approach aims to investigate syntactic change separately from dimensions of change such as sociolinguistic and psycholinguistic factors. Three possible directions of syntactic change will be considered: a reduction of complexity, an increase of complexity or a general stability of the system.

1.2.2. Subject pronouns: simplification, stability and complexification

The distribution of subject pronouns in null-subject languages is one of the main topics in the field of studies on bilingualism and heritage languages. Changes in the distribution of subject pronouns are very common and may go in different directions.

Most studies analysed processes of **simplification** in the distribution of subject pronouns in null-subject languages in bilingual speakers (e.g. Sorace and Filiaci, 2006; Sorace and Serratrice 2009; Tsimpli *et al.* 2004; Sorace *et al.* 2009). This is particularly evident when the null-subject language is spoken aside a non-null-subject language. There is a tendency to extend the use of overt subjects to discourse contexts in which null subjects are expected. This simplification depends on the fact that bilingual speakers generally opt for the simpler system (Platzack 2001; Hulk and Müller 2000): while non-null-subject languages have only one option (the overt subject pronoun), null-subject languages have more than one option; the distribution of overt and null subject pronouns depends on the interaction between grammatical and discourse factors. Montrul (2004, 2008) showed that a similar process of simplification is found in heritage speakers of Spanish in contact with English:

(1) Heritage Spanish (Montrul, 2004: 134)

Ella / #*pro* no notaba que el lobo estaba en la cama.
 she *pro* not noticed that the wolf was in the bed
 ‘She did not notice that the wolf was in the bed.’

In heritage Spanish, a referential null subject is pragmatically non felicitous, leading to a more generalised use of overt subject pronouns in cases in which a null subject would be preferred in monolingual Spanish.

In other studies, null subjects in heritage languages seems to be stable (Flores and Rinke, 2020; Nagy et al., 2011; Rinke and Flores, 2018) and the syntactic and discourse-related factors involved in their distribution are easily mastered by heritage speakers (Carvalho and Child, 2011; De Souza et al., 2018). The data presented in this study show that this is also the case of heritage southern varieties, such as Argentinian Neapolitan:

(2) Argentinian Neapolitan

Issə / *pro* non m’ à vistə.
 he *pro* not me has seen
 ‘He did not see me.’

In heritage Neapolitan, both the overt subject pronoun and null *pro* are acceptable, given the correct interpretation. In this case it is possible to talk about the **stability** of heritage languages, as opposed to their simplification. Aikhenvald (2006: 22) shows that if two languages in contact share a construction, contact may reinforce the productivity of a structure. This is precisely the case of (2): both the heritage language, Neapolitan, and the contact language, Spanish, allow for null subjects

A less investigated possibility is the **complexification** of heritage grammars, as the case in which new structures are introduced or additional constraints are added to existing structures in heritage languages. This is the case of heritage northern varieties, such as heritage Venetan, in which the number of possible configurations of different types of subject pronouns is bigger than in Italian Venetan. The

10 The syntax of subject pronouns in heritage languages

configurations available in Brazilian Venetan are presented in (3); notice that Venetan has two overt types of subject pronouns: strong and clitic.

(3) Brazilian Venetan

a. **Lu** no me ga vedesto.

he not me has seen

b. **Lu** no'l me ga vedesto.

he not he.SCL me has seen

c. No'l me ga vedesto.

not he.SCL me has seen

d. **El** no me ga vedesto

he.SCL not me has seen

e. *pro* no me ga vedesto.

pro not me has seen

'He did not see me'.

In (3)a), a subject pronoun *lu* is overtly realised. In (3)b) two subject pronouns appear: the subject pronoun *lu* and the subject clitic (*el*). In (3)c) only a subject clitic is realised after preverbal negation *no*. In (3)d), the same subject clitic is realised before preverbal negation *no*. Finally, in (3)e) there are no overt subject pronouns. Options (3)d) with a subject clitic preceding preverbal negation, and (3)e), without any overt subject pronouns, are not grammatical in Italian Venetan. In Brazilian Venetan, however, all these possibilities are attested, evidencing a complexification of the system. According to Aikhenvald (2006: 32), pre-existing structural similarity or existence of a lookalike in the contact languages are crucial factors in the diffusion of new patterns. The possibility of a complexification of the system will be discussed in detail for Brazilian Venetan in Chapter 3.

This dissertation aims to show that the changes attested in heritage languages with respect to subject realisation, regardless of their direction (towards a simplification or a complexification of the system), depend on the rich set of grammatical properties at the interface between syntax and discourse. At least some

of these properties are encoded in syntax by means of specific discourse-related features. Such features reflect what Miyagawa (2010: 158) defines as the *expressiveness* of language: agreement and movement involve some operation on formal features taking place in narrow syntax; at least some instances of these operations involve features that depend on the way speakers structure the information according to what they want to express to correctly deliver a message to their interlocutors. This study focusses precisely on the discourse-related features involved in the distribution of different types of subject pronouns, investigating how they interact with syntax and with information structure and what happens to such features when languages get in contact with others.

1.3. Information structure and syntax: two approaches

Information Structure was defined by Chafe (1976) as the way the information is packaged by the speaker to satisfy the interlocutors' communicative needs. Kamp (1981) proposed that sentences should not be interpreted individually but should be considered as part of the discourse; the theoretical framework that he developed, Discourse Representation Theory (DRT), is a dynamic semantic approach to the interpretation of phenomena such as anaphora and tense. The fundamental claim of DRT is that we build a mental representation of the discourse while we speak or listen and every sentence adds some information to our mental representation. Interpretation processes can never take in individual sentences but must take into account the previous discourse.

In this respect, one fundamental notion of information structure is Common Ground, which can be divided into Common Ground Content, the information mutually shared by speaker and addressee (Stalnaker 1974), and Common Ground Management, the way this information is continuously modified and updated (Krifka 2008).

Within the generative framework, several influential approaches to the role of information structure in grammar were proposed, focussing in particular on the role of CP and on the concepts of topicality, anaphoricity and accessibility. The main

12 The syntax of subject pronouns in heritage languages

claims made by the cartographic approach (Rizzi 1997; Rizzi and Shlonsky 2007) and by the minimalist approach (Chomsky 1995, 2008; Miyagawa 2005, 2010) are summarised in this section.

1.3.1. The cartographic approach to the left periphery

The observation that discourse-related factors have an effect on the syntax of a clause is fundamental in the cartographic approach to the structure of the left periphery of the sentence, developed in Rizzi (1997) and subsequent work (see for instance Benincà and Poletto 2004, Frascarelli 2007, Brunetti 2009).

In the generative tradition, starting at least from Bresnan (1970), the structural representation of a clause is taken to consist of three structural levels: the lexical layer, the inflectional layer and the complementiser layer. However, while Bresnan (1970) proposed that each sentence is introduced by a single node C, Rizzi (1997) showed that the complementiser layer is more articulated, hosting among others, topicalised and focalised elements, interrogative elements and relative pronouns. In other words, the left periphery of the sentence is split into a series of functional projections, each corresponding to a single feature specification, which is overtly or abstractly expressed. Working in the same framework, Rizzi and Shlonsky (2007) proposed that syntactic movement to A'-positions is triggered by the satisfaction of a "criterial feature", defined as the requirement of a head encoding a feature that determines the interpretation of the category bearing it (wh-, topic, focus...). The movement of an element to the left periphery must be motivated by the satisfaction of a criterion and the structure of the complementiser layer is rich and articulate to host the different kinds of phrases moved there.

The C(omplementiser)-layer can have different functions; one of them is related to the articulation of the clause into topic and comment (focus).

At this point, topic is simply defined topic as a generally preposed element that expresses old information that is already available and salient in previous

discourse²; focus, on the other hand, introduces new information or a contrast with respect to previously introduced information³.

Many languages distinguish the form of topics from the one of foci. In Romance languages, for example, a common strategy to identify topics involves a resumptive clitic coreferential with the topic, a construction known as Clitic Left Dislocation (CLLD).

(4) Italian

Il tuo libro, l' ho letto.
 the your book it.CL have.1SG read.PRT
 'Your book, I read it.'

The object clitic *lo* is obligatory in the Italian example. This is not true for foci:

² In Chapter 2, I will show that not all topics necessarily express old information. Frascarelli and Hinterhölzl (2007) proposed that there are three dedicated topic projections in the C-domain, each one hosting a different type of topic: ShiftP hosting Aboutness-Shift Topic, ContrP hosting Contrastive Topic and FamP hosting Familiar Topic:

(i) (adapted from Frascarelli 2007)

[ForceP [ShiftP [FocP [ContrP [FamP [FinP [IP ...]]]]]]]

This hierarchy of topic projections in the C-system correlates with specific discourse functions assigned to each topic type. In particular, the Aboutness-Shift Topic introduces a new topic, while the Contrastive Topic induces alternatives to previously introduced topics.

³ Contrastive focus is not further discussed in this dissertation. As it will become clear in Chapters 4 and 5, the analysis of referential properties of subject pronouns requires them to be uttered with a 'plain intonation', while contrastive focus is generally associated with emphasis and a different intonation. I refer to Cruschina (2011) for an exhaustive analysis of contrastive focus.

14 The syntax of subject pronouns in heritage languages

(5) Italian

IL TUO LIBRO ho letto (, non il suo).
the your book have.1SG read.PRT not the his.
'Your book, I read (, not his).'

In Italian, the focus is normally expressed by preposing the focal element and assigning it special focal stress. No resumptive clitics appear in this structure.

The topic in (4) and the focus in (5) are located in dedicated functional projections, TopP and FocP, respectively headed by Top and Foc and belonging to the complementiser system. In Rizzi's approach, a constituent endowed with topic or focus features must end up in a Spec/head configuration with Top or Foc. In Rizzi's terms, topic and focus are criteria, and movement to the dedicated projections in the C-system is triggered by the satisfaction of one of these criteria. Constituents cannot move freely to the left periphery.

The topic-focus system is integrated in the force-finiteness system, which have dedicated projections as well:

(6) (adapted from Rizzi 1997)

[ForceP Force [TopP Top [FocP Foc [FinP Fin [IP ...]]]]]

In Rizzi's approach, while Force and Finiteness are present in all clausal structures, Topic and Focus are present in a structure only if needed⁴. If present, the topic-focus field is enclosed in between force and finiteness.

⁴This kind of articulation of the C-system has been widely exploited for the account of a number of phenomena in Romance varieties, including the distribution of subject clitics in northern Italo-Romance varieties (Poletto 2000), which will be discussed in Chapter 3. Some aspects of such analyses can be maintained, while others need to be reshaped in the light of the data from heritage varieties.

1.3.2. Towards a feature-based approach

This section presents a minimalist alternative to the cartography of discourse-related projections proposed by Rizzi (1997). Minimalism and cartography have a different take on the encoding of discourse in syntax: while cartography describes the complex set of discourse- and interpretation- related projections in the left periphery, minimalism focusses on the question of which discourse-related features are involved in the syntactic computation and how. Chomsky (2008) proposed that discourse factors are encoded in narrow syntax in the form of edge features, uninterpretable features that may be realised on phase heads and trigger movement of a constituent for discourse or interpretive purposes. For instance, C may optionally carry an edge feature and trigger movement of a constituent with discourse import.

In the Minimalist Programme (Chomsky 2001, 2008), lexical items enter the derivation with a series of functional features that are probed by the agreement features of T and V. These two heads encode the agreement features responsible for the valuation of the case features of the DP subject and object. In turn, these two DPs have interpretable ϕ -features that value the uninterpretable ϕ -features under T and V. Following Chomsky (2001), such uninterpretable features are inherited from C and v respectively; in other words, the agreement features in C and v percolate down to T and V. The feature valuation process is implemented through the structural operation of agree, by which a probe P searches for a suitable goal G.

Several studies (Miyagawa 2005, 2010, 2017; Aboh 2010; Jiménez-Fernández 2010) show that the minimalist proposal can be extended to discourse-related features too. Miyagawa (2005) proposes that certain movement operations can involve an agreement relation with a discourse feature in the C-system. What is particularly interesting for this approach, is the idea that the different types of agreement (ϕ -agreement or discourse-agreement, specifically) are available to all languages, in accordance with Chomsky's Uniformity Principle.

Miyagawa proposes that ϕ -agreement and δ -agreement (agreement with discourse-related features, such as topic and focus) are respectively associated with a ϕ -feature or a δ -feature (discourse feature) in the C-domain, as already proposed by

16 The syntax of subject pronouns in heritage languages

Carstens (2003). Such a feature can percolate down from C to T and is therefore matched with a feature on a constituent in the syntactic structure, such as the thematic subject. Jiménez-Fernández (2010) proposed that some languages employ both types of features.

For instance, in the case of Spanish subject-verb agreement, the EPP feature may work in conjunction with ϕ -features, triggering movement of the subject to Spec-TP (7a), or with δ -features, such as topic, triggering movement of a topic-marked constituent to Spec-TP (7b).

(7) Spanish (Jiménez-Fernández 2010: 34)

- a. Susana vendió la moto.
 Susana sell.PAST.3SG the motorbike
- b. La moto(,) la vendió Susana.
 the motorbike it.CL sell.PAST.3SG Susana
 ‘Susana sold the motorbike.’

Either ϕ - or δ - features on T can be combined with the EPP feature and attract the relevant constituent to Spec-TP. Discourse features are directly encoded in syntax and the lexicon of a language contains categories which are endowed with δ -features.

In this respect, Aboh (2010) proposes that information structure starts in the numeration in the form of discourse-related lexical items. The discourse features encoded in these lexical items display a very specific syntactic behaviour and are therefore comparable to other formal features like ϕ -features. Numeration is defined as a set of pairs (LI, i), where LI is a lexical item and i is its index, understood as the number of times that LI is selected (Chomsky 1995). The computational system maps the set of items in the numeration to a sound representation (PF) and to a meaning representation (LF). What is mapped to these representations needs to correspond to what was originally part of the numeration: no further items can be added in the course of the derivation, as this would violate the Inclusiveness Condition (Chomsky 1995). Therefore, Aboh proposes that the information structure of a linguistic expression is already pre-determined in the numeration by means of discourse features. This is

compatible with the Strong Uniformity Principle (Miyagawa, 2010), the notion that all languages are uniform and have a complete featural set. The two types of grammatical features taken into account by Miyagawa, φ - and δ - features, both play a role in syntactic operations.

1.3.3. *Beyond C: discourse features in v and D*

According to Chomsky (2008), both C and v are phase heads. Jiménez-Fernández and Spyropoulos (2013) show that the parallel holds also when it comes to feature inheritance and δ -features. Also in the case of the v phase, V can inherit φ - or δ -features from the v head. This is exemplified by Spanish small clauses; the different word order in (8) depends on discourse properties of the small clause.

(8) Spanish (Jiménez-Fernández and Spyropoulos 2013: 188)

a. Considero muy lista a Susana.
 consider.1SG very smart to Susana

b. Considero a Susana muy lista.
 consider.1SG to Susana very smart
 ‘I consider Susana very smart.’

In (8a) *muy lista* is a topic and therefore appears before the focus *a Susana*; in (8b) *a Susana* is a topic, while *muy lista* is a focus. The different word order inside the small clause suggests that the behaviour of v is parallel to that of C: the authors propose that, at both phasal levels, the phase complement (T and V) inherits δ -features from the phase head (C and v). The effects of this inheritance were shown for C-T in Section 1.3.2: δ -features inherited by T from C, trigger the movement of the topic constituent to Spec-TP; as far as v -V, the case of small clauses shows that δ -features inherited by V from v trigger the movement of the topic part of the small clause to Spec-VP.

18 The syntax of subject pronouns in heritage languages

In this study, discourse feature inheritance is not discussed in the terms of Miyagawa (2010) and Jiménez-Fernández (2010). However, the proposed analysis stems from the notion that discourse features have an effect on different functional heads. Miyagawa's approach is extended to subject pronouns, aiming to show that the referentially specific interpretation of a pronoun is linked to the presence of a discourse feature on D.

1.4. Research questions and structure of the dissertation

1.4.1. *Research questions*

Several theoretical questions will be addressed in this study.

Throughout the dissertation **the role of language contact** in the distribution of overt and null subject pronouns in heritage languages is discussed. It will be shown that heritage languages undergo feature-reassembly, which allows for different syntactic properties and distribution of subject pronouns. In such perspective, heritage languages are not seen as simplified or incomplete systems; contact is rather seen as a source of innovation that may even lead to a complexification of the heritage grammar.

A question that stems from the discussion of feature-reassembly regards the **internal structure of pronouns**. If variation in the distribution of subject pronouns can be accounted for at the featural level, concepts such as structural deficiency and classes of pronouns (Cardinaletti and Starke 1999) become unnecessary. It will be demonstrated that all types of pronouns have the same internal structure and their different properties can be ascribed to a difference in their featural composition.

This proposal is particularly relevant in the discussion on the **nature of subject clitics** in Venetan, which have been shown to have a different internal structure than subject pronouns (Rizzi 1986, Brandi and Cordin 1989, Benincà 1994, Poletto 2000). However, in an approach in which all types of pronoun have the same structures, subject clitics will be analysed as regular pronouns; the difference is not in

their internal structure, but in the fact that they sometimes lack a discourse-related feature that is otherwise realised in strong pronouns.

The discussion will also focus on how the proposed analysis can capture the **antecedent-selection properties** of different types of subject pronoun. It will be shown that the interpretation of pronouns depends on the presence or the absence of a discourse feature, labelled [uR(eferential)].

Finally, the question of how **extra-syntactic factors** are involved in the distribution of different types of pronouns will be addressed. It will be proposed that the Salience Structure Hypothesis, according to which referential ambiguity is solved at the pragmatic level by means of the salience value of potential antecedents, is established via Context Scanning. This hypothesis provides further support for the fact that the realisation of [uR] on pronouns is strictly connected to the distinction between salient and non-salient antecedents, as it is licensed on pronouns only when some update to the salience structure is required.

1.4.2. Chapter outline

In Chapters 2 and 3 of this dissertation, a theoretical approach to different types of subject pronouns and the role of discourse features in their distribution will be developed; in Chapters 4 and 5, an application of this model to the antecedent selection properties of different types of subject pronouns will be proposed.

Chapter 2 presents the approach to subject pronouns in null-subject languages adopted in this work. The hypothesis is that all subject pronouns have the same internal structure and differences in their interpretation depend on a discourse feature realised in the D-head. This feature is defined as [R] (referential): when subject pronouns encode this feature, they are overt and referentially specific enough to obviate or switch reference; when subject pronouns lack this feature, they refer to the most salient discourse antecedent and are normally not phonologically realised. The hypothesis will be tested on heritage southern Italo-Romance varieties and, partially, on Venetan.

20 The syntax of subject pronouns in heritage languages

In Chapter 3, the approach is extended to null-subject languages such as Venetan, which allows for overt pronouns even when the [R] feature is missing. This is the case of subject clitics. Most works on subject clitics claim that they are agreement markers rather than real pronouns; however, analysing subject clitics from the perspective of information structure, it will be shown that their behaviour is clearly pronominal. The hypothesis is that Venetan subject clitics can be analysed as pronouns that may or may not be referentially specific enough to obviate or switch reference. Specifically, when they encode [R], they behave as regular overt pronouns; when they lack [R], they are phonologically realised counterparts of null *pro*.

In Chapter 4, the Position of the Antecedent Hypothesis ('PAH'; Carminati 2002) is tested on the three types of Venetan subject pronouns.

(9) Position of the Antecedent Hypothesis (PAH; Carminati 2002)

The null pronoun prefers an antecedent which is in the Spec-IP position, while the overt pronoun prefers an antecedent which is not in the Spec-IP position.

The study involves three Venetan varieties: Italian Venetan, Argentinian Venetan and Brazilian Venetan. The results of the study show that the antecedent preferences of subject clitics in heritage Venetan varieties indeed match those of overt pronouns. Moreover, the predictions made by the PAH are captured by the model proposed in Chapters 2 and 3, supporting the idea that at least some discourse-related features display clear syntactic effects. Finally, the role of contact in shaping the distribution of the discourse feature [R] in different varieties of Venetan will be discussed; it will be shown that there is not a direct effect of the contact language, but rather a process of feature-reassembly independent of the contact language, as the one described by Lardiere (2008).

Chapter 5 focusses on the role of salience in determining antecedent selection by different types of pronouns. A study on antecedent selection in Venetan is presented, showing that the PAH is too strong and that the salience structure of a sentence is the main factor determining the correct interpretation of a pronoun at the discourse level. It will be shown that subject pronouns are sensitive to syntactic and

discourse factors alike: the interaction of these factors leads to precise conditions of antecedent selection. It will be concluded that, in contact situations, these conditions are not lost, but rearranged and may lead to a complexification of the system.

Chapter 6 concludes the dissertation, summarising the theoretical proposals made in the previous chapters on the structure and distribution of subject pronouns and on the role of contact; finally, this chapter outlines some suggestions for future research that emerged from the data collected for the present dissertation.

22 The syntax of subject pronouns in heritage languages

Chapter 2. Referentiality and the internal structure of subject pronouns

2.1. Introduction

This chapter presents an approach to the distribution of third-person null and overt subject pronouns in null subject languages. Cardinaletti and Starke's (1999) tripartite model of structural deficiency, as well as related approaches on pronoun classes (Déchaine and Wiltschko 2002, Van Koppen 2012, Weiß 2015), will be reconsidered. The classification of pronouns in terms of structural deficiency proposed by Cardinaletti and Starke gives rise to rigid classes of pronouns. This classification fails to capture the full variation in syntactic and interpretive properties displayed by different types of subject pronouns in the Italo-Romance varieties under investigation. A simplified model is put forward, involving only one structure and one feature [R(eferential)] optionally realised on D, that can capture the variation in the realisation of different subject pronouns in the languages under investigation. The approach adopted here allows for a simple way of deriving both null and overt subjects along the lines described in Holmberg (2005), Sheehan (2006) and Roberts (2009).

This chapter also touches upon the extent to which the study of subject realisation is relevant for the understanding of interface-related issues. In particular, the interface between syntax and discourse has been shown to be somehow problematic for bilingual speakers of a null-subject language: they do not show native-like competence regarding the discourse-pragmatic constraints regulating the use and distribution of overt subjects. The fundamental idea of the Interface Hypothesis developed by Sorace and Filiaci (2006) is that bilingual speakers tend to realise some pragmatically infelicitous overt subject pronouns, in contexts that would require a null subject in monolingual speakers of the same language. The study presented in this

24 The syntax of subject pronouns in heritage languages

chapter involves heritage Italo-Romance languages spoken in different countries and shows that the distribution of null and overt subjects can be easily mastered by bilingual speakers. This results from the fact that referentiality is a universal property of human languages and is encoded in a [R(eferential)] feature assigned to a Logophoric Centre in CP, as proposed by Bianchi (2003). The division of labour between syntax and discourse gives rise to different featural combinations that in the end lead to the realisation of null or overt subjects and heritage (bilingual) speakers are as good as monolinguals in understanding and assigning different discourse values to referring expressions. Processes observed in heritage languages do not depend on a generalised tendency to simplify interface conditions, but on language specific combinations of features that emerge in different varieties of the same language; the number of the featural combinations that each variety grammaticalises in pronouns may hence decrease or increase; the different featural combinations available trigger a difference in the distribution of different types of pronouns.

The analysis presented here focusses on third person subject pronouns. First and second person are excluded from the analysis because they are not ambiguous with respect to referential properties: they are part of the Common Ground and they are always available to the speaker and the hearer being therefore referentially unambiguous (Erteschik-Shir 2007). Since the main issues addressed in this dissertation regard precisely cases of referential ambiguity, first and second persons are not taken into account.

The chapter is organised as follows: Section 2.2 presents an overview of the study of null subject and null-subject languages; Section 2.3 includes reviews of previous approaches on the internal composition of pronouns; moving from Cardinaletti and Starke (1999) and Déchaine and Wiltschko (2002), the approach to discourse features is presented and the [R] feature is introduced and discussed. In Section 2.4, some relevant interface and pragmatic theories on accessibility and salience structure of a sentence are discussed; the aim of this overview is not to answer the question of how the link between syntax and discourse is established: the aim is rather to propose a model for antecedent selection that can accommodate the variation displayed by heritage varieties. In Section 2.5 the derivation of null and overt subjects

is considered; the version of derivation of null and overt subjects adopted follows the lines of Holmberg (2005). In Section 2.6 the study on heritage Italo-Romance varieties is presented, with a focus on the fact that the interface conditions on the realisation of overt and null subjects are intact in these languages; the advantages of this newly-developed approach are presented. Section 2.7 concludes the chapter.

2.2. Null and overt subjects: an overview

Consistent null-subject languages are defined as languages in which a full referential subject can remain unexpressed. This possibility was formalised by linguists in the '70's (Perlmutter, 1971) and the '80s (Taraldsen, 1980). However, it was Rizzi (1982) who first attributed all these properties to a single cause, the null subject parameter. The parameter was formulated as following:

(1) The null-subject parameter (Rizzi 1982: 143)

- a. INFL can be specified as [+pronoun]
- b. INFL which is [+pronoun] can be referential.

Only a pronominal INFL can license a null subject.

(1) states that null subjects are licensed in null-subject languages because inflection in these varieties has a pronominal status.

There are four properties that consistent null-subject languages such as Italian show simultaneously: coreferentiality (2), free inversion (3), *that-t* effect (4) and rich agreement on the verb (5):

(2) Italian

Mia figlia_j non parla bene in dialetto. Però **pro_j** capisce tutto.
 my daughter not speak.3SG well in dialect but pro understand.3SG all
 'My daughter does not speak the dialect well. But she understands everything.'

26 The syntax of subject pronouns in heritage languages

(3)

Suona **il telefono**.
ring.3SG the telephone
'The telephone rings.'

(4)

La ragazza_j che *pro_k* hai detto che *pro_j* è brava in matematica.
The girl that *pro*.2SG have.2SG said that *pro*.3SG is.3SG good in math
'The girl that you said is good at math.'

(5)

Parl- o	'I speak'
Parl- i	'You speak'
Parl- a	'He/She speaks'
Parl- iamo	'We speak'
Parl- ate	'You speak'
Parl- ano	'They speak'

Properties in (2-5) will be discussed again in the study on heritage Italo-Romance languages presented in Section 2.6. It is clear, from these examples, that Italian has all the properties of consistent null subject languages in the traditional sense: in (2), the null subject is interpreted as being co-referential with the subject of the preceding clause; in (3), the subject is in a postverbal position; in (4), the subject is extracted from an embedded clause headed by an overt complementiser; the paradigm of the present tense of the verb in Italian in (5) shows rich agreement on the verb for all persons.

Aside from null-subject languages (languages that allow for a null subject) and non-null-subject languages (languages that do not allow for a null subject), there are radical null-subject languages (languages that can leave the subject unexpressed

even without displaying rich agreement) and partial null-subject languages (languages in which null subjects are restricted to specific structures or persons).

Holmberg (2005) defines consistent null-subject languages as those systems featuring null definite subject pronouns and partial null-subject languages as those systems not allowing null definite pronouns. Brazilian Calabrian (6) is a consistent null-subject language, while Brazilian Portuguese (7) is a partial null-subject language that does not allow null definite third person subjects:

(6) Brazilian Calabrian

pro tenə tre annə.
pro have.3SG three years
'He is three years old.'

(7) Brazilian Portuguese (Holmberg 2005: 553)

*(**Ele**) ganhou na loto.
he won.3SG on.the lottery
'He won the lottery.'

According to Holmberg (2005, 2009), this difference depends on the fact that the finite verb in consistent null-subject languages encodes a D feature and therefore yields a definite interpretation; this is the case of *tenə* in Brazilian Calabrian in (6). Conversely, the finite verb in partial null-subject languages does not have a D feature, therefore a pronoun with its own valued D feature needs to be merged in order to obtain a definite interpretation; this is the case of *ganhou* in Brazilian Portuguese in (7). A sentence without an overt third person subject is possible in Brazilian Portuguese, but it gives rise to a generic interpretation:

(8) Brazilian Portuguese (Kato 1999: 5)

pro conserta sapatos aqui.
pro repair.3SG shoes here.
'One repairs shoes here.'

28 The syntax of subject pronouns in heritage languages

The verb *conserta* in (8) does not have a D feature, therefore the null subject yields an indefinite (generic) interpretation. What is implied here is that *pro* too does not carry D: this point will be discussed in detail in the next section. I come back to the pronominal (D) status of the verb in null subject languages in Section 5.

Languages that never allow for null subjects are defined as non-null-subject languages. One of them is French (9):

(9) French

*(I) boit du vin.
he drink.3SG of wine
'He drinks wine.'

According to Holmberg (2009) the verb *boit* in (9) not only lacks a D feature, but it also lacks ϕ features; therefore, an overt subject must be merged in all contexts.

The last possibility is represented by radical (or discourse) null-subject languages, such as Mandarin Chinese.

(10) Mandarin Chinese (Holmberg 2005: 558)

__ méi chī zǎofàn.
__ no eat breakfast
'(I/You/He) have not had breakfast.'

According to Huang (1984), the type of null element displayed by these languages belongs to a different category from the one displayed by consistent null subject languages.

All the heritage languages included in the present study are consistent null-subject languages. The discussion of contact-induced change refers to contact with null-subject, partial null-subject and non-null subject languages; however, the distribution of null and overt subjects presented in this chapter refers mainly to consistent null-subject languages.

2.3. The internal structure of pronouns

The analysis of the structure of subject pronouns adopted here focusses on the different internal structure of null and overt pronoun and stems from Cardinaletti and Starke (1999) and Déchaine and Wiltschko (2002); like them, the present study provides an analysis of the internal structure of pronouns in terms of featural composition; unlike them, this study does not adopt a model that resorts to structural deficiency for some pronominal subjects. It is argued for the existence of an [R] feature that can capture the distribution of third-person strong overt pronouns and null *pro*. The variation displayed by heritage languages can be better explained if we analyse subject distribution as dependent on the internal featural composition of the DP, without resorting to structural deficiency.

Let us start by summarising how the distribution of third-person overt and null subjects in consistent null-subject languages is accounted for in Cardinaletti and Starke's (1999) model. In this model, personal pronouns are divided into three different classes according to their level of structural deficiency: strong pronouns (11)a), weak pronouns (11)b) and clitic pronouns (11)c).

(11) Italian

- a. Dirò tutto **a lei**.
 tell.1SG everything to her
 'I will tell everything to her.'
- b. Dirò **loro** tutto.
 tell.1SG them everything
 'I will tell them everything.'
- c. **Le** dirò tutto.
 tell.1SG her everything
 'I will tell her everything.'

30 The syntax of subject pronouns in heritage languages

Cardinaletti and Starke (1999) propose that the weak pronoun in (11)b) is structurally deficient with respect to the strong pronoun in (11)a) in that it lacks a set of morphological, syntactic, semantic and prosodic properties that the strong pronoun has; the clitic pronoun in (11)c) is structurally deficient with respect to the weak pronoun (11)b) in and the strong pronoun (11)a) in lacks a set of properties that both the weak pronoun and the strong pronoun have. In the remainder of this section a relevant difference between strong and weak pronouns is introduced to explain what structural deficiency means in practice.

According to Cardinaletti and Starke (1999) overt subject pronouns may be strong or weak, while *pro* belongs to the weak class. This implies that it should be possible to highlight some asymmetries with regard to the syntactic distribution of pronouns belonging to different classes. The trigger of such asymmetries is structural deficiency: the more deficient a pronoun is, the less features (projections in Cardinaletti and Starke's terms) it encodes. Weak pronouns are structurally deficient with respect to strong pronouns and this determines a restriction in their distribution, as shown below. Cardinaletti and Starke show that *pro* has the semantic properties of a weak pronoun, in that it normally must refer to an entity that is salient in the discourse. Weak pronouns can be overt as well; we will return to overt weak pronouns in Chapter 3.

The reasons why there can be a null pronoun (*pro*) as a sentential subject are discussed in Section 2.5. Following Sheehan (2006) and Roberts (2009), *pro* is defined as a phonologically null counterpart of a weak overt pronoun: *pro* is null because of a principle of economy that deletes copies of identical feature bundles at PF.

One relevant property that is encoded in strong pronouns and lacked in null subjects is the referential index. While strong pronouns (12)a) can have their own range of referents (hence, they can be referentially specific), *pro* (12)b) cannot, therefore it always needs to be associated with a salient antecedent in the previous discourse in order to be able to refer (the notion of salience and its relevance for the model will be discussed in Section 2.4).

(12) Argentinian Calabrian

a. **Iddri** anu truvetə a soluzionə.
 they have.3PL found the solution.

b. **pro** anu truvetə a soluzionə.
 pro have.3PL found the solution.

‘They found the solution.’

If no salient antecedent is found in previous discourse, the interpretation of *pro* in (12)b) remains ambiguous and can lead to a generic interpretation. This interpretive property of referential null subjects is formally captured by a different model, developed in Déchaine and Wiltschko (2002). Déchaine and Wiltschko propose that pronouns can be of three types: DP, φ P or NP. Similarly to what is claimed by Cardinaletti and Starke, in Déchaine and Wiltschko’s model, each pronoun type is associated with a syntactic projection, as shown in (13). With respect to their semantic and referring properties, DP-pronouns function like regular definite referring expressions, while φ P-pronouns lack inherent semantics. NPs are constants with interpretive properties that depend on their inherent semantics and syntactic properties of nouns. NPs are left aside for the moment, as they are not relevant for the present study (see Déchaine and Wiltschko 2002 for an analysis of Japanese *kare* as an NP-pronoun).

(13)

a. [**DP** [φ P [NP] (DP-pronouns)

b. [φ **P** [NP] (φ P-pronouns)

c. [**NP**] (NP-pronouns)

Both Déchaine and Wiltschko (2002) and Cardinaletti and Starke (1999) come to the same conclusion, namely that strong pronouns are referential DPs; the referential properties of strong pronouns are encoded in the DP layer; weak pronouns, conversely, encode only φ features, while they lack the DP layer and the referential properties associated with it. Hence, in this line of analysis, the lack of referential

32 The syntax of subject pronouns in heritage languages

properties of *pro* is syntactically represented as the absence of a portion of structure: the DP layer.

(14)

a. [DP *she* [φ P [NP]

b. [φ P *pro* [NP]

Déchaine and Wiltschko show that this structural difference explains reference tracking properties of strong pronouns such as obviation and switch reference. In most studies on Romance pronouns, obviation is intended as a discourse-sensitive mechanism that distinguishes third person participants from each other, while switch reference takes place when the subject of a dependent clause is distinct from the subject of the main clause. According to Déchaine and Wiltschko, both properties imply the reference to a different subject with respect to the one previously introduced in the discourse. Strong pronouns are definite referential expressions and, as such, are used as obviatives and markers of reference to a different subject; *pro*, as a weak pronoun is a variable and, as such, it tracks a default reference to the proximate subject (intended as the same subject as in previous discourse).

2.3.1. Introducing the [R] feature

In the remainder of this chapter, it will be shown that the different interpretation of null and overt subjects can be captured through a single primitive, in the form of a discourse feature encoded on the D head. I call this feature [R(eferential)]. To start with, discourse features are here conceived in Miyagawa's (2005) sense, as features encoding discourse-related properties that are involved in syntactic operations on par with φ features, as already discussed in Chapter 1. Aboh (2010) claims that information structure starts in the numeration in the form of discourse-related features that trigger a specific syntactic behaviour. He argues that this follows from the notion that what is mapped to PF and LF needs to correspond to what was originally part of the numeration: no further items can be added in the course of the derivation, as this

would violate the Inclusiveness Condition (Chomsky 1995). In this study, I will assume that [R] starts in the numeration as a property of some lexical items, such as referentially specific DPs.

As far as overt subject pronouns are concerned, Longobardi (1994) showed that they occupy the D head of a DP⁵; It is proposed that this is true for all overt and null subject pronouns. Strong pronouns (15)a) have their own referential range, hence are not dependent on a salient discourse antecedent and encode an [R] feature in D. In the case of null subjects (15)b) (weak pronominal elements, according to Cardinaletti and Starke, 1999), the D head lacks the [R] feature.

(15)

a. [DP_[R] she [φ P [NP]]]

b. [DP *pro* [φ P [NP]]]

In both cases, the D head of the DP encodes a [D] feature that expresses the semantic notion of definiteness. According to Lyons (1999), the main property of definiteness is identifiability, intended as the ability of the interlocutors to locate a possible referent for the pronoun. While definiteness is considered here to be a property of all subject pronouns, [R] is an optional feature that specifies what referent is being referred to by the pronoun.

As will be shown in Sections 2.4 and 2.5, the model presented in this section explains how syntax establishes that the interpretation of *pro* is dependent on a salient antecedent in previous discourse, while an overt pronoun can switch the reference to a less salient discourse antecedent. Unlike what is claimed in Cardinaletti and Starke (1999) and Déchaine and Wiltschko's (2002), I maintain that all pronominal forms have the same syntactic structure, consisting of three layers: D, φ and N. The difference between overt (strong) and null (weak) pronouns is not in their structure, but in their featural composition; weak pronouns lack [R], the discourse feature

⁵ The idea that pronouns are determiners goes back to Postal (1969). See also Alexiadou et al. (2007) in this respect.

34 The syntax of subject pronouns in heritage languages

encoding the referential properties; strong pronouns encode [R], which is responsible for phenomena such as topic shift or switch reference.

This approach does not contradict Holmberg's (2005) analysis of null subjects as ϕ Ps encoding only ϕ -features; this analysis will be discussed in Section 2.5; informally, the featural composition of ϕ Ps allows only for a default interpretation, making them obligatorily coreferential with the most salient antecedent. The default interpretation results precisely from the lack of [R], the feature that makes overt pronouns referential enough to switch the reference to a less salient antecedent. In other words, the presence of [R] allows reference to a specific discourse antecedent, crucially, not the most salient one. The presented approach supports the view proposed in Manzini (2014) and Pescarini (2018) that the rigid organisation of pronouns into classes claimed by Cardinaletti and Starke (1999) is not necessary; referential weakness and deficiency are captured by the distribution of [R] in pronouns.

A question that is not addressed in Cardinaletti and Starke (1999) and Déchaine and Wiltschko (2002) regards the way syntax establishes that the interpretation of *pro* is dependent on a salient antecedent in previous discourse, while an overt pronoun can switch the reference to a less salient discourse antecedent. In the remainder of this chapter, it will be shown that the interpretation of *pro* and overt subject pronouns can be expressed by means of [R].

2.4. The interpretation of null and overt subjects: topic, linking and logophoricity

Several approaches to information structure and, more specifically, to the realisation of null and overt subjects in null-subject languages claim that the interaction between different domains of language takes place in CP, the locus in which syntactic and pragmatics or discourse information is exchanged (Platzack 2001). Subject pronouns agree for ϕ -features with an element in CP (a null topic in a dedicated ShiftP according to Frascarelli 2007; a context linker in Sigurðsson 2011). In Frascarelli's approach, the null subject also agrees for a discourse feature [+aboutness] with a null topic in

ShiftP. This implies that discourse features have an effect on syntax and participate in the same type of dependencies as ϕ -features, as shown also by Miyagawa (2005, 2010). Frascarelli's analysis postulates the presence of a discourse feature determining the null realisation of the subject in Italian; as far as overt pronouns are concerned, Frascarelli claims that they are inserted to obviate coreference with respect to the current Aboutness Shift Topic; in other words, they can act as a new Aboutness Shift Topic. I discuss Frascarelli's approach in Section 2.4.1.

A similar analysis is presented in Sigurðsson (2011, 2013)⁶: his definition of Context Linking is however more conceptual and strictly linked to the idea of context scanning and logophoric participants, as I discuss in Section 2.4.2. Logophoricity is also a fundamental component in Bianchi (2003): she adopts the idea of a Logophoric Centre functioning as the centre of deixis, hence defining pragmatic conditions that regulate the reference to entities in the discourse, as I discuss in Section 2.4.3.

2.4.1. Frascarelli (2007): The A-topic and the Topic Criterion

In Frascarelli (2007), null subjects are licensed through an Agree relation: an Aboutness-shift topic ('A-topic') in the C-domain agrees in person and number with the finite verb. The A-topic is merged in the left periphery and is endowed with a [+aboutness] edge feature. The author further proposes that null subjects are identified through a Topic Criterion:

(16) Topic Criterion (Frascarelli 2007: 721)

- a. The high topic field in the C-domain contains a position in which the [+aboutness] feature is matched with the local third person null subject.
- b. When continuous (i.e. not changed from the previous sentence), the [+aboutness] topic can be null (i.e. silent).

⁶ See also Sundaresan (2013) for a similar approach on the role of perspective in anaphoric interpretation in Tamil.

36 The syntax of subject pronouns in heritage languages

Therefore, the null topic is the antecedent for the null subject. A topic can only remain silent when it is part of a topic-chain: the first topic must be overt, while its copies in the following sentences can (but do not need to) remain silent. The A-topic position, however, is always syntactically represented in the C-domain, either overtly or covertly.

This is shown in (17): the antecedent of a null subject is a null A-topic located in the CP of the clause that immediately contains the null subject (17)c). The null A-topic is a copy of another A-topic (which can be null or overt) in the locally preceding discourse (17)b). Therefore, the reference of the null subject comes from a spelled-out DP in the preceding discourse, via a chain of A-topics.

(17) Argentinian Calabrian

- a. **A wuagliona mia_j** un parla tantu. Però **pro_j** capisce tuttu.
the daughter mine not speak.3SG much but pro understand.3SG all
'My daughter does not speak a lot, but she understands everything.'
- b. [_{ShiftP} **A wuagliona** [_{TP} [_T parla...]]]
- c. [_{ShiftP} < **A wuagliona_k** > [_{TP} **pro_k** [_T capisce...]]]

In (17)b) there is an overt newly introduced A-topic in ShiftP (the position dedicated to A-topics in Frascarelli 2007). In (17)c), which immediately follows (17)b), the A-topic is maintained and is, therefore, silent. Similarly, Frascarelli claims that overt pronouns are inserted to obviate coreference with respect to the current Aboutness Shift Topic. They can act as new Aboutness Shift Topic and, as such, they can appear in the left periphery (ShiftP in Frascarelli's terms).

(18) Argentinian Sicilian

- a. **Iddu_j** vinni n'anno prima. Poi **pro_j** mannau a chiamare a famigghia.
he came.3SG an year earlier then pro sent.3SG to call the family
'He had come one year earlier. Then he called the family.'
- b. [_{ShiftP} **Iddu** [_{TP} [_T vinni...]]]
- c. [_{ShiftP} < **Iddu_k** > [_{TP} **pro_k** [_T mannau...]]]

In Frascarelli's system, a strong subject pronoun can act as a topic; however, this system is problematic in view of the analysis of *pro* as a weak pronoun (Cardinaletti and Starke, 1999; Déchaine and Wiltschko, 2002): the idea that a discourse feature is assigned to *pro* is somehow at odds with its structural deficiency and with its "default" interpretive properties.

Unlike Frascarelli, I define the "default" interpretation of *pro* as the property of being always dependent on a salient antecedent; in this respect, *pro* can be intended as the default choice in null-subject languages. This idea is implicit in the original definition of the Avoid Pronoun Principle (Chomsky 1981), according to which overt pronouns should be avoided whenever possible. Regardless of the adopted analysis on its internal structure, null *pro* in consistent null-subject languages can encode only ϕ features, while information-structural properties are realised on (strong) overt pronouns. If we assume that *pro* is licensed under specific discourse conditions, then the difference between consistent and radical null subject languages cannot be maintained. Recall that, according to Huang (1984) and Holmberg (2005), the null category licensed under discourse conditions in Mandarin Chinese, for instance is not the same as *pro* in consistent null-subject languages. Based on this information, I propose that in consistent null-subject languages it is the overt pronoun that needs to be licensed, not the opposite. Besides, as it will be shown in Chapter 4 and 5, the fact that strong pronouns in null-subject languages are necessarily used to obviate coreference is not as clear-cut as Frascarelli's model predicts. The situation is very nuanced and can be defined in terms of preference (Carminati 2002): the *preferred* interpretation of strong subjects is indeed that involved in obviation (and switch reference), but this does not exclude that they can be allowed also when these conditions do not apply. This becomes particularly evident in contact and heritage varieties, as illustrated in (19).

(19) Argentinian Sicilian

Me patri_j era rispettatu. **Iddu**_j avea n'ato muodo.
my father was.3SG respected he had.3SG another way.

38 The syntax of subject pronouns in heritage languages

“My father was respected. He had a different behaviour.”

I leave this discussion for following chapters, in which it will be shown that the approach proposed in this dissertation can capture the idea of preference resorting again to the internal structure of the pronoun and to [R].

In sum, Frascarelli’s (2007) model correctly predicts the distribution of null subjects, but it is not clear about the distribution of overt pronouns. Besides, the assignment of an [aboutness] feature to the null subject contradicts the idea that *pro* in consistent null-subject languages is a weak pronoun and, as such, lacks discourse properties. The approach adopted here builds on Frascarelli’s model, in a slightly different perspective: what requires licensing is not *pro*, but the overt subject, precisely because of the additional discourse-related information it carries.

2.4.2. Sigurðsson (2011, 2014): Context Scanning

The fundamental concept in Sigurðsson’s (2011, 2014) Context-linked Grammar is that, at some level of cognition, grammar is context-sensitive and planned (Sigurðsson, 2014: 176): while narrow syntax (including the operation Merge) is context-blind, context-sensitive grammar makes sure that context-dependent items (indexical or deictic items, like subject pronouns) and categories fit their context. Sigurðsson maintains that all clauses contain context-related categories such as the logophoric participants (the speaker - ΛA , the addressee - ΛP , third-person participants - topic), speech time (ST) and speech location (SL). These categories are represented by speech event features, defined as C/edge features, that reflect or relate to discourse properties and are found in a split clausal head C:

(20) (adapted from Sigurðsson 2014)

$C \supseteq \{\text{Top}, \Lambda A, \Lambda P, \text{ST}, \text{SL}, \dots\}$

According to Sigurðsson (2014), *C*/edge features are silent, but they have overt clause-internal effects. They are syntactic heads acting as separate probes, but they may bundle up.

Unlike Frascarelli (2007), Sigurðsson (2011) believes that overt arguments need to match a *C*/edge feature in their local *C*-domain in order to be interpreted. Any finite *C*-domain has its own set of *C*/edge linkers (CL_n), a set of features that can be independently valued or valued in relation to a preceding category. This is shown for example in (21).

(21) Brazilian Calabrian

a. **Iddru** ha lavuratu assai.

he have.3SG worked much

‘He worked a lot.’

b. [CP...{ **CL_n**]_i...[TP...**Iddru**_i...

The pronoun *iddru* in Spec-TP (21) agrees with its local CL_n in CP. Crucially, in this system, the only type of agreement allowed in syntax is ϕ -agreement. The reference of the CL_n is decided by Context Scanning, an operation that can happen under long-distance agreement (in a subordinate clause, for example) or by extrasyntactic means. CL_n matching and context scanning yield context linking:

(22) (adapted from Sigurðsson 2011)

[CP [Force [CL_n [... [TP... [phon]/ \emptyset ...
Context scanning ↗ ↘ *C/edge linking* ↗

Sigurðsson’s approach represents a formalisation of the assumption that referential arguments link to their linguistic or deictic context via their *C*-domain. However, while the ϕ -computation of the CL_n values is completed in syntax, referential properties are established through Context Scanning, a conceptual operation taking place at the discourse level. Sigurðsson (2011, 2014) does not further discuss how the scanning takes place and, in general, the way discourse and context properties are

40 The syntax of subject pronouns in heritage languages

encoded in syntax is still rather poorly understood. This study adopts Sigurðsson's view of Context Scanning, an interface operation that provides syntax with the necessary discourse-related information in order to license an overt subject. Sigurðsson proposes that a pragmatic feature is responsible for the linking of all arguments to the context or the discourse. A similar proposal in this respect was already presented in Bianchi (2003), Speas (2004) and D'Alessandro (2004, 2007).

2.4.3. The Logophoric Centre and the [R] feature: accessibility and salience

The idea of logophoricity employed by Sigurðsson (2011, 2014) is defined as a relationship between a referring expression and a referent, an entity whose speech, thoughts or perspective are reported. Bianchi (2003) proposed that speech event, the deictic centre of a sentence, can be defined precisely as a Logophoric Centre. Every clause is anchored to a Logophoric Centre, a speech or mental event that includes discourse participants and constitutes the centre of deixis, relating directly to the context or discourse setting in which the sentence is uttered. The idea of a Logophoric Centre is exploited in Bianchi (2003) to account for different types of control clauses. Control is always mediated by a Logophoric Centre syntactically represented in FinP. Bianchi distinguishes obligatory and non-obligatory control, in that the first is anaphoric, while the second is pronominal. In other words, in obligatory control, the controller needs to be an argument of the immediately previous discourse (generally the previous sentence), while in non-obligatory control, it needs not. Bianchi proposes that in the case of non-obligatory control, the Logophoric Centre licenses a referentially independent DP carrying an [R] feature⁷.

In this study, I propose that a similar mechanism is involved in the distribution of null and overt subject pronouns in null-subject languages. More specifically, depending on the interpretation, the Logophoric Centre may or may not be assigned an [R] feature. Recall that another occurrence of [R] is realised on strong

⁷ See also Landau (2004) on the role of [R] in non-obligatory control contexts.

pronouns and triggers a referentially specific reading of the type involved in switch reference and topic shift. If D lacks [R] the referring expression will be weak and, hence, corefer with the most salient discourse antecedent. Before defining the type of relationship that is established between the Logophoric Centre in CP and the pronoun in Spec-TP, I will give a short overview on salience as intended in pragmatic and discourse studies.

2.4.3.1. Salience and accessibility

Salience is a commonly used concept in linguistics. Boswijk and Coler (2020), in their overview of the different uses of the term in linguistics, define salience as a property that makes some information stand out in the discourse; in other words, salient information is prominent in the discourse, therefore easily recalled and accessible by all interlocutors. The concept of salience was proposed in Ariel (1990) as one of the fundamental ingredients of Accessibility Theory. This theory deals with the use of context in utterance interpretation; among other phenomena, it develops an account of intra- and inter- language differences in the occurrence of null thematic subjects. This account stems from the idea that the distribution of null and overt subjects inside the same language and across languages depends on the accessibility of their referents in the context and it requires a refined distinction between different types of elements and features encoded in them. Accessibility Theory is based on the idea that speakers choose between different referring expressions to mark accessibility differences; different referring expressions mark different degrees of accessibility.

The degrees of accessibility depend on different factors; here the focus is on their salience in the discourse, intended as their topical or non-topical role, as discussed in previous sections. Mental representations are not equally salient (so not equally accessible) to the participants in the discourse at different stages of the discourse and different context-retrieving expressions point to different levels of salience in discourse: a topical constituent is more salient and more accessible than a non-topical one. Another relevant aspect for the definition of different degrees of accessibility is the depth of storage in memory, which is strictly dependent on the

42 The syntax of subject pronouns in heritage languages

notion of givenness (Chafe 1976): speakers use some referring expressions to refer to “old” or given information, evoking representations that were already introduced in the discourse; other referring expressions evoke new representations that are not referred to previously in the current discourse. The preference for a referring expression in relation to givenness depends on the fact that, once evoked, concepts remain active in the memory for a while; the distance between the last mention and the current mention is crucial, as later mentions are more accessible than earlier ones. When the representation is not highly accessible, speakers favour ‘bigger’ linguistic expressions: these are the markers of low accessibility. On the other hand, when the representation is highly accessible in the memory, the ‘smallest’ linguistic expressions available in the system are preferred: the markers of high accessibility.

(23) Accessibility hierarchy (adapted from Ariel 1990)

Accessibility	·	- Full name
	↑	- Demonstrative
	↑	- Stressed pronoun
	↑	- Unstressed pronoun
	+	- Null

Accessibility marking depends crucially on the ease of retrieving the intended antecedent. In particular, high accessibility markers are used when the concept to be evoked is very salient and recent: high accessibility markers are preferred in subsequent mentions of the topic of current discourse, being the unmarked forms in discourse. Ariel suggests that the number of markers may vary across languages and that the same marker does not necessarily refer to the same level of accessibility across languages.

Ariel elaborates a precise coding pattern for accessibility. The criteria used in coding the “degree of accessibility” are informativity (the emptier the marker is semantically, the higher accessibility it signals), rigidity (the less uniquely referring an expression is, the higher accessibility it signals) and attenuation (the smaller an

expression is phonologically, the higher accessibility it signals). Rigidity and informativity seem to be particularly relevant in the definition of [R]: the higher informativity and rigidity of a strong pronoun with respect to a weak pronoun follow precisely from the presence of [R] on the former. These criteria therefore define the distribution of different pronominal forms in each language. In the case of null subjects in consistent null-subject languages, it is clear that their choice follows from the idea that a null element should be always preferred, whenever possible; the use of overt subjects depends more specifically on the need to obviate or switch reference, as suggested also in Frascarelli (2007).

The assignment of [R] is obtained via a scanning operation, defined along the lines of Sigurðsson (2011, 2014). Such scanning operation identifies the potential antecedents in previous discourse and organises them in a set. At the discourse level, each sentence has a salience structure that consists of a simply ordered set of ranked discourse referents carrying different salience values; some referring expressions (strong pronouns, in the case of the present study) can update the salience structure and change the salience value of discourse referents; the main factors that determine the updating process are the descriptive content of the referring expression, the distance from the antecedent, as well as subjecthood and topicality.

Building on Von Stechow's (2000) dynamic-semantic model, in the next section I propose a formal definition of salience as a context-dependent choice function that takes a set of potential antecedents and yields one of such elements. The interpretation of pronouns crucially depends on such function, which reflects the salience structure of a discourse.

2.4.3.2. Reference and context in dynamic semantics

In dynamic semantics, discourse is understood as a string of sentences realised one after the other. Each sentence represents an instruction to update the context with some new information; in other words, every sentence has a context change potential. The notion of context is fundamental in dynamic semantics, as it is that component of discourse that influences the interpretation of new sentences with respect to previous

44 The syntax of subject pronouns in heritage languages

ones. This is shown in Chierchia (1992: 132): if we utter a sentence S_1 and then we continue with a sentence S_2 , the meaning of the two sentences is obtained by composing the context change potential of S_1 with that of S_2 ; the context change potential of S_2 represents a possible continuation of the context change potential of S_1 . I refer to the abundant literature on the topic for the details of this and related proposals (see in particular: Chierchia 1992; Groenendijk and Stokhof 1991a, 1991b; Heim 1982, 1990; Kamp 1981).

Von Heusinger (2000, 2002) discusses the dynamic-semantic model with respect to accessibility of the antecedents of pronouns and indefinite DPs. He proposes that there exists a set of potential antecedents of referring expressions.

Here, this set will be called Δ and the potential antecedents will be called d_x , where x is a number that identifies the salience value of the antecedent, starting from d_1 (the most salient among potential antecedents retrievable from context). A referring expression such as a pronoun refers to an element d_x of an ordered set $\Delta = \{d_1 \dots d_n\}$ of possible discourse antecedents associated with the same descriptive content but associated with different salience value: specifically, d_1 is associated with the highest salience value (it is the most salient potential antecedent), and the salience value gradually decreases in all the following elements d_{x+1} of the set.

It is proposed that the set Δ of possible discourse antecedents is established by Context Scanning, along the lines of Sigurðsson (2011). At this point, the Logophoric Centre in CP comes into play, mediating the relationship between this set and the pronominal form in Spec-TP.

At the discourse level, the Logophoric Centre will encode the selection of an element from the set of possible antecedents. The selection of the first element d_1 in the set implies that the Logophoric Centre does not license a referentially specific interpretation and the subject will corefer with the most salient discourse antecedent.

However, as proposed in Von Heusinger (2000), the salience value can change in the flow of discourse because of different factors, such as the syntactic construction or the situational knowledge. In this case, an element d_{x+1} of the set is selected, which implies an update of the salience structure of the sentence: the Logophoric Centre licenses a referentially specific interpretation and the subject will

refer to a less salient discourse antecedent. This notion stems from the idea that some expressions (some types of pronouns, in the case under analysis) have a salience-change potential that changes or updates the salience structure. This update can be thought as a choice function that defines the rearrangements of the elements of ordered set \mathcal{A} into an ordered set \mathcal{A}' . This “update” function takes a choice function and a set \mathcal{A}' and yields a new choice function. This model predicts that if an element $d_{\neq i}$ of the set is selected, the salience value of this element needs to be updated, leading to an update in the salience structure; in this case, the Logophoric Centre licenses a referentially specific interpretation on the subject, in that reference has to be switched. I refer to Von Heusinger (2000) for the details of this proposal. Possible developments and further applications of the model sketched above are a matter of future research.

2.4.4. The assignment of [R]

At this point, it is necessary to clarify how [R] is encoded in syntax. If the antecedent selected in discourse corresponds to the element d_i in the set \mathcal{A} , the Logophoric Centre will not encode any extra [R] feature; if the selected antecedent corresponds to an element $d_{\neq i}$ in the set \mathcal{A} , the Logophoric Centre will encode [R]. A possible way to explain the encoding of [R] is given in Adger and Ramchand (2005). They propose that a syntactic feature needs to be interpreted at the LF interface as a predicate abstraction (in Heim and Kratzer’s 1998 sense: a feature whose semantic value is defined as a function); I take this feature to correspond to [R] on the Logophoric Centre; at the same time, there is another occurrence of this [R] feature on the pronominal form that is interpreted with respect to [R] on the Logophoric Centre. In practice, this model captures the fact that pronouns are always referentially dependent on an assignment function. Building on Adger and Ramchand (2005), I propose that [R] is a feature that establishes a correspondence between interpretation and syntactic realisation of pronouns. In sum, [R] on the Logophoric Centre is interpreted as a predicate abstraction, while [R] on the pronoun is syntactically dependent on [R] on the Logophoric Centre.

46 The syntax of subject pronouns in heritage languages

Let us now see how the assignment of [R] to the pronoun works in syntactic terms. Recall that [R] is assigned to the Logophoric Centre as the result of a function responsible for the assignment of an interpretation (that of referential specificity, as discussed in Section 2.4.3.1). [R] will be assigned to the Logophoric Centre only if this particular interpretation is invoked. The [R] feature on the Logophoric Centre checks an [R] feature on the subject DP. I define this syntactic relationship as [R]-checking:

(24) [R]-checking

An uninterpretable feature [uR] on a syntactic object Y must be checked against an interpretable [R] feature on syntactic object Z c-commanding Y.

The pronoun carries [uR], while the Logophoric Centre carries [R]. Being uninterpretable, [uR] needs to be deleted before transfer to LF⁸. The derivation follows a specific downward licensing configuration, defined by Wurmbrand (2010) as Reverse Agree. A similar notion is used by Zeijlstra (2012) in his definition of Agree as an operation that applies when an interpretable feature [F] c-command an uninterpretable feature [uF]. Zeijlstra's definition of Agree is presented in (25):

(25) (*Reverse*) Agree: α can Agree with β iff:

- a. α carries at least one uninterpretable feature and β carries a matching interpretable feature.
- b. β c-commands α .
- c. β is the closest goal to α .

⁸ In Chomsky's (1995) terms, deletion implies that the feature becomes invisible at LF, but is not erased from the computation. Erasure of a feature is intended as a stronger form of deletion, in which an element is eliminated entirely from the computation. This is not the case of [uR], which is deleted at LF but remains available for operations at PF.

In the case under analysis, α corresponds to the pronominal form in Spec-TP, carrying [uR] and β corresponds to the Logophoric Centre carrying [R]. The Logophoric Centre c-commands the pronominal form; the Logophoric Centre is the closest goal to the pronoun. [R] on the Logophoric Centre checks [uR] on the subject DP (which will be assigned the referential properties defined by Context Scanning) and the derivation can proceed.

If, however, Context Scanning does not assign [R] to the Logophoric Centre, there will be no feature to be checked on the subject DP: the ϕ features of a weak pronoun such as *pro* are sufficient to identify the antecedent and the derivation can proceed.

The interaction between discourse, information structure and syntax in the definition of the factors that lead to the realisation of null and overt subject pronouns is an aspect that is still poorly understood. In this section I presented a model that helps in better understanding this complex interface phenomenon. The model adopted in this study captures Frascarelli's (2007) account of an Agree relation between a null element in CP and the referring expression (the pronoun) in Spec-TP, showing however that such relationship is necessary for overt pronouns, rather than *pro*. Besides, it was proposed a slightly modified version of Sigurðsson's (2011) Context Scanning; in particular, it was shown that discourse-related features such as [R] participate in syntactic operations: [uR] on the subject DP needs to be checked against [R] on the Logophoric Centre in order to be correctly interpreted at the interface.

2.5. The derivation of null and overt subjects

The approach presented in Section 4 fits the analyses proposed in Holmberg (2005), Sheehan (2006) and Roberts (2009) on the pronominal nature of T. A generally accepted property of null-subject languages is the presence of a pronominal feature on finite T heads; this means that finite verbal morphology is pronominal and that a null subject entering an Agree relationship with such T head can be interpreted as definite. According to Holmberg (2005), the null subject is specified for interpretable ϕ features and moves to Spec-TP just like overt subjects. In Holmberg's view, *pro* is

48 The syntax of subject pronouns in heritage languages

simply a pronoun that is not overtly realised, a phonologically null counterpart of the overt pronouns.

Roberts (2009) proposes that subject pronouns are deleted at PF when their featural content is identical to that of the T head they agree with. This phenomenon was previously defined in Sheehan (2006) as *Deletion under feature-identity*. Sheehan further proposes that the pronoun in Spec-TP, rather than the verb in T, is deleted because the verb contains more information than the pronoun. Conversely, PF deletion is not a viable option in the case of strong pronouns, in that they contain an additional [R] feature that is not present on T, so their featural composition is not identical.

Holmberg (2005) defines *pro* as a ϕ P; this definition comes from Déchaine and Wiltschko (2002) discussed in detail in Section 3. Recall that a ϕ P is an element that encodes only valued ϕ features and lacks the D(efinite) layer, the substructure required for a definite referential reading. In the approach adopted in this dissertation, however, both strong pronouns and *pro* have the same DP structure. In strong pronouns, D is endowed with a [uR] feature; as for *pro*, D lacks [uR]. The presented approach does not pose any problems for the idea that finite T encodes D. *pro* also encodes D, but no extra features are added, hence its featural composition represents a proper subset of the features of T and it can be deleted at PF.

This analysis captures the difference in the interpretation of strong and weak subjects without resorting to structural deficiency, building on the idea that null subjects generally refer to the most salient discourse antecedent. In this respect, Cole (2010) proposes that the linking of referring expressions proceeds along the lines of accessibility theory; for instance, he explains the different distribution of first/second- and third-person null subjects with the idea that the referents of first and second person pronouns are always salient in the discourse (see also Erteschik-Shir 2007 in this respect), while referents of third persons may be less salient. The higher probability that a first or second subject is null follows from the fact that thematic null subjects are recovered by subject-verb agreement up to the point of *Morphological*

Maximality.⁹ Morphological Maximality is intended as the maximum point up to which subject-verb agreement occurs in one language. In the case under investigation, Morphological Maximality is represented by φ agreement. First and second persons are generally sufficiently informative and rigid (in accessibility terms) to identify an antecedent¹⁰, while this is not always true for third person. Third person requires an [uR] feature to identify the antecedent, which results in the more frequent overt realisation of the third person pronoun.

The derivation of null and overt subjects in consistent null subject languages then follows an identical path, regardless of φ -composition of the pronoun (as proposed, in a slightly different fashion, in Sigurðsson 2011, 2014). This derivation is easily captured in Holmberg's (2005) approach, as introduced in Section 2.2.

Consider example (26):

(26) Argentinian Neapolitan

a. *pro* ricevə accussì.

pro said.3SG like this

b. *Essə* ricevə accussì.

she said.3SG like this

'She talked like this.'

Recall that both null and overt pronouns have the structure of a DP and follow the same derivation path. The difference is at the internal featural composition of the pronouns, in that D in overt pronouns encodes [uR], while in *pro* it does not, hence it can be deleted under φ -feature identity with T. In both cases, the pronominal DP is merged in Spec-vP and subsequently moved to Spec-TP for Agree with finite T. Recall that, at this point of the derivation, Agree is responsible only for the valuation

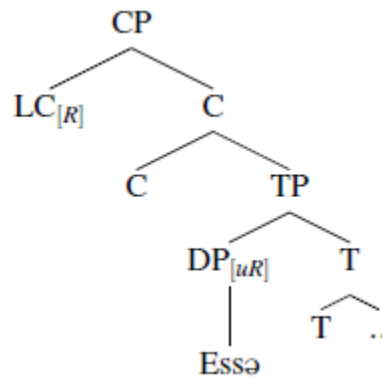
⁹ Morphological Maximality is, in a way, reminiscent of Silverstein (1976). In his analysis too, surface nominative subjects are deleted under co-reference with subjects of higher clauses. Co-reference, specifically for third person, is then expressed by a null form.

¹⁰ In section 6 I discuss the case of Brazilian Venetan first person pronouns, which supports Cole's proposal.

50 The syntax of subject pronouns in heritage languages

of ϕ -features; once CP is merged, the Logophoric Centre (LC) in CP otherwise checks [uR] on the DP if a referentially specific interpretation is required:

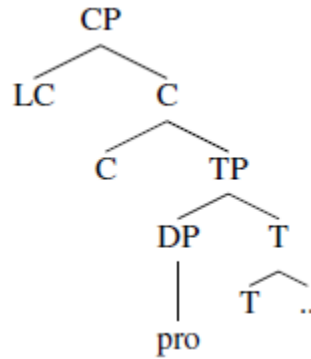
(27)



In (27) a discourse referent $d_{\neq i}$ was selected from the set \mathcal{A} of possible discourse antecedents. Following the model presented in Section 2.4, the Logophoric Centre is hence assigned an [R] feature that checks the closest occurrence of [uR]: the one on the DP subject in Spec-TP. Once the relationship is established, the derivation can proceed. I leave the question of how mismatches are solved (such as the case in which [R] is present on the Logophoric Centre but not on the pronoun in Spec-TP) for Chapter 4.

The context in which the Logophoric Centre is not assigned [R], which results from the selection of a salient referent d_i at the discourse level, makes a pronoun lacking [uR] acceptable. This is the case of *pro*:

(28)



The derivation of null subjects in (28) parallels that of overt subjects in (27). In the case of null subjects, however, the Logophoric Centre is not assigned [R], so a DP lacking [uR] evokes the correct interpretation and the derivation can proceed. Notice that the presence of [uR] on the pronoun but not on the Logophoric Centre would cause the derivation to crash at the LF interface, since the uninterpretable feature cannot be checked and deleted.

2.5.1. Summary

The goal of this section was to show that the approach presented in this study does not contradict or challenge previous analyses of the derivation of null subjects. In particular, the approach is consistent with the idea that null subjects are deleted at PF under feature identity with T. This deletion is allowed in case ϕ -composition of the pronoun in Spec-TP is identical to that of the T head. This is precisely the case: despite being a DP, *pro* does not encode any extra feature. Hence, the null subject is simply a phonologically null variant of a weak pronoun.

In addition, the adopted approach has the desired advantage of shedding some light on the reasons why, in some contexts, PF deletion is blocked. Sheehan (2006) suggested that deletion is not possible because of an extra feature that makes the featural composition of the pronoun not identical to that of the verb in T. This feature is formally defined as [uR], a discourse feature realised in D; [uR] makes the

52 The syntax of subject pronouns in heritage languages

pronoun referentially specific enough to switch the reference to a less salient antecedent, establishing a relationship with a Logophoric Centre encoding [R] that checks [uR] on the pronoun.

2.6. Subject realisation in heritage languages

The data presented in this section come from an investigation on heritage Italo-Romance languages. The data collection was carried out in different countries in the spring of 2019. In this section I give an overview of the study, while the details of the data collection are presented in Appendix A.

The languages under investigation are spoken by communities of descendants of Italian immigrants in Argentina, Belgium, Brazil and Canada. Six different heritage Italo-Romance varieties are included in this study, of which five are from southern Italy and one is from northern Italy.

As far as heritage southern varieties are concerned, three of them belong to the upper southern Italo-Romance group (Eastern Abruzzese, Neapolitan, Northern Calabrian) and the remaining two varieties belong to the extreme southern Italo-Romance group (Southern Calabrian, Sicilian). The study of southern varieties involved 6 heritage speakers in Argentina, in Belgium and in Brazil and it consisted in a semi-guided production task. 192 sentences were considered in total in this study. The details on the methodology, the speakers and the task will be presented in Appendix A.

The heritage northern Italo-Romance variety (Venetan) is also a null-subject language; however, the system of Venetan is more complex, because of the presence of a second paradigm of reduced pronominal forms, subject clitics. Because of this difference, the data from the two groups will be presented separately most of the times, even though speakers of the two groups of languages were asked to perform the same task. In this chapter I only marginally refer to northern varieties; I leave the discussion on their system and on different types of pronouns in Venetan for next chapters.

The study on Venetan involved 24 heritage speakers of heritage varieties of the language in Argentina, Brazil and Canada and it consisted in a semi-guided

production task; 1308 sentences were considered in total in this study. The details on the methodology, the speakers and the task are presented in Appendix A.

2.6.1. The interpretation of null subjects in heritage Italo-Romance

The study presented in this section includes the results of a preliminary exploratory investigation on heritage Italo-Romance varieties in Argentina, Belgium, Brazil and Canada. In these countries, Italo-Romance varieties are spoken in contact with the dominant languages, respectively: Argentinian Spanish, Belgian French, Brazilian Portuguese and Quebec French. The preliminary investigation had the goal to prepare the field for the more specific investigation on interpretive properties of subject pronouns discussed in Chapter 4 and 5. The task was a semi-guided production: informants were asked to tell a short episode from their childhood.

Building on a previous study on subject realisation in heritage Friulian (Frasson *et al.*, 2021), the research question addressed in this preliminary investigation is how heritage null-subject languages deal with interface properties of different types of subject pronouns. In this chapter it was shown that consistent null subject languages are those languages in which referential null subjects are identified through a combination of syntactic and discourse-related properties (Frascarelli 2007; Cole 2009; see also Cognola and Casalicchio 2018). It was shown that heritage Italo-Romance varieties behave consistently with this definition.

This fact challenges the observation, discussed in various studies, that linguistic phenomena related to the C-domain are more vulnerable in bilingual speakers (Platzack, 2001; Hulk & Müller, 2000). This observation was later formulated as the Interface Hypothesis (Sorace and Filiaci, 2006), which claims that interface constructions, intended as those phenomena in which grammar interacts with other domains of language such as discourse and pragmatics, are particularly vulnerable in bilingual language acquisition. The vulnerability results from the increased cognitive load required by the integration of information between different modules, which is particularly problematic for bilinguals, who are constantly switching between languages.

54 The syntax of subject pronouns in heritage languages

The expression and the interpretation of null and overt subjects is one of the most described phenomena in the research on the Interface Hypothesis. In particular, it has been shown that bilingual speakers of a null-subject language do not show native-like competence regarding the discourse-pragmatic constraints regulating the use and distribution of null subjects and overt subjects (e.g. Sorace and Filiaci, 2006, Montrul, 2008; Sorace and Serratrice, 2009).

With respect to the predictions made by the Interface Hypothesis, this study aims to check whether the discourse properties that regulate the distribution of subjects are indeed problematic for heritage speakers. This study shows that the introduction of [R] in the structure of pronouns can capture the variation at the information-structural level displayed by heritage varieties at different levels: cross-linguistic variation, variation among speakers of the same language and intra-speaker variation.

Although Section 2.6.2.2 briefly refers to the variety of Venetan spoken in Italy, the study presented in this chapter does not have the goal to compare heritage varieties to a monolingual baseline of the same language (which does not exist for Italo-Romance varieties; even in Italy, all speakers of these languages are bilinguals) nor to define the effect of contact with a dominant language. The main goal of the study is to define heritage languages with respect to the distribution of null and overt subjects and draw conclusions on the relevance of interface factors involved in the phenomenon.

The results of the investigation are presented in the next section.

2.6.2. Heritage Italo-Romance varieties as consistent null subject languages

The study shows that heritage Italo-Romance varieties are consistent null-subject languages. This is shown by the following examples of heritage varieties spoken in different countries:

(29) Argentinian Eastern Abruzzese

pro è na buonissima wuaglieunə.

pro be.3SG a very good girl

‘She is a very good girl.’

(30) Brazilian Calabrian

pro è itə più viciniə a mmia.

pro be.3SG gone more close to me

‘He came closer to me.’

(31) Belgian Sicilian

Quannu *pro* mi vitti, *pro* mi dissi...

when pro me saw.3SG pro me said.3SG

‘When she saw me, she told me...’

(32) Quebec Venetan

pro ga scominsià contar na storia.

pro have.3SG started tell.INF a story

‘She started to tell a story.’

(29) is an example of heritage Eastern Abruzzese spoken in Argentina, in contact with Spanish; (30) is an example of heritage Calabrian spoken in Brazil, in contact with Portuguese; (31) is an example of heritage Sicilian spoken in Belgium, in contact with French. (32) is an example of heritage Venetan spoken in Quebec, in contact with the local variety of French. In all contact situations, null subjects are accepted and are, in fact, the preferred option in most cases.

An even more striking fact was already discussed in Section 2: heritage Italo-Romance varieties behave consistently with the predictions made by Rizzi in (1982) about the null subject parameter. Consider examples (33-35):

56 The syntax of subject pronouns in heritage languages

(33) Argentinian Calabrian

A **wuagliona mia_j** un parla tantu. Però **pro_j** capisce tuttu.
the daughter mine not speak.3SG much but pro understand.3SG all
'My daughter does not speak a lot, but she understands everything.'

(34) Argentinian Calabrian

Sona **u telefonu**.
ring.3SG the telephone
'The telephone rings.'

(35) Belgian Sicilian

Na fimmina_j, che **pro_k** ci sentiva diri sempri a papà che **pro_j** era forti pi ballari.
A woman that pro.1SG him heard.1SG say always to dad that pro.3SG was.3SG strong for dance
'A tiny woman, that, as I always heard my father say, was a very good dancer.'

Heritage Italo-Romance varieties have all the properties of consistent null-subject languages in the traditional sense: in (33), the null subject refers to the subject of the preceding clause; in (34), the subject is postverbal; in (35), the subject is extracted from an embedded clause headed by an overt complementizer. In other words, heritage Italo-Romance varieties are an instance of pro-drop languages, as per the first definition of the term (Rizzi 1982). In the remainder of this section, different discourse-related variables are checked, in order to test the predictions made by the Interface Hypothesis, starting from southern varieties in next section.

2.6.2.1. Subject realisation in heritage southern varieties

This section presents the results of the study on subject realisation in southern heritage varieties. In this study, different varieties are grouped together and distinguished only by country in which the data were collected. Figure 1 shows that all southern varieties behave uniformly with respect to the realisation of overt and null subject.

Figure 1. Distribution of lexical, overt pronominal and null subjects in each heritage southern variety. Sentences in total: 192.

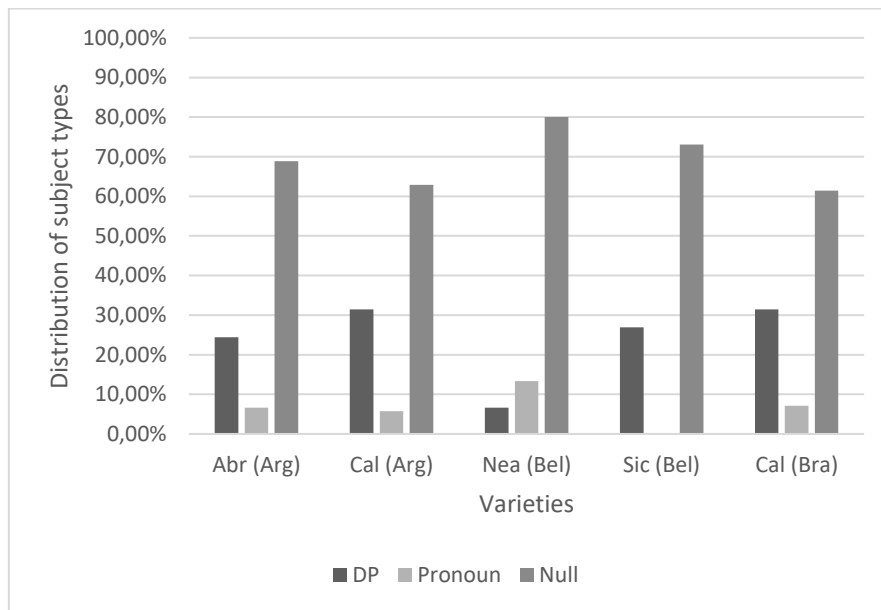


Figure 1 shows that all varieties in the two countries display a similar pattern. All varieties display null subjects, which are always the preferred choice. In consideration of the relatively small sample analysed, the most logical choice for this study is to group the heritage languages according to the country in which they are spoken. Therefore, in this section I will not refer to single southern varieties in each contact situations, but more generally to southern varieties in Argentina, in Belgium and in Brazil.

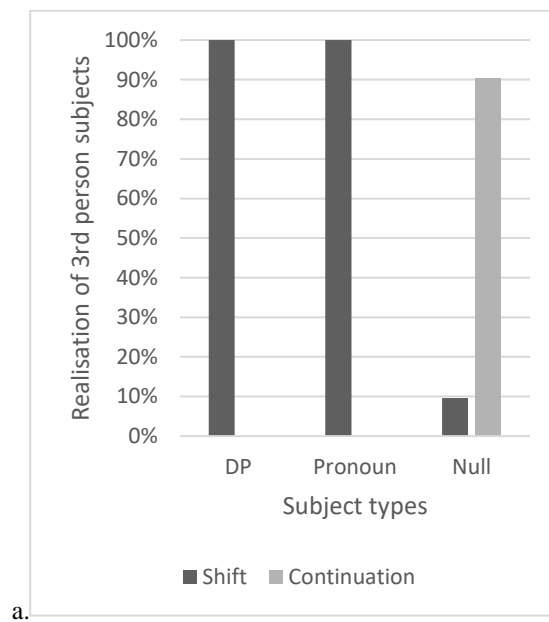
The results, even after grouping the varieties according to the country in which they are spoken, show that null subjects represent the majority of the occurrences.

- b. CONT *pro_j* ha partutu giuvinottu (...).
 pro have.3SG left young
- c. CONT Però *pro_j* ha iutu due volte a l' Italia.
 but pro have.3SG gone two times to the Italy
- ‘Dad worked all his life. He left when he was young. But then he went back to Italy twice.’

In (36)a) the topic shift is signalled by an overt subject. A null subject is used for continuation in (36)b) and (36)c).

The realisation of null and overt subjects in topic shift and continuation contexts is shown in Figure 3:

Figure 3. Realisation of third person subjects in southern varieties in Argentina (a), Belgium (b) and Brazil (c), classified as topic shift and continuation. Sentences in total: 192.



60 The syntax of subject pronouns in heritage languages

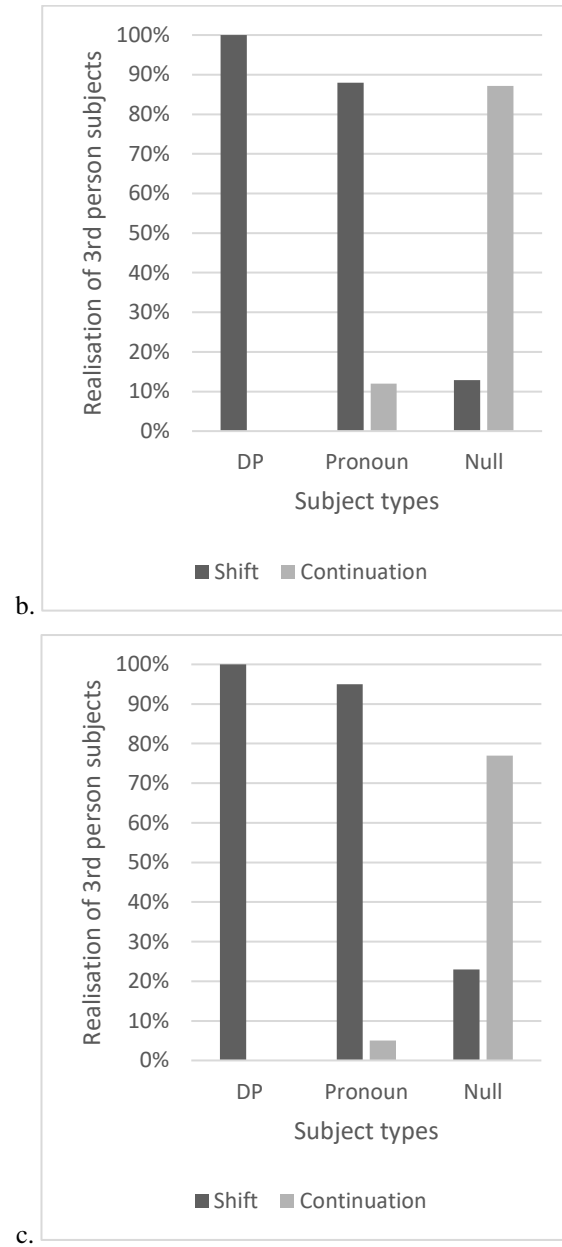
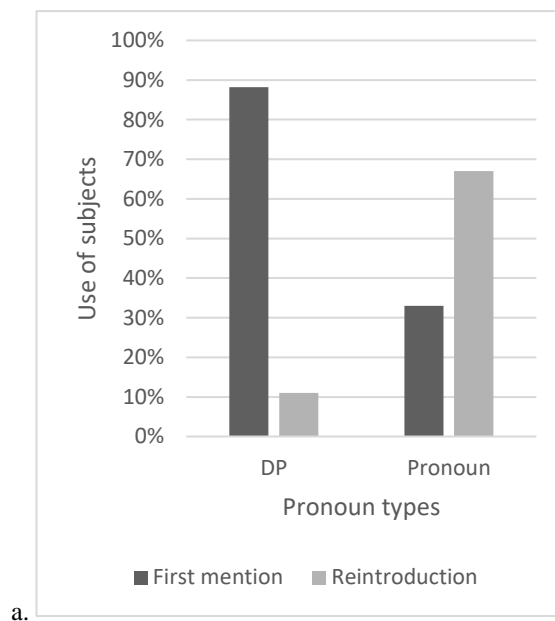


Figure 3 shows a clear preference for overt subjects in a context of topic shift in all contact situations, while null subjects are realised more often in topic continuation.

As far as the distinction between lexical and pronominal subjects, the data were coded for a second discourse-related variable, similarly to what is done in Frasson *et al.* (2021): “MENTION”, with the possible values “FIRST MENTION”, “CONSECUTIVE MENTION” or “REINTRODUCTION”. In this case only the cases of topic shift were included, in order to get a deeper look into the different realisation of the very first mention of a referent in the discourse and the reintroduction of an already known referent after a shift: both cases qualify as topic shift, but the reintroduced referent is expected to be more accessible than the newly introduced one. Consecutive mentions of the same referent were excluded as they qualify as topic continuation, as already shown in Figure 3. The realisation of lexical and pronominal subjects in different mentions is shown in Figure 4:

Figure 4. Use of lexical and pronominal third person subjects; different mentions in the discourse in southern varieties in Argentina (a), Belgium (b) and Brazil (c). Sentences in total: 192.



62 The syntax of subject pronouns in heritage languages

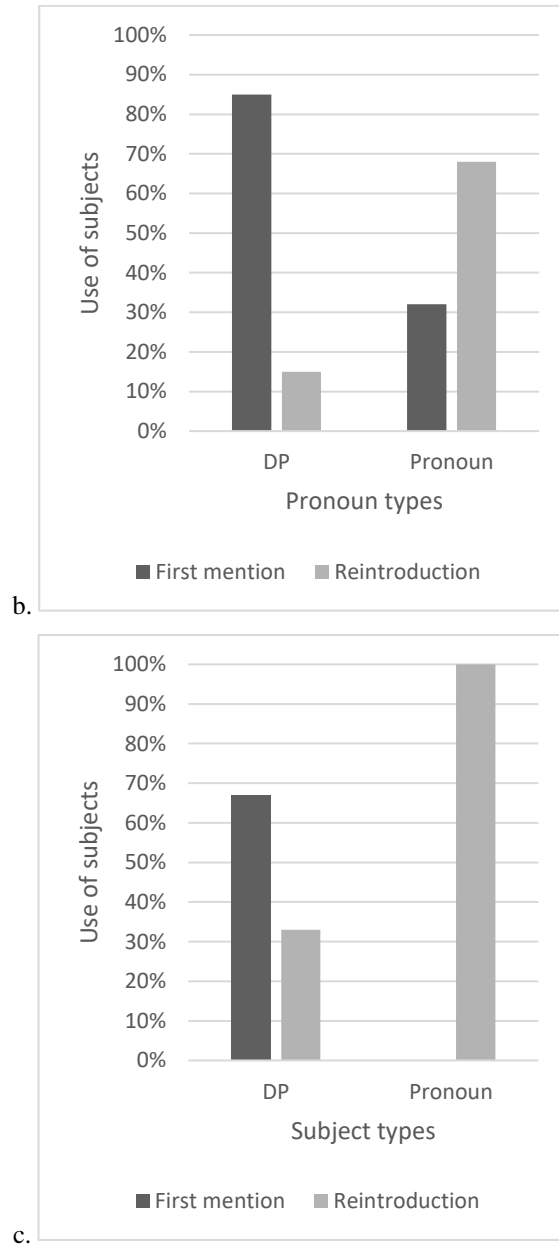


Figure 3 and 4 refer only to the distribution of third person. As already pointed out, first and second person pronouns were excluded because they are intrinsically more

salient in discourse and therefore are most likely to allow for *pro* in the context of topic shift or first mention too.

In general, TOPIC SHIFT is the preferred context for lexical and pronominal subjects. More specifically, FIRST MENTION is the preferred context for lexical subjects and REINTRODUCTION is the preferred context for overt subject pronouns. Recall that REINTRODUCTION implies the shift to a previously introduced referent. CONSECUTIVE MENTION corresponds to instances of topic continuation and is the preferred context for null subjects; conversely, lexical DPs and overt subject pronouns are strongly disfavoured in this context. Consider example (37):

(37) Brazilian Calabrian

- a. 1ST **Mio papà_j** è venutə nel 1949.
 my dad be.3SG come.PRT in.the 1949
- b. CONS **pro_j** s'è spusetə colla mia mamma per la procura
 pro REFL.be.3SG married with.the my mom for the attorney
- c. SHIFT e mamma_k è venutə u febraiə del '53.
 and mom be.3SG come the February of.the 53
- d. SHIFT Io_y sunu netə u novembro del '53.
 I be.1SG born the November of.the 53
- e. REINTR Po' **iddrə_j** ha mannatə a ciamè i parentə.
 then he have.3SG sent to call.INF the relatives

'My dad_j arrived in 1949. Then he_j married my mom through an attorney and mom_k came in February 1953. I_y was born in November. Then he_j called my relatives.'

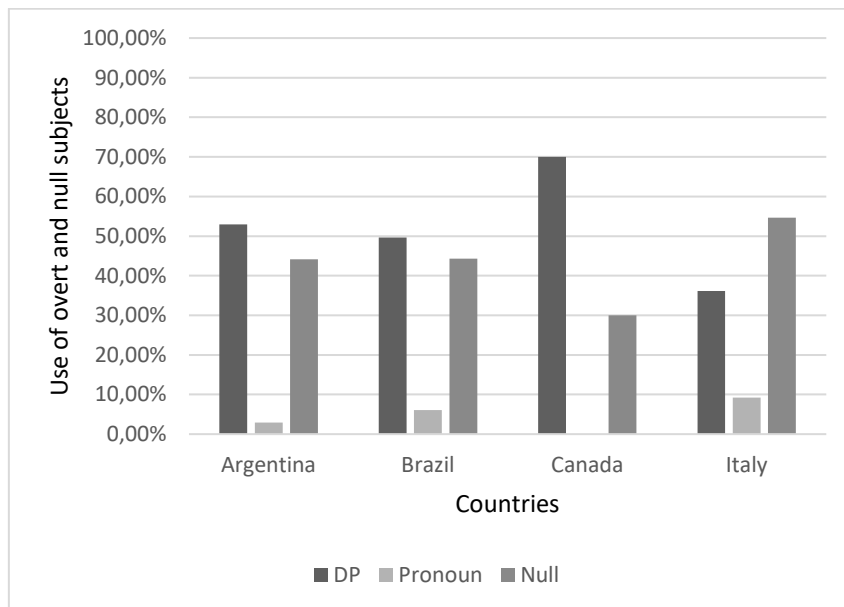
In (37)a), the referent *my dad* is introduced for the first time. In (37)b) the null subject is used for topic continuation. In (37)c) the topic is shifted to *mom* and in (37)d) to *I*. Finally, in (37)e), the referent *my dad* is reintroduced with an overt pronoun.

In conclusion, data from heritage southern varieties exhibit a regular null subject pattern: the distribution of null and overt subjects with respect to discourse variables is expected for consistent null subject.

2.6.2.2. *Subject realisation in Venetan*

This chapter focussed mainly on null subjects in heritage southern varieties. The main reason behind this choice is that southern varieties are consistent null-subject languages without subject clitics; as it will be shown in next chapters, the presence of subject clitics in Venetan adds some complexity to the system. In this section it is shown that, despite displaying slightly fewer null subjects than southern varieties, Venetan varieties spoken in Argentina, Brazil and Canada too are null-subject languages. In this section, the situation of heritage Venetan varieties is compared to that of Italian Venetan. The data from Italian Venetan come from the Microcontact Atlas (<https://microcontact.hum.uu.nl/#contributions>) and include spontaneous productions by 6 native speakers. Recall that the varieties used by speakers in Italy cannot be considered as ‘homeland’ in its traditional understanding: also in the Italian setting, speakers of Venetan are generally dominant in another language (Italian). These data are added for the sake of clarity and comparison with the heritage varieties.

Figure 5. Use of overt and null pronouns in heritage Venetan varieties and Italian Venetan in the different contact situations (countries). Sentences in total: 1308.



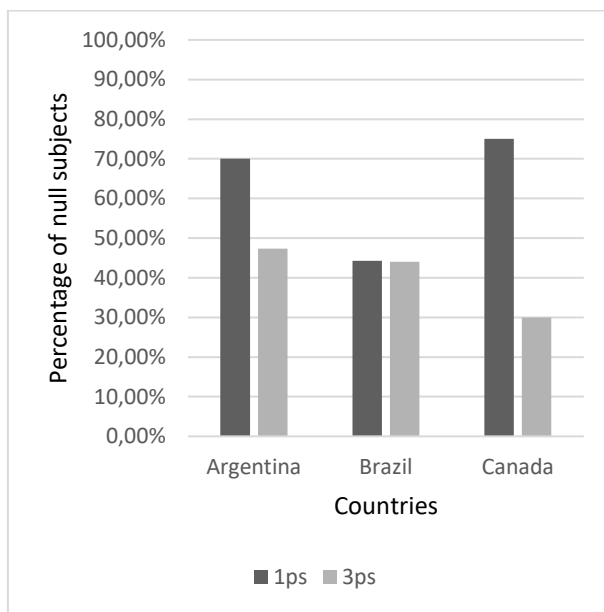
The heritage informants in Argentina, Brazil and Canada leave subjects null, but the number of null subjects is generally slightly smaller than that of overt subjects. This pattern does not parallel that of Italian Venetan, in which null subjects outnumber overt lexical subjects, a behaviour that resembles that of heritage southern varieties.

The fact that null pronominal subjects have a more limited distribution in heritage northern varieties was already noticed in Frasson (2021) with particular reference to first person pronouns in Brazilian Venetan; the change was analysed as the result of a process of attrition or incomplete acquisition in contact with Brazilian Portuguese, as it will be shown in Chapter 3; in that analysis, Brazilian Venetan was seen as involved in an ongoing change from consistent to partial null subject.

The present study focusses on third person and the reasons behind the more limited distribution of first person pronouns in Brazilian Venetan will not be discussed. However, for the sake of comparison, looking at the distribution of null first and third person subjects in the present study on heritage Venetan, it is possible to notice that the distribution of first person null subjects is indeed more limited in the variety of Venetan spoken in Brazil. This is not true for Venetan spoken in Argentina and Canada, which display a very similar pattern to the one identified for southern varieties.

66 The syntax of subject pronouns in heritage languages

Figure 6. Percentage of null subjects per country; first and third persons in heritage Venetan in the different contact situations (countries). Sentences in total: 1308.



As for third person, the distribution of null and overt subjects¹¹ in the context of topic shift and continuation in heritage Venetan is that of a consistent null subject language (and not that of a partial null subject language as Brazilian Portuguese). Just like in heritage southern varieties, in heritage Venetan, third-person null subjects are disfavoured in the context of topic shift.

(38) Brazilian Venetan

- a. SHIFT **Me nono_j** ga comprà a tera.
 my grandfather have.3SG bought the land
- b. SHIFT Na volta a roba prinsipae_k iera darghe un toco de tera ai fioi.
 a time the thing main was.3SG give.them.CL a piece of land to.the children
- c. SHIFT Se i_y vea a tera, vea da magnar.
 if they had.3PL the land had.3PL to eat

¹¹ Here I focus only on strong pronouns in Venetan. Subject clitics will be discussed in next chapter.

d. SHIFT Lora **lu**_j ga comprà ste tere.

so he have.3SG bought these lands

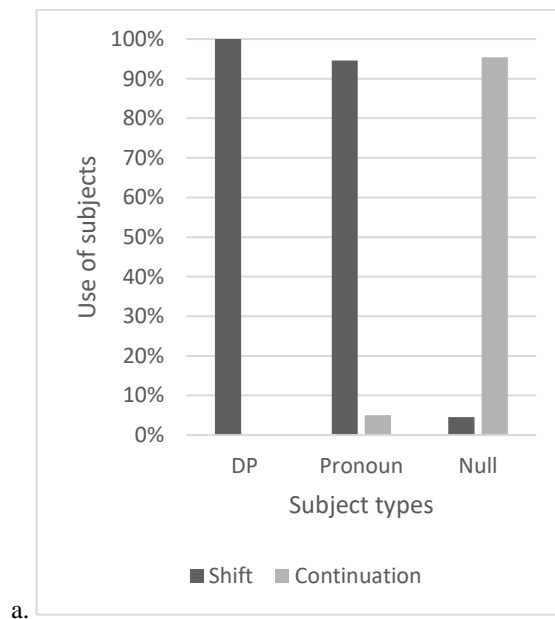
e. CONT **pro**_j ga scominsià vendar el formento par pagar ste tere.

pro has.3SG started sell.INF the wheat for pay.INF these lands

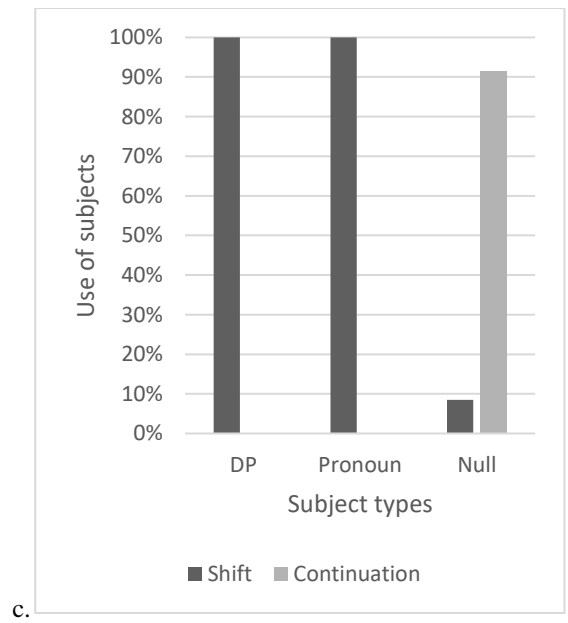
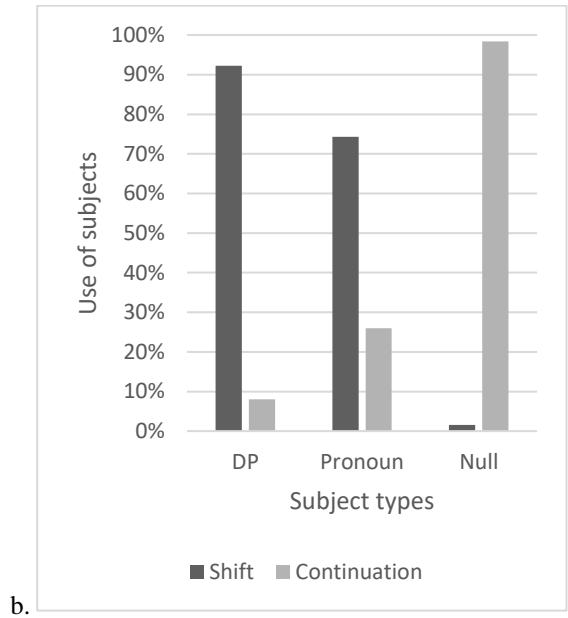
“My grandfather_j bought the land. Once the main problem_k was giving a piece of land to the children. If they_y had land, they could eat. So, he_j bought these lands. He_j started selling wheat to pay these lands.”

The distribution of third person subjects in heritage Venetan in topic shift and continuation is shown in Figure 7:

Figure 7. Realisation of third person subjects in topic shift and continuation in heritage Venetan in Argentina (a), Brazil (b) and Canada (c). Sentences in total: 1308.



68 The syntax of subject pronouns in heritage languages

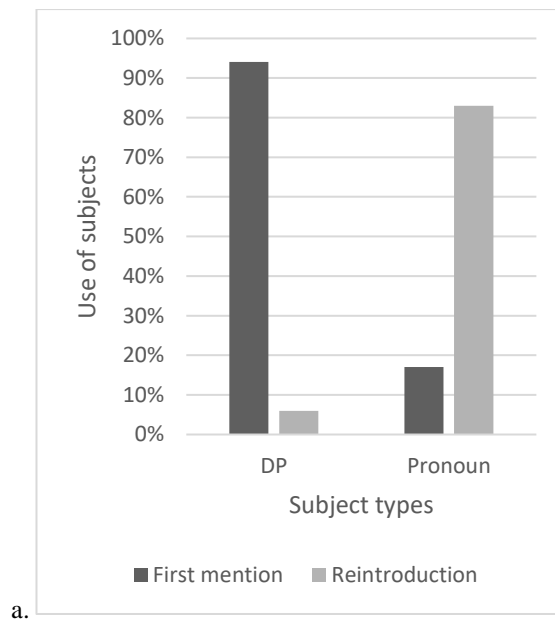


As in the case of southern varieties, the distribution of lexical and pronominal subjects shows a clear preference for these type of subject in a context of topic shift in all

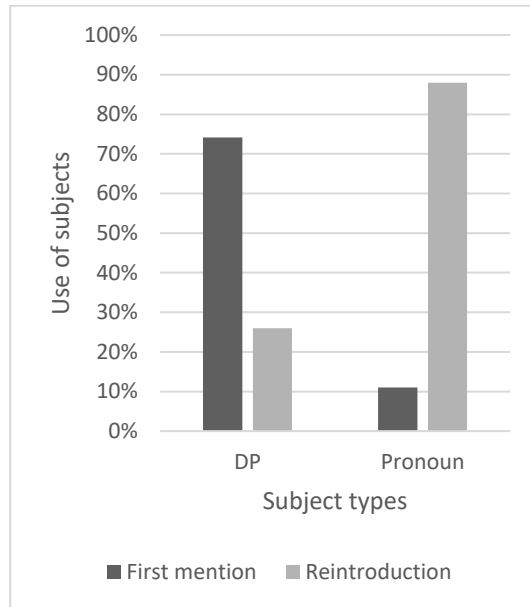
contact situations, while null subjects are realised significantly more often in topic continuation.

Example (38) shows that also with respect to the variable MENTION, Venetan behaves on par with southern varieties: the lexical subject *my grandfather* (38)a) is used for the first mention of the referent; the overt pronoun *lu* (38)d) is used for the reintroduction of the referent after a shift; the null subject (38)e) is used in consecutive mentions of the same referent. Figure 10 shows the distribution of lexical and pronominal subjects in the three heritage varieties of Venetan with respect to the variable MENTION, distinguishing again FIRST MENTION from REINTRODUCTION:

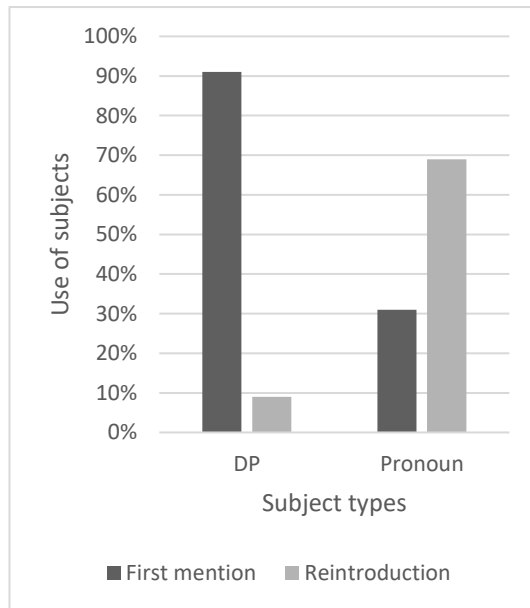
Figure 8. Use of lexical and pronominal third person subject pronouns; different mentions in the discourse in heritage Venetan in Argentina (a), Brazil (b) and Canada (c). Sentences in total: 1308.



70 The syntax of subject pronouns in heritage languages



b.



c.

Also in this case, Venetan behaves almost identically to heritage southern varieties; differences are explained through the different use of the same accessibility markers in heritage varieties.

In conclusion, heritage Venetan has a distribution of null and overt subjects that patterns with consistent null-subject languages with respect to discourse-related variables. Therefore, the proposal presented in Frasson (2021) of a transfer from Brazilian Portuguese to Brazilian Venetan cannot be generalised to all persons and varieties. The pattern displayed by Brazilian Venetan with respect to third person mirrors that of other heritage varieties and is easily captured in the approach on the structure of pronouns presented in this chapter. The case of Venetan and of southern varieties is discussed in the next section.

2.6.3. *Topic shift and switch reference: the role of [uR]*

The picture that emerged from Section 2.6.2 is quite straightforward: all the languages under analysis are consistent null-subject languages, in that they do not restrict the use of null subjects to some persons or conditions. This final section addresses some relevant points that emerged from the observation of the data.

First of all, the distinction between topic shift and continuation and the different mentions of the discourse referents is discussed. Consider again example (37), repeated as (39):

(39) Brazilian Calabrian

- | | | | | |
|-----------|----------------------------|----------------------|------------------|---------------------------|
| a. 1ST | Mio papàj è | venutə | nel | 1949. |
| | my dad | be.3SG | come.PRT | in.the 1949 |
| b. CONS | pro_j s'è | spusetə | colla | mia mamma per la procura |
| | pro | REFL.be.3SG | married with.the | my mom for the attorney |
| c. SHIFT | e | mamma _k è | venutə | u febraiə del '53. |
| | and mom | be.3SG | come | the February of.the 53 |
| d. SHIFT | Io _y sunu | netə | u | novembro del '53. |
| | I | be.1SG | born | the November of.the 53 |
| e. REINTR | Po' iddrəj ha | mannatə | a ciamè | i parentə. |
| | then he | have.3SG | sent | to call.INF the relatives |

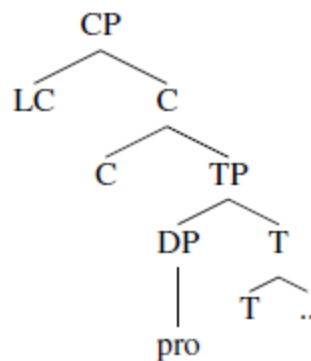
72 The syntax of subject pronouns in heritage languages

‘My dad_j arrived in 1949. Then he_j married my mom through an attorney and mom_k came in February 1953. I_y was born in November. Then he_j called my relatives.’

In (39)a), the referent *my dad* is introduced for the first time. In (39)b) the null subject is used for topic continuation. In (39)c) the topic is shifted to *mom* and in (39)d) to *I*. Finally, in (39)e), the referent *my dad* is reintroduced with an overt pronoun. The approach presented in this chapter can account for the realisation of *pro* in (39)b) and *iddru* in (39)d) in very simple terms.

In (39)b), the set of possible discourse antecedents \mathcal{A} includes only one possible choice: *mio papà* in (39)a). The set can be represented as $\mathcal{A} = \{mio\ papà\}$. Since it is the only element in the set, the antecedent has the highest salience value. There is no change in the salience structure in (39)b), hence the Logophoric Centre in (39)b) is not assigned [R]: this is a case of topic continuation, as the topic of (39)a) is maintained in (39)b). The pronoun in (39)b) does not encode [uR] and undergoes PF deletion under feature identity with the verb and is not phonologically realised, resulting in *pro*. Since there is no [uR] feature that has to be checked and deleted on the pronouns and also no [R] feature on the Logophoric Centre, the derivation can proceed.

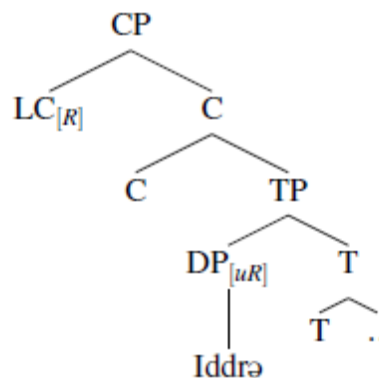
(40)



The case of (39)e) is slightly different, since more potential discourse antecedents have been introduced. Context Scanning defines the set of possible

discourse antecedents $\Delta = \{io, mia\ mamma, mio\ papà\}$, in which the element with the highest salience value corresponds to *io*, followed by *mia mamma* and finally by *mio papà*, which has the lowest salience value. The Logophoric Centre, obviously, cannot select the two elements with the higher salience value, as they do not have the required ϕ -composition: the first element is a first person, and the second is a third person feminine, while the required interpretation is that of a third person masculine. The only possible choice with this ϕ -composition is *mio papà*, which has the lowest salience value. Therefore, the salience structure of the discourse needs to be updated. The Logophoric Centre is assigned [R] as a result of the model presented in Section 2.5; [uR] gets checked against [R] on the pronoun *iddru*. The uninterpretable feature is deleted and the derivation is assigned the correct interpretation at the LF interface. Recall that [uR] remains visible at PF, therefore the pronoun cannot be deleted at PF.

(41)



At the DP level, there is no need to resort to structural deficiency or to rigid classes of pronouns: the difference is captured by the presence of [uR] on D. This system is flexible enough to capture the phenomenon described by Sorace (2011) with respect to the Interface Hypothesis and to the fact that bilingual speakers of a null subject language do not show native-like competence regarding the discourse-pragmatic constraints regulating the use and distribution of null subjects and overt subjects.

In conclusion, the distribution of null and overt subject pronouns in heritage Italo-Romance varieties is that of consistent null-subject languages. I propose that the

74 The syntax of subject pronouns in heritage languages

Logophoric Centre is a universal property of null subject languages and that the distribution of [R] can, potentially, capture also the more complex distribution of subject pronouns in partial null subject languages.

Finally, with respect to interface phenomena in bilingual speakers, data from heritage Italo-Romance varieties show that discourse properties do not represent a problem for speakers of heritage null subject languages.

2.7. Conclusion

This chapter discussed the internal structure of subject pronouns, referring specifically to strong subject pronouns and to null weak subjects. It was proposed that both ‘strong’ and ‘weak’ pronouns have the structure of a DP; the two types of pronouns are distinguished by an [uR] feature encoded by D in strong pronouns and lacked in weak pronouns. The [uR] feature makes the pronoun referentially specific enough to switch the reference to a non-salient discourse antecedent. Weak pronouns lack [uR] and will co-refer with the most salient discourse antecedent. In consistent null subject languages, weak subject pronouns can undergo PF deletion under ϕ -feature identity with finite T, resulting in a phonologically null *pro* (Holmberg 2005, Sheehan 2006, Roberts 2009).

This approach can easily capture the distribution of null and overt subjects in heritage Italo-Romance varieties. These languages are well-behaved null subject languages, as shown by the fact that the distribution of null and overt subjects depends on a precise interaction of agreement and discourse factors. Particularly in the case of first person (the most salient in discourse), ϕ -features can identify a null subject (as per Cole’s Morphological Maximality). This is true also for third person, but the reference is sometimes less accessible; in those cases, an overt pronoun needs to be realised. Also in heritage varieties, the difference between the overt and the null element is the presence of [uR] in the former.

This approach combines traditional views on the interaction of discourse and syntax in the realisation of null subjects (Frascarelli 2007; Sorace *et al.* 2009;

Referentiality and the internal structure of subject pronouns 75

Sigurdsson 2011, 2014) with the analysis of null subjects presented in Holmberg (2005), Sheehan (2006) and Roberts (2009).

In next chapter this proposal is extended to subject clitics. It is shown that the factors determining their realisation are more complex in the case of heritage languages.

76 The syntax of subject pronouns in heritage languages

Chapter 3. The internal structure and the status of subject clitics

3.1. Introduction

This chapter extends to subject clitics the analysis proposed in Chapter 2. A study on subject clitics in Friulian and Venetan is presented, which shows that the analysis of subject pronouns proposed in the previous chapter captures the different distributions of subject clitics in heritage Friulian and Venetan quite neatly; it also allows to dispense with Cardinaletti and Starke's (1999) tripartite model of structural deficiency. In previous studies on northern Italo-Romance varieties, subject clitics are defined as pronominal forms that are grammaticalised as agreement markers and are realised in the inflectional field (Renzi and Vanelli, 1983; Rizzi, 1986; Brandi and Cordin, 1989; Benincà, 1994; Poletto, 1993, 2000). Roberts (2010) proposes that subject clitics are heads encoding ϕ -features that cliticise on a T head. Roberts's (2010) approach extends Baker's (1988) proposal on noun incorporation to cases of head movement such as the cliticisation process that leads to the realisation of subject clitics of the Friulian and Venetan type. However, the idea that some subject clitics encode also discourse-related properties has always been present in the literature on subject clitics, starting from Benincà (1994) and Poletto (2000). It will be shown that such properties emerge more evidently in heritage varieties of Venetan; this fact was already shown for Friulian in Frasson *et al.* (2021).

In this chapter it is maintained that subject clitics in Italian varieties of Friulian and Venetan are the overt counterpart of weak *pro*, in that they lack a [uR] feature. Subject clitics in heritage Friulian and Venetan allow also for a strong pronominal interpretation, as shown by their distribution with respect to different topic conditions; in my terms, they encode a [uR] feature that makes them referentially

78 The syntax of subject pronouns in heritage languages

specific to obviate or switch reference to a non-salient discourse antecedent; this possibility allows subject clitics to have a pronominal distribution, as shown by different tests on their position and realisation in the structure. The alternation between the weak pronominal and the full pronominal use is due to the presence or absence of the [uR] feature on the clitic.

The goal of this chapter is to show that subject clitics are not simple agreement markers, as proposed in previous studies, nor that they can be assigned to the class of severely deficient elements in Cardinaletti and Starke's model. They are instead regular subject pronouns. Building on previous studies on discourse-related properties of subject clitics, it will be shown that their analysis as markers of ϕ -agreement leaves too many open issues on their distribution and realisation also in Italian varieties of the languages under investigation. Therefore, subject clitics are better analysed as pronouns that in the Italian varieties of Friulian and Venetan must be always adjacent to the verb, while in heritage varieties display a more autonomous behaviour in that they can be separated from the verb. This difference is crucially dependent on the availability of a [uR] feature for subject clitics in heritage varieties, but not in Italian varieties.

Like Chapter 2, this chapter focusses mainly on third person subjects, as their referents can encode different salience values. First and second person referents, conversely, are intended as "stage topics" (Erteschik-Shir 1997): they belong to the 'common ground', the part of the information state which is always shared by the speaker and the hearer. For this reason, they are not included in this analysis.

Recall also that this work adopts Miyagawa's (2005, 2010, 2017) Strong Uniformity Principle: discourse-related and information-structural factors are encoded in syntax by means of discourse features that in some languages play a role in syntactic operations.

This chapter is organised as follows. Section 3.2 extends the analysis of the internal structure of pronouns to Friulian and Venetan subject clitics. Section 3.3 presents some studies suggesting that discourse- and context- related factors affect the distribution of subject clitics in Venetan varieties, supporting the idea that they cannot be simple ϕ -heads. Section 3.4 presents the study on heritage Venetan and compare

the results to a previous study on heritage Friulian reported in Frasson *et al.* (2021). Section 3.5 defines the discourse and syntactic properties of subject clitics, with particular reference to heritage varieties of Venetan, in which subject clitics display a clear pronominal behaviour. It is shown that an analysis of subject clitics as pronouns is a desirable result in minimalist terms, for reasons of economy of derivation and representation. Finally, Section 3.6 addresses two phenomena that are related to subject clitics but not strictly relevant for the present analysis, hence not further discussed in the dissertation: the realisation of second person subject clitics and subject clitic inversion. Section 3.7 concludes the chapter.

3.2. Subject clitics and [R]

The two varieties considered in this study are Friulian and Venetan. Friulian is a Rhaeto-Romance language, while Venetan is an Italo-Romance language; they are both spoken in north-eastern Italy, as well as by communities of heritage speakers in Argentina, Brazil and Canada. Their subject pronoun systems consist of two different paradigms: a tonic one and a clitic one. This chapter focusses mainly on subject clitics.

The present section introduces the analysis of subject clitics following the approach presented in Chapter 2. Recall that, in the present approach, the rigid distinction of classes of pronouns presented in Cardinaletti and Starke (1999) is no longer needed. I follow Déchaine and Wiltschko's (2002) pronominal structure model, but unlike them I further propose that all pronominal forms share the same structure, consisting of three layers: D, φ and N. The relevant property that distinguishes strong pronouns from weak pronouns (such as null *pro*) is referentiality; this property is represented as a discourse-related feature [uR], realised on D in strong pronouns in order to make them referentially specific enough to switch the reference to a non-salient antecedent. Weak pronouns are associated with a 'default' salient discourse antecedent, so they do not encode [uR].

(1)

a. [DP_[uR] she [φ P [NP]]]

b. [DP *pro* [φ P [NP]]]

Recall also that *pro* is simply a phonologically null counterpart of an overt weak pronoun. In consistent null subject languages, the null pronoun results from an operation of deletion taking place at PF (Sheehan 2006, Roberts 2009). Subject pronouns are deleted at PF when their featural content is identical to that of the T head they agree with. This is precisely the case of *pro*, whose φ features are identical to those of the verb. Sheehan (2006) also proposes that the pronoun in Spec-TP, rather than the verb in T, is deleted because the verb contains more information than the pronoun. This predicts, for the present analysis, that PF deletion is not a viable option in the case of strong pronouns, precisely because of the additional [uR] feature that is not present on T: their featural composition is not identical. This prediction is borne out, as showed in Chapter 2.

Aside from the two classes already presented (strong and weak pronouns), Cardinaletti and Starke (1999) propose a third class: clitic pronouns. This chapter extends the discussion started in Chapter 2 to overt weak pronouns and to clitics, showing that also this distinction is not supported by empirical data on subject pronouns. Firstly, clitic pronouns in Friulian and Venetan will be defined in Cardinaletti and Starke's model; then, the case of the Brazilian Portuguese weak subject pronoun *clê* will be used to illustrate the more subtle distinction existing between them and subject clitics in Friulian and Venetan.

3.2.1. *Friulian and Venetan subject clitics*

Friulian and Venetan have both strong and clitic subject pronouns; the two paradigms display morphological differences, as shown in Table 1:

Table 1. Tonic and clitic subjects in Friulian and Venetan.

	Friulian		Venetan	
	Tonic	Clitic	Tonic	Clitic
1sg	jo	i, o	mi	-

2sg	tu	tu	ti	Te
3sg	M: lui F: je	M: al F: e	M: lu F: ela	M: el F: la
1pl	no	i, o	nantri	-
2pl	voi	i, o	valtri	-
3pl	lor	a	M: lori F: lore	M: i F: le

According to previous analyses (Benincà, 1994; Brandi and Cordin, 1989; Rizzi, 1986; Poletto, 1993, 2000), subject clitics do not display regular pronominal properties; they rather behave as agreement markers, as part of verbal inflection. In Poletto (1993) subject clitics are defined as lexical items generated in Spec-vP and subsequently moved to a complex Infl head that hosts them and the verb.

Subject clitics, similarly to weak pronouns, have a more restricted distribution with respect to tonic pronouns. They cannot occur in dislocation (2)a) or isolation (2)b), they cannot be coordinated (2)c) or modified by adverbs (2)d). Wherever the tonic pronoun is licensed in these contexts, the subject clitic is not.

(2) Venetan

a. Ela/*La, ze bela.

she she.SCL be.3SG beautiful

‘She is beautiful.’

b. Chi ze che me ciama? – Ela/*La.

who be.3SG that me.OCL call.3SG she she.SCL

‘Who is calling me? – Her.’

c. Marco e lu/*el.

Mark and he he.SCL

‘Mark and him.’

82 The syntax of subject pronouns in heritage languages

d. Solche lu/*el.

only he he.SCL

‘Only him.’

So far, it seems that subject clitics are not different from weak pronouns, as defined in Chapter 2 in Déchaine and Wiltschko’s terms. Both weak pronouns and subject clitics are compatible with an analysis that considers them as ϕ P-pronouns. Recall that, also in Holmberg (2005) and Sheehan (2006), weak pronouns are intended as ϕ Ps that undergo PF deletion under ϕ -feature identity with T. The referential properties of subject clitics are not distinguished from those of weak pronouns. Subject clitics, just as *pro*, display semantic deficiency and are expected to refer to an entity that is already salient in the discourse, as was shown in Chapter 2.

This identity between subject clitics and weak pronouns is at odds with Cardinaletti and Starke’s proposal, as well as previous studies on northern Italo-Romance varieties, according to which subject clitics and weak pronouns belong to different classes. Subject clitics belong to a third class of pronouns, in that they are more structurally deficient than weak pronouns: they are heads, rather than phrases. In other words, Cardinaletti and Starke’s model includes a class of strong elements and two classes of deficient elements (weak and clitic) that differ from each other by the level of structural deficiency they display: the more deficient a pronoun is, the less features (projections in Cardinaletti and Starke’s terms) it encodes. Weak pronouns are structurally impoverished with respect to strong pronouns; clitics are structurally impoverished with respect to both weak and strong pronouns. Properties that are lacked by clitics are a subset of properties lacked by weak pronouns. In Chapter 2, however, it was proposed that there is no need to assume structural deficiency and that all pronouns are DPs. In this chapter it will be shown how the model applies to subject clitics.

According to previous analyses of subject clitics (Rizzi, 1986; Brandi and Cordin, 1989; Poletto, 1993, 2000; Benincà, 1994), the restricted distribution of subject clitics depends precisely on their impoverished structure. Rizzi (1986), building on Kayne (1975, 1983) proposed some tests in order to check the different

distribution of subject clitics and phrasal pronouns. Subject clitics, for instance, always need to be adjacent to the verb: no adverbs or other non-clitic material can be inserted between the verb and the subject clitic.

(3) Friulian

Al (*simpri) dizeve (simpri) cussì.
he.SCL always said.3SG always this
'He always talked like this.'

In addition, subject clitics of the Friulian and Venetan type obligatory double overt subjects: this property depends on the head nature of subject clitics; if they are realised in the inflectional field, they leave the subject position free for a phrasal subject to (optionally) appear.

(4) Friulian

Marie e à comprât il pan.
Mary she.SCL have.3SG bought the bread
'Mary bought bread.'

Finally, according to Cardinaletti and Starke (1999), the cooccurrence of more clitics in a clitic cluster may trigger specific phonotactic processes. In some Friulian varieties, for example, subject clitics are dropped in clitic clusters:

(5) Friulian

a. (*Al)=lu à cjalât.
he.SCL=it.OCL have.3SG watched
'He watched it.'

b. (*E)=si è pierdude.
she.SCL=REFL.CL be.3SG lost
'She got lost.'

84 The syntax of subject pronouns in heritage languages

- c. (*A)=no son fruts.
they.SCL=not be.3PL children
'They are not children.'

These properties of subject clitics are not displayed by weak pronouns, such as the reduced subject pronoun *cê* in colloquial Brazilian Portuguese¹², as shown by Kato (1999, 2000). This form allows for non-clitic material to be inserted (6)a, it can double a lexical subject only when this is dislocated in the left periphery (therefore, when the canonical preverbal subject position, Spec-TP, is empty) (6)b and does not cluster with clitics (6)c):

(6) Brazilian Portuguese

- a. *Cê sempre fala isso.*
you always speak.2SG this
'You always say this.'
- b. *Você, cê ri de todo mundo.*
you you laugh.2SG of all world
'You make fun of everyone.'
- c. *O que que cê me fala?*
the what that you me.OCL speak.3SG
'What is it that you are telling me?'

Cardinaletti and Starke state that a difference like the one identified between Friulian and Venetan subject clitics on the one hand and Brazilian Portuguese *cê* on the other depends precisely on the fact that the former are heads, while the latter is a

¹² The same analysis proposed here for Brazilian Portuguese reduced subject pronouns holds for French subject pronouns. See Rizzi (1986) and Sportiche (1996) for an analysis of French data.

phrase. In other words, while subject clitics are severely deficient elements, the weak pronoun *cê* is “mildly” deficient: despite displaying an impoverished distribution with respect to dislocation, isolation, coordination and modification (on a par with subject clitics), it is a phrasal pronoun.

However, the approach adopted here dispenses with structural deficiency; it is claimed that all pronouns have the structure of a DP. Hence, for subject clitics too, there should be no need to claim that they belong to a class of severely deficient elements with an extremely reduced featural composition. The goal of this chapter is to show that this is indeed the case. In Chapter 2 it was proposed that an [uR] feature can be assigned to D in order to obtain a referentially specific interpretation. If subject clitics are simple φ -agreement markers, they should not encode an extra discourse feature and their realisation should not depend on discourse factors. In the present model, subject clitics should have the following structure:

(7)

[DP *SCI* [φ P [NP]]]

This is however not yet the full picture. In what follows, it will be shown that it is not possible to ascertain whether subject clitics are simple φ -agreement markers; it is also proposed that subject clitics can encode a [uR] feature, yielding an interpretation comparable to that of strong pronouns, in that it is strictly dependent on discourse-related and information-structural considerations. This is shown in (8):

(8)

[DP_[uR] *SCI* [φ P [NP]]]

The idea that subject clitics belong to a class of severely deficient elements is too strong, as it would not capture the possibility of their strong pronominal behaviour in heritage varieties. Several facts on the distribution and the interpretation of subject clitics cannot be captured by the analyses proposed in previous studies.

3.3. Subject clitics, pragmatics and discourse

The idea that tonic pronouns and subject clitics in varieties such as Venetan and Friulian belong to different classes and display properties that crucially distinguish them is maintained in many studies (Brandi and Cordin 1981, 1989; Benincà 1983, 1994; Rizzi 1986; Poletto 1993, 2000; Manzini and Savoia 2005). One of the assumptions of these studies is that subject clitics belong to the class of severely reduced elements (in Cardinaletti and Starke's model), lacking referential properties and therefore the [uR] feature introduced in Chapter 2. However, some of the aforementioned studies show that the distribution of subject clitics is not always that of obligatory agreement markers.

3.3.1. *The restriction on doubling*

One first piece of evidence against an analysis of subject clitics as pure agreement markers comes from a constraint on doubling in Venetan. Pescarini (2020) shows that the distribution of subject clitics in Venetan is not that of obligatory agreement markers, but of pronouns. As already noticed by Benincà (1994), Venetan subject clitics do not double a DP subject when this is postverbal (9)a) and they optionally double a preverbal subject (9)b):

(9) Venetan (Benincà 1994: 20)

a. (*El) riva to fradeo.
he.SCL arrive.3SG your brother
'Your brother is arriving.'

b. Mario (el) compra na casa.
Mario he.SCL buy.3SG a house
'Mario is buying a house.'

In (9)a), the realisation of a subject clitic is blocked by the presence of a postverbal subject; this restriction supports the idea that the subject clitic is realised as the external argument of the verb and targets the same position as the lexical subject *to fradeo* (Spec-TP) rather than a complex Infl-head as proposed by Poletto (1993). In (9)b), optionality depends on discourse-related properties of the DP subject: the subject clitic is realised when the DP subject is topicalised, as a form of resumption, otherwise it is dropped.

Pescarini (2020) observes that these constraints provide fundamental evidence against Cardinaletti and Starke (1999)'s model, in that the phonological deficiency of subject clitics does not necessarily imply their structural deficiency. As we will see in this chapter, the definition of classes of pronouns is problematic, in that the alternation between different pronominal forms may depend on external factors, rather than following from a strict distinction related to the internal structure of pronouns.

3.3.2. *The Venetan discourse-related clitic*

The observation that subject clitics can have a function related to the information structure of the sentence is not totally new. A very well-known case of this type is the clitic *a* in Venetan. The syntactic properties of this element have been widely discussed in the literature on subject clitics (see in particular Benincà 1983, 1994; Poletto 1993, 2000). This clitic evolved from a first person singular nominative pronoun in Latin; however, Benincà proposes that this element cannot be considered a subject clitic on par with the others. This is shown by the fact that the clitic *a* is not constrained by syntactic conditions as the ones involved in the realisation of other subject clitics, but to rather “pragmatic” conditions, to express surprise or emphasis. Besides, the *a* clitic can optionally appear with all grammatical persons, even with postverbal subjects, in which case subject clitics are normally not realised in Venetan, as shown in Section 3.1. In (10)a) there is no subject clitic, since the subject appears postverbally, but the clitic *a* can be optionally realised; (10)b), with a third person masculine subject clitic and a postverbal subject, is not grammatical; (10)c) shows

88 The syntax of subject pronouns in heritage languages

that the clitic *a* can be optionally realised even in cooccurrence with another subject clitic.

(10) Venetan (Benincà 1994: 20-21)

a. (A) riva Giorgio.

A.CL come.3SG George

‘George is coming.’

b. *El riva Giorgio.

he.SCL come.3SG George

‘George is coming.’

c. (A) te parli sempre.

A.CL you.SCL talk.2SG always

‘You are always talking.’

To support the idea of a pragmatic role of this clitic, Benincà shows that *a* is the only clitic that can appear in imperative clauses:

(11) Venetan (Benincà 1994: 25)

A sona pì pian!

a.CL play.IMP more quiet

‘Play more quietly!’

Benincà proposes that the *a* clitic in (11) does not refer to the subject, but it rather has the function of marking the whole sentence as “new information”.

The idea that not all subject clitics have precisely the same featural composition is also crucial in the work by Poletto (2000). In her proposal, most subject clitics are found in the inflectional field: in Roberts’ (2010) terms, they are heads encoding ϕ -agreement and realised on the verb, sharing the same ϕ features of the

subject. However, there is a restricted group of clitics that in Poletto's approach are found in a higher functional field, possibly the CP.

(12) (adapted from Poletto 2000: 36)

[FP1 invariable SCI/deictic SCI [NegP [FP2 number SCI/person SCI]]]

The elements found in FP1 have little or no connection to the ϕ features of the subject and they rather have a pragmatic or deictic function. Subject clitics encoding the ϕ features of the subject, however, are obligatorily realised in a lower functional field FP2, inside the TP.

Hence, the idea that subject clitics can encode discourse features has been already explored. The discourse-related interpretation identified by Benincà (1994) and Poletto (2000) for some clitics should therefore be extended to the whole paradigm of subject clitics.

3.3.3. *The Štivorian context-related expletive clitic*

A first support for the idea that extra-syntactic factors influence the realisation and the interpretation of all subject clitics, including the ones traditionally assigned to the inflectional field, was presented in a study by Casalicchio and Frasson (2019) on Štivorian, a Venetan-Trentino variety used by a community of heritage speakers in Bosnia-Herzegovina. Štivorian displays a phenomenon that is rare among Italo-Romance varieties (see: Tortora, 1997 for Borgomanerese; Bentley *et al.* 2015 for Sardinian) and not attested, to my knowledge, in other Venetan varieties. In existential constructions, two types of expletive clitics are used: the locative clitic *ghe* (13a) or the dummy clitic *l'*¹³ (13b).

¹³According to Poletto (1993), *l'* is not a subject clitic, but an 'auxiliary clitic', a dummy form realized exclusively with 3rd person singular present form of the verb 'to be'. According to Garzonio and Poletto (2011), the realization of *l'* is independent of the subject but may depend

90 The syntax of subject pronouns in heritage languages

(13) Štivorian (Casalicchio and Frasson, 2019)

a. Gh' è tanti ucraini qua ncora.

LOC.CL be.3 many Ukrainians here still

‘There are still many Ukrainians here.’

b. Su quele tere l' è i serbi.

on those lands SCL be.3 the Serbs

‘The Serbs are on those lands.’

The pattern shown in (13) is not attested in Venetan and Trentino varieties spoken in Italy, in which all existential clauses use either the locative or the dummy clitic¹⁴, but not both. The choice of the clitic in Štivorian is not optional; it is strictly dependent on the context and is determined by speech location: the locative clitic (13)a) is used when the event place matches the place of the speaker; the dummy clitic (13)b) is used otherwise.

Casalicchio and Frasson (2019) propose that speech location can be represented in the Logophoric Centre in C (see also Bianchi 2003) that, as showed in Chapter 2, can contain also spatial coordinates¹⁵ and licenses an appropriate expletive

on the auxiliary itself. According to Parry (1995), the *l'* element may be the grammaticalization of an object clitic as part of the verbal morphology.

¹⁴ Note that Štivorian, just as the varieties considered in this study, is a heritage language. It is spoken in contact with Serbian, which is the dominant language for all the speakers of the variety.

¹⁵ The presence of a null locative argument in C is not new; it was proposed for unaccusative constructions with a postverbal subject (see in particular: Benincà 1988; Pinto 1997; Sheehan 2010, 2016; Tortora 1997, 2014; Bentley & Cruschina 2018). According to Francez (2007), the location itself, which is held to be part of utterances about the existence or presence of an entity, is always implicit and determined by context.

clitic on T¹⁶. I come back to the meaning and the role of the Logophoric Centre in Section 5.

In the context of the present study, the case of Štivorian is relevant in that it shows that the realisation of inflectional subject clitics depends also on discourse- and context- related factors. In the following section it is shown that such considerations do not apply exclusively to expletive constructions, but also to referential ones.

3.3.4. Subject clitics in Brazilian Venetan

Before moving to the present study, I introduce a study, presented in Frasson (2021), about the nature of subject clitics in Brazilian Venetan. This study shows that Brazilian Venetan subject clitics tend to mirror the behaviour of reduced subject pronouns of colloquial Brazilian Portuguese (Duarte 1995, Negrão and Müller 1996, Kato 2000, Kato and Duarte 2014), in cases in which these pronouns are realised. Hence, for instance, Brazilian Venetan subject clitics can appear before the preverbal negation marker, unlike the expected post-negation position described in the literature for Italian Venetan. This position mirrors the one found in Brazilian Portuguese for reduced pronoun.

(14) Italian Venetan (Frasson, 2021)

No'l me ga dito gnente.
not he.SCL me.OCL have.3SG told nothing
'He did not tell me anything'

(15) Brazilian Venetan (Frasson, 2021)

El no me ga dito gnente.
he.SCL not me.OCL have.3SG told nothing

¹⁶ See also Ritter & Wiltschko (2014): they propose that T is not only a Tense head, as generally assumed, but it may also encode Space. They propose a coincidence [*u* coin] feature that may express either time, space or participant coordinates.

92 The syntax of subject pronouns in heritage languages

‘He did not tell me anything.’

(16) Brazilian Portuguese (Frasson, 2021)

Cê não me disse nada.
you not me.CL tell.2SG.PAST nothing
‘You did not tell me anything.’

Besides, Brazilian Venetan displays interpolation: a non-clitic element, like an adverb, can be inserted between the clitic and its verbal host (17). Interpolation is not possible in Italian Venetan: in this variety, temporal adverbs like *sempre* appear postverbally (18). However, the same adverb can appear preverbally in Brazilian Portuguese (19):

(17) Brazilian Venetan (Frasson, 2021)

El sempre parlea cusita.
he.SCL always spoke.3SG like this
‘He used to speak like this.’

(18) Italian Venetan

El parlea sempre cussi.
he.SCL spoke.3SG always like this
‘He used to speak like this.’

(19) Brazilian Portuguese (Ambar, Gonzaga and Negrão 2004)

O João sempre come bolos.
The John always eat.3SG cakes
‘John always eats cakes’.

This difference suggests that, in such contexts, Brazilian Venetan subject clitics behave as weak pronouns rather than agreement markers, being therefore more autonomous with respect to the verb than their Italian Venetan counterparts.

Further evidence of the identity between Brazilian Portuguese weak pronouns and weak-pronominal subject clitics in Brazilian Venetan comes from dislocation facts. According to Cardinaletti and Starke (1999), reduced forms, both weak and clitic, cannot be dislocated. In fact, in Brazilian Venetan, as in Brazilian Portuguese, when the subject pronoun is dislocated, it needs to appear in its strong pronominal form:

(20) Brazilian Venetan (Frasson, 2021)

LU/(**EL*), no me ga dito gnente.
he he.SCL not me.OCL have.3SG told nothing
'He did not tell me anything.'

Example (20) shows that, in a contrastive environment, a strong pronoun appears. As predicted by Cardinaletti and Starke (1999), a reduced form cannot be realised in dislocation.

Frasson (2021) concludes that the contact language, Brazilian Portuguese, affects the nature of subject clitics in Brazilian Venetan. In this chapter it is shown that the analysis needs to be refined. In particular, the contact language may not be that important in shaping the change: it is shown that subject clitics in different cases of contact, with languages that do not have reduced pronominal forms such as Spanish, display similar pronominal properties. Besides, in the remainder of this chapter, it will be proposed that the trigger for such anomalous behaviour of subject clitics in heritage Venetan is connected to the optional presence of [uR] on subject clitics, and not on a distinction between weak and clitic elements. Before moving to the analysis, Section 4 introduces the data from the study that will be used to support the hypothesis tested in this study.

3.4. The interpretation of subject clitics in heritage Friulian and Venetan

This study compares the heritage Friulian data presented in Frasson *et al.* (2021) to the Venetan data collected in Argentina, Brazil, Canada and Italy that were already introduced in Chapter 2.

Recall that the study involves 31 speakers of Venetan in four different countries: 2 speakers in Argentina, 22 in Brazil, 1 in Canada and 6 in Italy. I refer to Argentinian, Brazilian and Canadian Venetan as heritage Venetan in order to distinguish them from Italian Venetan. The details on the methodology, the speakers and the task are summarised in Section 3.4.1 and fully presented in Appendix A.

The Venetan data will be compared to the findings presented in Frasson *et al.* (2021) on Friulian.

3.4.1. Presentation of the study: tasks and methodology

The data were collected through interviews carried out during a fieldwork in Argentina, Brazil and Canada. At the beginning of the interview, all informants were asked to answer a sociolinguistic questionnaire that included questions on their background: their exposure to Venetan and to the contact language, their education and family situation.

Following the sociolinguistic questionnaire, informants were required to perform a preference task and a semi-guided production task: the two tasks are described in the following sections.

Before introducing the data and discussing the results, a disclaimer is in order. The data presented in this chapter have been collected during an exploratory fieldwork that had the main goal of locating communities and speakers of the varieties under investigation in Argentina, Brazil and Canada. Even though a data collection were carried out during the fieldwork, there are some missing pieces of information (for instance: some informants did not carry out both tasks or did not complete one of them), as it will emerge in the following sections. Because of such discrepancies,

some data had to be excluded, as it will be specified in the remainder of this section. The details of the data collection are described in Appendix A and B.

3.4.1.1. Preference task

A short preference task was used to test the use of subject clitics in five syntactic contexts, exemplified in (i-v). Participants had to choose between two proposed sentences, one with a subject clitic and one without it. Five items were used to test the hypothesis already presented in Frasson (2021) and Frasson *et al.* (2021) that subject clitics in heritage varieties allow for a pronominal interpretation, a fact that results in the possible drop of the subject clitics even in contexts in which they would be obligatory in the same varieties as spoken in Italy.

The behaviour in contexts (i-v) are taken to distinguish a subject clitic (an agreement marker) from a subject pronoun, as already proposed in Rizzi (1986), Brandi and Cordin (1989) and Cardinaletti and Starke (1999). Contexts i, ii and iii were also tested for Friulian in Frasson *et al.* (2021).

i. the doubling of a pronominal or lexical subject. Recall that in Venetan a subject clitic obligatorily doubles a pronominal or lexical subject (21)a); subject clitic drop (21)b) is a first piece of evidence of pronominal behaviour:

(21)

- a. La Maria la ga cromptà el pan.
the Mary she.SCL have.3SG bought the bread
- b. La Maria ga cromptà el pan.
the Mary have.3SG bought the bread
- ‘Mary bought bread.’

ii. the doubling of a topicalised subject: subject clitics obligatorily double lexical and pronominal subjects also in this case (22)a); the absence of a subject clitic, again, evidences a pronominal behaviour (22)b):

96 The syntax of subject pronouns in heritage languages

(22)

- a. Marco, ieri, el ga magnà massa.
Mark yesterday he.SCL have.3SG eaten too much
- b. Marco, ieri, ga magnà massa.
Mark yesterday have.3SG eaten too much
'Yesterday Mark ate too much.'

iii. the repetition in both conjuncts in coordinated structures: if subject clitics are obligatory agreement markers, they should be realised on every coordinated finite verb (23)a); if the subject clitic is realised only in the first conjunct, it is most likely pronominal (23)b):

(23)

- a. El magna e el beve.
he.SCL eat.3SG and he.SCL drink.3SG
- b. El magna e beve.
he.SCL eat.3SG and drink.3SG
'He is eating and drinking.'

iv. the position with respect to preverbal negation: a subject clitic usually follows preverbal negation (24)a); if it precedes it, it is most likely a pronoun (24)b):

(24)

- a. No 'l me ga visto.
not he.SCL me.OCL have.3SG seen
- b. El no me ga visto.
he.SCL not me.OCL have.3SG seen
'He has not seen me.'

v. the interpolation of an adverb: subject clitics cannot be separated from the verb by non-clitic elements such as adverbs (25)a); if an adverb can be inserted, they display a more autonomous pronominal behaviour (25)b):

(25)

a. El dizea sempre cussì.

he.SCL said.3SG always like this

b. El sempre dizea.3SG cussì.

he.SCL always said like this

‘He always talked like this.’

For heritage Venetan varieties, participants were presented with auditory stimuli recorded by a native speaker of Italian Venetan. Each stimulus contained a sentence pair testing one of the phenomena in (i-v). The sentences were randomised. After listening to the sentences, participants were asked to choose the one they preferred.

3.4.1.2. *Semi-guided production task*

For the semi-guided production task, informants were asked to tell a short story about their childhood. The interviewer asked the informants to say something about their childhood. The recorded data was subsequently transcribed and annotated. The general details on the corpus of heritage Venetan speech are described in Appendix B.

Production data was used to support elicited data in the questionnaire and to check whether the use of clitics depends also on other linguistic or non-linguistic features.

All sentences containing finite verbs were coded for the syntactic variables CLITIC USE, PERSON and SUBJECT TYPE. CLITIC USE is understood as the actual realisation of subject clitics by the participants in contexts in which a subject clitic is expected. For PERSON, only third person subject clitics were included in the model,

98 The syntax of subject pronouns in heritage languages

while first and second persons were excluded: Venetan does not have first person subject clitics, while the exclusion of second person is discussed in Section 6. Finally, SUBJECT TYPE is understood as the type of subject doubled by a subject clitic: lexical (DP) subject, null subject, or strong pronominal subject.

Each sentence was also coded for the discourse-related variable TOPIC. This variable was used to show whether there was a continuation or a shift in topic with respect to the previous sentence. If the topic of the sentence in which the subject clitic appeared was the same as in the previous sentence, the occurrence of the subject clitic was coded as *cont*; otherwise, if the topic of the sentence was not the same as previous sentence (including the cases of reintroduction of an old referent after some time and introduction of a totally new referent), the sentence was coded as *shift*.

For the heritage varieties, 824 sentences were coded, of which 540 contained a subject clitic and 284 did not. For the Italian Venetan data, 119 sentences from the Microcontact atlas were coded, of which 98 contained a subject clitic and 21 did not.

3.4.2. Results of the study

3.4.2.1. Preference task

In the context of non-topicalised DP subjects, the Argentinian and the Canadian informants¹⁷ chose the sentence in which the DP subject is not doubled by a subject clitic. Most Brazilian informants chose the sentence without a subject clitic too:

Table 2. Realisation of SCIs with non-topicalized DP subjects in heritage Venetan.

	Argentina	Brazil	Canada	Total
La Maria la ga cromptà el pan. the Mary she.SCL has bought the bread	0	7	0	7
La Maria ga cromptà el pan. the Mary has bought the bread 'Mary bought bread.'	1	15	1	17

¹⁷ Bear in mind that the results for Argentina and Canada represents the choices made by only one speaker per country.

The internal structure and the status of subject clitics 99

Total	1 ¹⁸	22	1	24
-------	-----------------	----	---	----

Despite the difference in numbers, Argentinian and Brazilian Venetan data seem to resemble the Friulian data reported in Frasson *et al.* (2021):

Table 3. Realisation of SCIs with non-topicalized DP subjects in Heritage Friulian (adapted from Frasson *et al.*, 2021)

	Argentina	Brazil	Total
Maria e à comprât il pan. Mary she.SCL has bought the bread	0	1	1
Maria à comprât il pan. Mary has bought the bread 'Mary bought bread.'	4	5	6
Total	4	6	10

In the case of topicalised DP subjects, the Argentinian informant chose the sentence without a subject clitic, while the Canadian informant chose the sentence with a subject clitic. In the case of Brazilian informants, most of them chose the sentence with a subject clitic:

Table 4. Realisation of SCIs with topicalized DP subjects in heritage Venetan.

	Argentina	Brazil	Canada	Total
Marco, ieri, el ga magnà massa. Mark yesterday he.SCL has eaten too much	0	16	1	17
Marco, ieri, ga magnà massa. Mark yesterday has eaten too much 'Yesterday Mark ate too much.'	1	6	0	7
Total	1	22	1	24

The choices made by Venetan informants in this case do not mirror the ones reported for Friulian in Frasson *et al.* (2021), especially in the case of Argentina:

¹⁸ A total of 2 Venetan speakers were interviewed in Argentina. However, only one could carry out the questionnaire.

100 The syntax of subject pronouns in heritage languages

Table 5. Realisation of SCIs with topicalized DP subjects in Heritage Friulian (adapted from Frasson et al., 2021)

	Argentina	Brazil	Total
Marco, òr, al à mangjât masse. Mark yesterday he.SCL has eaten too much	3	6	9
Marco, òr, à mangjât masse. Mark yesterday has eaten too much 'Yesterday Mark ate too much.'	1	0	1
Total	4	6	10

In the case of coordinated structures, the Argentinian and Canadian informants chose the sentence with the repeated subject clitic; this is true also for most Brazilian participants, although quite a few of them chose the sentence with only one clitic (in the first conjunct):

Table 6. Realisation of SCIs in coordinated structures in heritage Venetan.

	Argentina	Brazil	Canada	Total
El magna e el beve. he.SCL eats and he.SCL drinks	1	13	1	15
El magna e beve. he.SCL eats and drinks 'He is eating and drinking.'	0	9	0	9
Total	1	22	1	24

As in the previous context, the choices made by Venetan informants are quite different from the ones reported for Friulian in Frasson *et al.* (2021):

Table 7. Realisation of SCIs in coordinated structures in Heritage Friulian (adapted from Frasson et al., 2021)

	Argentina	Brazil	Total
Al mangje e al bef. he.SCL eats and he.SCL drinks	2	5	7
Al mangje e bef. he.SCL eats and drinks 'He is eating and drinking.'	2	1	3
Total	4	6	10

The internal structure and the status of subject clitics 101

In the case of interpolation, most Brazilians chose a sentence with an interpolated adverb; the Argentinian and the Canadian informants chose the sentence with an adverb in its (non-interpolated) expected position:

Table 8. Realisation of SCLs in coordinated structures in heritage Venetan.

	Argentina	Brazil	Canada	Total
El dizeva sempre cussi. he.SCL said always this	1	8	1	10
El sempre dizeva cussi. he.SCL always said this 'He always talked like this.'	0	14	0	14
Total	1	22	1	24

With respect to the position of the negation marker (preceding or following the clitic), most Brazilians and the Canadian informant chose a sentence with a negation marker following the clitic; the Argentinian informant, on the other hand, chose a sentence with a negation marker preceding the clitic:

Table 9. Position of the negation marker with respect to the subject clitic in heritage Venetan.

	Argentina	Brazil	Canada	Total
No 'l me ga visto. not he.SCL me.OCL has seen	1	7	0	8
El no me ga visto. he.SCL not me.OCL has seen. 'He did not see me.'	0	15	1	16
Total	1	22	1	24

In sum, it seems that in heritage Venetan subject clitics display pronominal behaviour, but the results are not conclusive, in that the agreement-like behaviour is also still possible. This may depend on the speaker or the context. A similar finding is shown by the result of the study on heritage Friulian presented in Frasson *et al.*

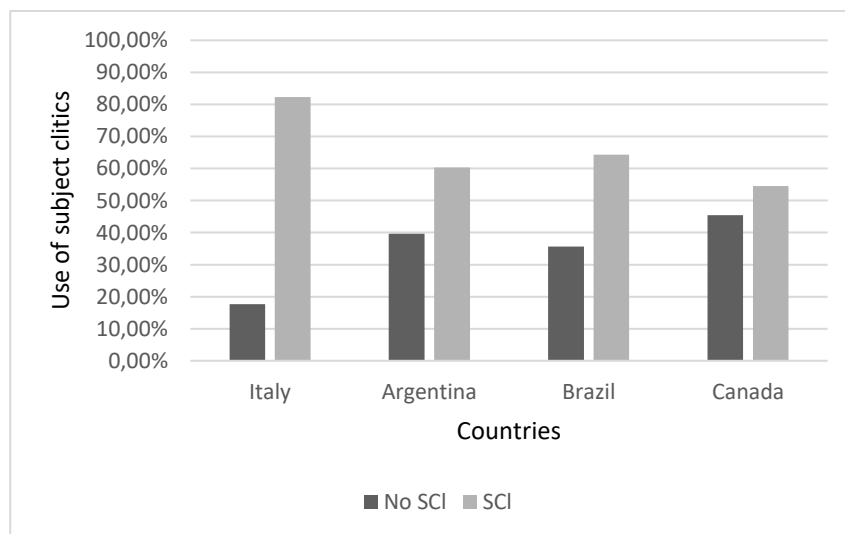
102 The syntax of subject pronouns in heritage languages

(2021). In general, there is a high inter-speaker variation¹⁹ in heritage Venetan, just as it is in heritage Friulian. The results of the spontaneous production task help clarify the results of the questionnaire, in particular with respect to information structure.

3.4.2.2. *Semi-guided production task*

The results of the semi-guided production task show that, in general, speakers of Venetan in Argentina, Brazil and Canada produce fewer subject clitics than speakers in Italy in contexts in which they are expected.

Figure 1. Use of Venetan subject clitics per country in expected contexts in different contact situations (countries). Sentences in total: 1308.

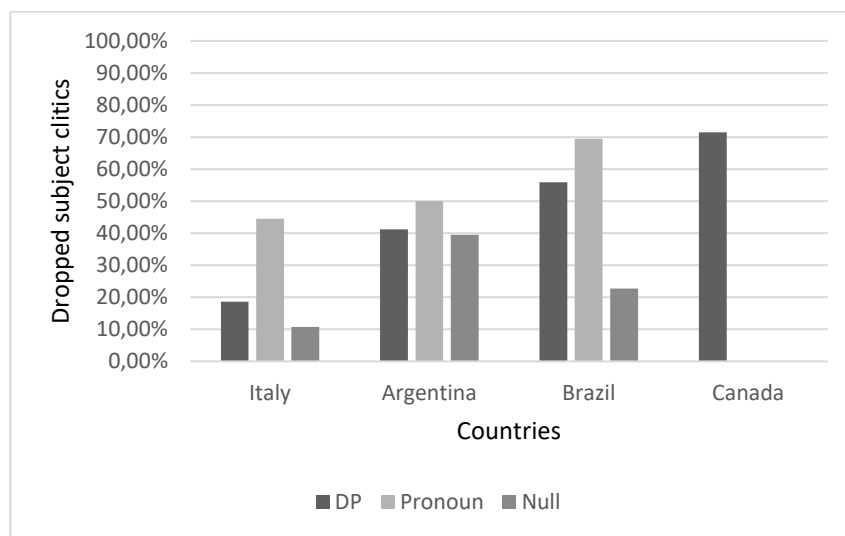


Such drop rate strengthens the conclusion that the distribution of subject clitics in Argentinian, Brazilian and Canadian Venetan is indeed different from the one of Italian Venetan.

¹⁹ As already discussed in Chapter 1, given the lack of a standard variety of Venetan and Friulian and the impossibility of receiving formal instruction in these languages, intra-speaker variation is also attested in all speakers, studies and contact situations discussed in this dissertation.

However, splitting the data according to whether the sentence, in which the subject clitic should appear, has a null or overt subject, highlights an effect of subject type both in heritage and Italian Venetan: subject clitics are most likely to be dropped in sentences featuring an overt pronominal subject, i.e. in pronoun-doubling contexts.

Figure 2. Percentage of dropped subject clitics per contact situation (country) in expected contexts with overt and null subjects. Sentences in total: 1308.



In fact, subject clitics are most likely to be dropped in contexts in which they are expected to double a pronominal subject also in Italian Venetan²⁰. The Canadian informant did not produce any sentence with a strong pronoun, so this piece of information is missing. In the case of lexical DP subjects, the drop rate is low in Italy while it is high in Argentina and very high in Brazil and Canada. In the case of null subjects, the Canadian informant never dropped subject clitics, while Italian and Brazilian informants dropped subject clitics considerably less with null subjects than with overt ones; the case of Argentina is yet different, as subject clitics are often

²⁰ The fact that Venetan subject clitics display a pronominal behaviour already in homeland varieties, in that they can be dropped in doubling contexts with lexical DP subjects, was first noticed by Benincà (1983, 1994). See section 3.1. in this respect.

104 The syntax of subject pronouns in heritage languages

dropped when no other overt subject is realised too; the Canadian informant never dropped subject clitics when no other overt subject is realised.

These data support what had already emerged from the forced choice task, in that they highlight the fact that varieties of Venetan spoken in the Americas generally allow for a higher rate of drop of subject clitics with respect to the Italian Venetan variety described in literature.

At this point, however, both in the case of the preference task and of the production it is not clear what factors trigger the higher probability of subject clitic drop in heritage Venetan. The study on heritage Friulian subject clitics (Frasson *et al.*, 2021) shows that subject clitics are most likely to be dropped in conditions of topic continuation, a behaviour that strengthens the idea that subject clitics allow for a pronominal distribution in heritage varieties. This is true also for heritage Venetan: as shown by the results from the spontaneous production task in the present study, the drop of subject clitics in heritage Venetan is strictly dependent on requirements of the syntax-discourse interface. This behaviour hence supports both the idea that they indeed display pronominal properties, as shown in Frasson (2021) for Brazilian Venetan and confirmed for southern Italo-Romance varieties in Chapter 2.

Frasson *et al.* (2021) showed that Argentinian and Brazilian Friulian subject clitics display a pronominal behaviour that emerges in specific discourse-related conditions: subject clitics can be dropped in topic continuation, as illustrated in (26).

(26) Argentinian Friulian (Frasson *et al.*, 2021)

a. SHIFT **I_j** ai tacat fevelà furlan, [...]

I.SCL have.1SG started speak.inf Friulian

b. CONT [...] dop **pro_j** ai vut ancje la fortune di sposà une fie di furlans.

then pro have.1SG had too the fortune of marry.INF a daughter of Friulians

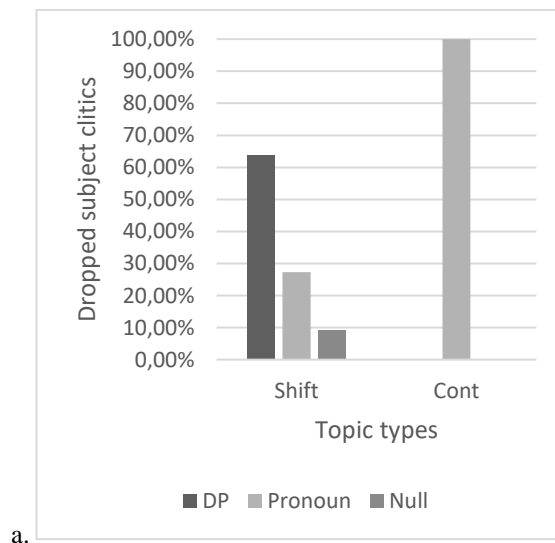
‘I started to speak Friulian, then I was lucky enough to marry a Friulian descendent.’

Example (26) shows that the subject clitic is dropped when the referent does not change from the previous discourse.

The heritage Venetan data presented in this dissertation show a behaviour comparable to that of heritage Friulian. In the remainder of this chapter, it will be shown that the realisation of subject clitics is regulated by discourse-related factors similar to those involved in pronoun realisation also in heritage Venetan.

With respect to the variable ‘TOPIC’, different subject types display different behaviours with respect to the realisation of subject clitics in topic shift and continuation in both considered varieties. Unfortunately, the Venetan data from Italy and Canada are not sufficient to be included in this analysis, so only the pattern evidenced by the Argentinian and Brazilian data will be presented.

Figure 3. Dropped subject clitics in Argentinian (a) and Brazilian Venetan (b); continuation and shift. Sentences in total: 1308.



- b. CONT Dopo *pro*_j se ga criado in Italia,
 then pro REFL has grown in Italy
- c. CONT *pro*_j ga catà el moroso e se ga maridà.
 pro has taken the boyfriend and REFL has married
- ‘My grandmother was born in Germany. Then she grew up in Italy, she found a boyfriend and got married.’

In (27)a) the referent *la nonna* is introduced for the first time and it is not doubled with a subject clitic; in topic continuation *pro* appears in (27)b) and (27)c); crucially, no subject clitics are realised in this example.

However, Figure 3 shows that subject clitics can still appear in topic continuation if no other overt subjects appear. Consider, for instance, example (28):

(28) Brazilian Venetan

- a. **I** **tozatei**_j ndea scola.
 the children went.3PL school
- b. No **i**_j savea mia parlar el portugues.
 not they.SCL knew.3SG not speak the Portuguese
- ‘Children went to school. They could not speak Portuguese.’

Here, the referent *i tozatei* is introduced in (28)a) as the topic of the sentence. In (28)b), a subject clitic *i* appears, referring to the topic in (28)a). Notice that the subject clitic appears after the negation marker *no* in (28)b).

As already noticed in Frasson (2021), however, the opposite order with respect to negation is also found in Brazilian Venetan. Consider example (29):

(29) Brazilian Venetan

- a. Me pare parlea che **ghe**_j (=a so noni) rivea el giomal de la Italia.
 my father spoke that they.DAT (= to his grandparents) arrived the newspaper from the Italy
- b. **Mi**_k no savaria parlar la sità. (...)
 I not know speak the city

108 The syntax of subject pronouns in heritage languages

- c. El giornal, parchè ij no savea come iera.
the newspaper because they.SCL not knew how was
'My father_j said that they_y (=his grandparents) received a newspaper from Italy. (...) I_k cannot say which city. (...) The newspaper, because they_y did not know how it used to be.'

What emerges from this example is that subject clitics in heritage Venetan can be used for topic shift and in that case they behave as pronouns. The subject clitic *i* in (29)c refers to the speaker's grandparents; this referent was previously referred to in (29)a with the dative clitic *ghe*. Hence, in the context of topic shift, heritage Venetan subject clitics display properties of pronouns, while they can be either dropped or realised in a position adjacent to the finite verb in the context of topic continuation. The next section expands the analysis introduced in Chapter 2 to the case of subject clitics. It will be shown that [uR] plays an important role in the distribution and the interpretation of subject clitics and that it captures the different facts shown in this section:

- that subject clitics behave as pronouns in the context of topic shift;
- that they have to be adjacent to the verb in the context of topic continuation;
- that they can be dropped.

I propose an alternative analysis to the one of subject clitics as agreement markers, in order to capture the case of heritage Venetan and explain the difference with respect to Italian Venetan.

3.5. Subject clitics as pronouns

This section extends to subject clitics the analysis put forward in Chapter 2 on the internal structure of pronouns. This analysis dispenses with Cardinaletti and Starke's structural deficiency and captures different referential properties through the introduction of [uR], a discourse-related feature encoded in strong pronouns and not present in weak pronouns. All pronouns have the structure of DPs but only some of them encode a [uR] feature in D. This [uR] feature makes the pronoun referentially

specific enough to switch the reference to a non-salient discourse antecedent. In this approach, the behaviour of subject clitics is not taken to depend on their agreement-like nature, but, again, on a difference in their internal featural composition. It is this variation in the featural composition, in particular regarding the [uR] feature that triggers a distribution that is different from that of strong pronouns. The data presented in Section 3.4.2 show that subject clitics in heritage Venetan (and Friulian, as showed in Frasson *et al.*, 2021) can optionally encode [uR], triggering a pronominal behaviour and challenging the traditional analysis of subject clitics as markers of pure φ -agreement. An analysis of subject clitics as heads realised in the inflectional field implies that they are severely deficient (in Cardinaletti and Starke's terms). Section 3.3 showed that this approach is too strong and cannot capture various phenomena related to the realisation of subject clitics in Friulian and Venetan. It was further proposed that subject clitics, just like strong and weak pronouns, are DPs.

3.5.1. *Salience and discourse*

As showed in Section 3.4, subject clitics display a clear pronominal behaviour, as evidenced by their distribution. Consider once again example (29), repeated as (30):

(30) Brazilian Venetan

- a. Me pare parlea che **ghe_j** (=a so noni) rivea el giornal de la Italia.
 my father spoke that they.DAT (=to his grandparents) arrived the newspaper from the Italy
- b. Mi_k no savaria parlar la sità. (...)
 I not know speak the city
- c. El giornal, parchè **ij** no savea come iera.
 the newspaper because they.SCL not knew how was
 'My father_j said that they_y (=his grandparents) received a newspaper from Italy. (...)
 I_k cannot say which city. (...) The newspaper, because they_y did not know how it used to be.'

110 The syntax of subject pronouns in heritage languages

The subject clitic *i* in (30)c) refers to the speaker's grandparents; this referent is semi-active, as it was previously referred to in (30)a) with the dative clitic *ghe*, despite not being topical.

Recall that the notion of activation of referents was introduced by Chafe (1987): all concepts are in different states of activation at particular times. They can be active, semi-active or inactive. An active concept normally has the properties of a given topic (in the terms of Givón 1983); in accessibility terms, as proposed by Ariel (1990), active concepts are usually indicated by a marker of high accessibility. Conversely, an inactive concept is generally a new referent, therefore it is indicated by a marker of low accessibility. A semi-active concept is at an intermediate level of activation, between active and inactive concepts: it is only peripherally active and it is indicated by a marker of intermediate accessibility. In the case of example (30), the subject clitic in (30)c) is referring to a semi-active concept, hence it is used as a marker of intermediate accessibility. In this case, the subject clitic switches the reference to a non-salient antecedent and appears before preverbal negation *no*. The subject clitic can switch reference if it encodes a [uR] feature. This means that the subject clitic will be used as a regular pronoun, allowing it to move across sentential negation as a pronoun; the unexpected position of the subject clitic with respect to preverbal negation is captured, as will be shown in Section 3.5.2.

Thus, the considerations on accessibility made in Chapter 2 hold for subject clitics as much as for strong pronouns. Recall that Accessibility Theory proposes that speakers choose between different referring expressions to mark accessibility differences; in turn, different referring expressions mark different degrees of accessibility. An accessible antecedent is defined as a salient referent in the discourse. Chapter 2 showed that strong subject pronouns mark a lower degree of accessibility than null subjects. The same degree of accessibility can be marked by subject clitics as well. As a result, they will behave on par with strong pronouns, rather than (weak) *pro*.

Strong pronouns, as markers of a lower degree of accessibility, are to be preferred to weak pronouns for switch reference or obviation. In the approach adopted here, these properties are connected to the presence of [uR]: the higher informativity

and rigidity of a strong pronoun with respect to a weak pronoun follow precisely from the presence of [uR] on the former. Subject clitics, as showed in example (30), can obviate or switch reference; in that case they will encode [uR]. However, subject clitics can also lack [uR]. If that is the case, they will still feature a D head; however, the head will carry only a [D] feature, not encoding [uR]. Subject clitics are DPs anyway, and their interpretive properties will mirror those of *pro*. Consider the following example: the subject clitic in (31)b) refers to a very active concept, in that the antecedent *i tozatei* is the topic of the previous sentence (31)a). However, a subject clitic is overtly realised in (31)b), while, in the model put forward in Chapter 2, in this case the pronoun should be deleted at PF under feature-identity.

(31) Brazilian Venetan

a. **I tozatei_j** ndea scola.

the children went.3PL school

b. No **i_j** savea.3PL mia parlar el portugues.

not they.SCL knew not speak the Portuguese

‘Children went to school. They could not speak Portuguese.’

In this case subject clitics do not undergo PF deletion, unlike what would be expected in consistent null subject languages. I will return to this issue in Section 3.5.2.

Chapter 2 addressed the question of how salience is encoded on different discourse antecedents and what leads to the assignment of [uR]. It was proposed that the assignment of [uR] follows the same model described in Chapter 2 for strong pronouns. Context Scanning identifies the potential antecedents in previous discourse and organises them in a set of ranked discourse referents carrying different salience values. Recall that salience is a context-dependent choice function (Von Heusinger, 2000) which selects among different potential discourse antecedents d_x that belong to an ordered set $\mathcal{A} = \{d_1 \dots d_n\}$; each referring expression refers to one of such elements d_x . Elements in the set are selected by a Logophoric Centre in CP, intended as the centre of deixis, a universal property of every sentence (Bianchi 2003). If the Logophoric Centre selects the most salient element d_l from the set \mathcal{A} , there will be no

112 The syntax of subject pronouns in heritage languages

update in the salience structure. The selection of another element $d_{\neq i}$ from the set implies that the salience value of this element needs to be updated and that the reference has to be switched. The Logophoric Centre now licenses a referentially specific interpretation.

Following Adger and Ramchand (2005), it was further proposed that an assignment function is responsible for the interpretation of syntactic objects. The Logophoric Centre is assigned an interpretable [R] feature only in case of switch reference or obviation. The uninterpretable [uR] feature on the subject DP is checked against the interpretable [R] feature on the Logophoric Centre and is checked before transfer to LF. This checking happens according to the mechanism outlined in Wurmbrand (2010) and Zeijlstra (2012), which was referred to as upward (Reverse) Agree:

(32) (*Reverse*) Agree: α can Agree with β iff:

- a. α carries at least one uninterpretable feature and β carries a matching interpretable feature.
- b. β c-commands α .
- c. β is the closest goal to α .

[uR] is checked but not deleted (Chomsky 1995) and remains visible at PF, thus blocking the deletion of the pronoun under feature identity.

Going back to subject clitics of the type in (30)c), i.e. subject clitics that refer to a semi-active referent, also in this case [uR] on the subject clitic is checked against [R] on the Logophoric Centre and the derivation can proceed. If the subject clitic lacks [uR], there will be no feature checking: the φ -features of the subject clitic will suffice to identify the antecedent; the subject clitic will therefore refer to the most salient antecedent. If the subject clitic encodes [uR] but the Logophoric Centre does not, checking cannot happen and the derivation crashes. The distribution of [uR] in different pronominal forms in Venetan is summarised in (33).

(33)

Full pronoun_[uR]: - inactive or semiactive referent ($d_{\neq I}$)

- switch reference

Subject clitic_[uR]: - semiactive referent ($d_{\neq I}$)

- switch reference

Subject clitic_[+R]: - active referent ($d_{=I}$)

- topic continuity

*pro*_[+R]: - active referent ($d_{=I}$)

. topic continuity

Before moving to the syntactic derivation of subject clitics, a disclaimer is in order. It was proposed that subject clitics and strong pronouns can both encode [uR] and hence display the same interpretive properties. It is possible that a more finely nuanced system is at work at the level of discourse and that more discourse-related features can be assigned to different forms. A wider cross-linguistic study to ascertain whether more fine-grained interpretive distinctions are visible at the syntactic level and in the distribution of different types of pronouns is beyond the scope of this study.²¹

²¹ I am aware that the distribution of strong pronouns and subject clitics is not precisely the same, as discussed in Section 3.3. Recall that, unlike strong pronouns, weak pronouns and subject clitics cannot be dislocated in the left periphery of the sentence, uttered in isolation, coordinated or modified. This is true not only for Friulian and Venetan, but also for other languages that display some phonologically reduced pronominal forms. In Section 3.3 we discussed the case of Brazilian Portuguese and in Section 3.5.3.2 we will discuss French. More examples can be found, such as Dutch, as we will discuss in Chapter 4, and the German dialects discussed in Weiß (2015). However, as shown by heritage Venetan subject clitics, phonological deficiency does not necessarily imply interpretive or syntactic deficiency, challenging the notion of classes of pronouns presented in Cardinaletti and Starke (1999). See also Manzini (2014) and Pescarini (2020) for similar proposals against a rigid classification of pronouns.

3.5.2. *The syntax of subject clitics*

The setup of subject pronouns proposed in Chapter 2 for consistent null subject languages, combined with a simple post-syntactic operation, captures the derivation of subject clitics and accounts for the apparent counterexamples that previous analyses of subject clitics as agreement markers somehow eluded. Besides, this approach has the desired minimalist goal of reducing structure-building operations to instances of Merge; finally, this analysis to subject clitics proves to be more economical at the representational level, dispensing, for instance, with the cartography of functional projections available for subject clitics proposed in Poletto (2000), as discussed briefly in Section 3.3.

Let us start by noticing that the derivation of sentences containing subject clitics proceeds in a fashion parallel to that of sentences with strong pronouns. Consider the following examples:

(34) Argentinian Venetan

Ela ga ciamà me pare.
she have.3SG called my father
'She called my father.'

(35) Brazilian Venetan

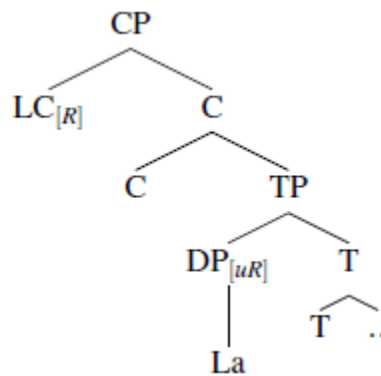
La ga ciamà el paron.
she.SCL have.3SG called the boss
'She called the boss.'

(34) shows that a strong pronoun does not have to be doubled by a subject clitic; this follows precisely from the idea that either the strong pronoun or the subject clitic checks the EPP feature on T, but not both. In (35), the subject clitic is merged with TP; it is not clear whether the subject clitic encodes [uR] or not, but it does not matter for the present discussion: recall that all pronouns are DPs and therefore behave

identically, from a syntactic viewpoint²². I will return to the interpretive and distributional difference between subject clitics and full DPs in Section 5.3.

In the analysis adopted in this dissertation, there is no need to claim that the strong pronoun in (34) and the subject clitic in (35) participate in different derivations. Both elements have the same internal structure. A sentence with a subject clitic is hence derived as shown in (36):

(36)



A discourse referent $d_{\neq l}$ was selected from the set \mathcal{A} of possible discourse antecedents via Context Scanning. The Logophoric Centre in CP is hence assigned an [R] feature and the [uR] feature on the subject *la* in Spec-TP can be checked against [R]. Once [uR] is checked, the pronoun gets to refer to $d_{\neq l}$.

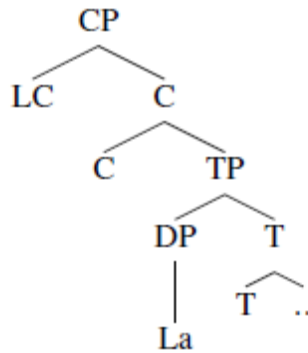
Consider now the possibility that the subject clitic *la* in (35) lacks [uR]. As shown in Section 3.2, this is the situation in which clitics are used as simple overt counterparts of *pro* in context of topic continuation. Since the subject clitic lacks [uR], it is not necessary for it to check against a context-defined [R], whether it exists or not.

²² The idea that subject clitics follow the same derivation path as strong pronouns and lexical DP subjects is not new. Poletto (1993) proposed that subject clitics are argumental and merged in Spec-vP, where they are assigned a thematic role.

116 The syntax of subject pronouns in heritage languages

However, if it lacks [uR], the subject clitic has a ϕ -featural setup which is exactly identical to that of T. Following Sheehan (2006), this means that it should undergo deletion under feature-identity for the overlapping ϕ -features. In this scenario, a subject clitics includes ϕ -features, while T includes the same ϕ -features and also verbal features; the features realised on the subject clitics are a subset of those of T. However, the subject clitic does not get deleted, yielding a derivation apparently identical to the one in (36).

(37)



In order to understand better what the presence or the absence of [uR] means for the realisation of subject clitics, I reconsider two phenomena already introduced in Section 3.3 with respect to Brazilian Venetan as described in Frasson (2021): preverbal negation and adverb interpolation. Section 3.5.3 showed that the distribution of subject clitics in these two contexts clarifies the role of [uR] and the implications it has in subject clitic realisation.

3.5.3. Why subject clitics do not get deleted at PF: adjacency to the verb

In the previous section we arrived at a conundrum, having to do with the non-deletion of a set of ϕ features (the subject clitic ones) which is identical to that of the T head. Specifically, a subject clitic carries morphological information on person, number and

gender that should be deleted by identity, according to Holmberg (2005), Sheehan (2006) and Roberts (2009), since they are also present on T. This proposal implies that *pro* in consistent null subject languages is simply a phonologically null counterpart of an overt weak pronoun, which deletes at PF under feature identity with finite T. If subject clitics lacking [uR] have the same structure and featural composition of *pro*, one may wonder why they do not undergo PF deletion. One observation that can help solve the issue regards the position of subject clitics. On the basis of their distributional properties (such as the position with respect to preverbal negation or the realisation in coordination, as discussed in Section 3.2), previous studies claim that subject clitics are agreement markers, heads realised on T. This analysis captures the fact that they can appear after the preverbal negation marker, that they are repeated in both conjuncts in coordinated structures and that adverbs cannot be interpolated between subject clitics and the verb. However, despite having the advantage of solving the problem of the position of subject clitics, this approach runs into serious problems, as shown in Section 3.3. Here, an alternative analysis for the peculiar distributional properties of subject clitics is put forward, which helps explaining why there is no PF deletion where it should be expected.

Kayne (1983) shows that French subject clitics, unlike object clitics, are in an argumental position and cliticise only at S-structure. Rizzi (1986) built on this assumption and defined French subject clitics as “phonological clitics” and subject clitics of the Venetan type as “syntactic clitics”, capturing the different distribution of the two subject clitic types in structural terms: phonological clitics are phrases, while syntactic clitics are agreement markers.

This approach runs into some problems. First of all, subject clitics do not double lexical DP subjects in Venetan, unless they are dislocated in the left periphery, as shown by Benincà (1994): the impossibility of doubling in this context depends on the fact that subject clitics are in Spec-TP just like ‘strong’ subject pronouns and are not located in the inflectional field. Besides, the data from heritage Venetan presented in Section 3.4.2 showed that subject clitics have the distribution of pronouns, rather than of obligatory agreement markers.

118 The syntax of subject pronouns in heritage languages

All these facts considered, I propose an alternative analysis that on the one hand captures the more autonomous behaviour of subject clitics in heritage varieties, simplifying, on the other hand, the whole discussion on the nature of subject clitics. In this system, the distribution of subject clitics in heritage and Italian Venetan is captured. In particular, this analysis has the advantage of explaining also the system of heritage Venetan varieties, in which subject clitics can be dropped under the right discourse-related conditions. Besides, in this model, the problem of explaining why Venetan requires two agreement markers on T (the subject clitic and the verbal morphology) simply does not emerge: subject clitics are pronouns, not agreement markers, and they cliticise at PF. Adjacency to the verb is intended as a PF process (Baker 1988, Marantz 1988, Bobaljik 1994) and represents an alternative to PF deletion, as illustrated below.

Baker (1988) proposed that cliticisation follows from the same consideration as another process: noun incorporation. Cliticization and incorporation involve both syntax and morphology in that it produces morphological units composed by elements that can be shown to be syntactically independent. The way this reasoning is connected to the case of subject clitics becomes immediately clear if we consider Baker's (1988) Condition of Morphological Identification²³:

(38) Condition of Morphological Identification (Baker 1988: 156)

If B is the NP position at the head of a chain, B bears a theta index at LF only if it bears a morphological index.

According to the condition in (38), only objects that are morphologically identified can be interpreted at LF. In other words, Baker's condition links the level of the semantic representation to that of the overt morphological realisation, in that an argument that stands in a relation with a theta-assigning item has to be

²³ For the sake of comparison, Roberts (2010) proposes that his approach on cliticisation and head movement is compatible with Baker (1988)'s approach on noun incorporation. I follow Roberts in this respect and I therefore propose that the condition of morphological identification proposed by Baker for noun incorporation also applies in the case of cliticisation.

morphologically realised. Baker proposes that the condition can be satisfied in four different ways: case assignment, agreement, adjacency and incorporation. The one that is relevant for subject clitics is adjacency. More specifically, there is a (post-syntactic) adjacency requirement on the realisation of subject clitics. In Section 3.3.1 it was shown that subject clitics are realised as the external argument of the verb and as such, they are assigned a theta-role by the verb (see also Poletto 1993 on the argumental status of subject clitics). So, subject clitics remain visible at PF when they are adjacent to the verb, satisfying the condition of morphological identification. If the requirement of adjacency at PF is violated, they will undergo deletion.

The relevance of this adjacency requirement is visible in the case of Friulian. Consider example (39). The subject clitic is realised when adjacent to the verb in (39)a) but is not realised when adjacency is impossible because of the presence of another element immediately preceding the verb (39)b).

(39) Friulian

- a. *(Al) à bussât lui.
 he.SCL have.3SG kissed him
- b. (*Al) lu à bussât.
 he.SCL him.OCL have.3SG kissed
 ‘He kissed him.’

Subject clitics do not cooccur with object clitics in Friulian; if an object clitic is realised in preverbal position, the subject clitic will be dropped (39)b). In other words, if the subject clitic is not adjacent to the verb, then it cannot be overtly realised. Subject clitic drop in this case does not depend on a property of clitic clusters in Friulian, but rather on the violation of the adjacency requirement: since object clitics in Friulian occur immediately before the verb, subject clitics cannot be adjacent to it and are therefore not overtly realised.

Summarising, two strategies are available. The first one, available mainly in heritage Venetan varieties, is a simple PF deletion under feature identity, as the one described by Sheehan (2006) and Roberts (2009). The second, available both in

120 The syntax of subject pronouns in heritage languages

heritage and Italian Venetan, is PF adjacency to the verb. This way, the realisation of subject clitics that lack [uR] is fully captured in the system presented in this study. In next section, I reconsider the cases, discussed in Section 3.3, of adverb interpolation and preverbal negation and show that this approach does not pose any problem for subject clitics encoding [uR].

3.5.3.1. *Adverb interpolation*

As shown in Section 3.3, temporal adverbs in Italian Venetan appear only postverbally:

(40) Italian Venetan

El parlea sempre cussì.

he.SCL spoke.3SG always like this

‘He used to speak like this.’

Since preverbal temporal adverbs are not allowed, it is impossible to check for potential restrictions on subject clitic realisation in this case.

Heritage varieties of Venetan, however, offer the opportunity to analyse in detail the conditions on the realisation of subject clitics, in that an adverb can appear between the subject clitic and the finite verb; a similar phenomenon of interpolation was described in Ledgeway and Lombardi (2005) for object clitics in southern Italo-Romance varieties. However, in the present approach, this phenomenon is captured, more simply, as an adverb appearing between the subject and the verb, possibly under the influence of Brazilian Portuguese and Argentinian Spanish, as discussed in Frasson (2021).

(41) Brazilian Venetan

El sempre parlea cusita.

he.SCL always spoke.3SG like this

‘He used to speak like this.’

In the case of (41), the subject clitic *el* behaves as a regular subject pronoun because of [uR]. The subject clitic will behave in all respects as a strong pronoun: at the interface with discourse, it can switch the reference to a non-salient discourse antecedent; at PF, it cannot undergo deletion because of [uR], which makes its featural composition non identical to that of T. According to Ledgeway and Lombardi (2005) and Cyrino (2013), the preverbal position of the adverb is a consequence of a different extent of movement of the verb and the pronoun. In the present analysis, this is compatible with the idea that the subject clitic is not realised on T.

Cyrino (2013) in particular argues that, in Brazilian Portuguese, T consists of two positions, a higher T1 and a lower T2 position, and that verbs can move only as high as T2, which is the reason why they appear after temporal adverbs. Verb movement in heritage Venetan will not be discussed here. However, Cyrino's analysis captures the difference between heritage and Italian varieties of Venetan very well.

(42)

[CP [TP1 *el*_[uR] [AdvP *sempre* [TP2 [T *dizea*]]...]]

The pronominal behaviour of the subject clitic is triggered by the presence of [uR], blocking PF deletion and not constraining the subject clitic to a position immediately adjacent to the verb.

Recall that subject clitics that lack [uR] are generally deleted at PF in heritage Venetan:

(43) Brazilian Venetan

pro *sempre* *parlea* *cusita*.
 pro always spoke.3SG like this
 'He used to speak like this.'

In this case, the subject clitic is not adjacent to the verb, hence it is deleted at PF under feature identity with T.

122 The syntax of subject pronouns in heritage languages

3.5.3.2. *The relative position of subject clitics and preverbal negation*

The second question addressed here regards the position of the subject clitic with respect to the preverbal negation marker *no*. Recall that, in Italian Venetan, subject clitics generally appear after negation (44).

(44) Italian Venetan

a. No i savea parlar.
not they.SCL knew.3PL speak.INF
'They could not speak.'

The relative position of the pronominal element with respect to preverbal negation has been traditionally used as a test to distinguish subject clitics of the Venetan type from subject clitics of the French type (or Brazilian Portuguese *cê*, as shown in Section 3.2.1), analysed as weak pronouns by Rizzi (1986) and Cardinaletti and Starke (1999). Unlike Venetan, French subject clitics precede preverbal negation:

(45) French

Ils ne savaient pas parler de certaines choses.
they not knew.3PL not speak.INF of certain things
'They did not know how to speak of certain things.'

The different position of the subject in the two languages with respect to preverbal negation is generally taken as evidence of the agreement-like nature of subject clitics in Venetan. However, this approach does not take into account the different nature of preverbal negation in the two languages. French preverbal negation marker *ne* is analysed as a head (Pollock 1989) or as an affix appearing inside the inflectional field (Ouhalla 1991), while the role of sentential negation in French is performed by postverbal *pas*. The subject clitic *ils* in (45), occupying Spec-TP, precedes preverbal negation *ne*.

The case of Venetan is different, as preverbal *no* is the sentential negation and it occupies the specifier of NegP, which dominates TP (Zanuttini 1998, Zeijlstra 2004).

(46)
[CP [NegP *no* [TP *i* [T *savea*] [vP ...]]]]

In a model that dispenses with pronoun classes, the order displayed in (46) does not depend on the agreement-like nature of subject clitics, but rather on the obligatory adjacency of the subject clitic to the verb and on NegP dominating TP. Therefore there is no need to assume that Venetan subject clitics, unlike their French counterparts, are realised in the inflection.

In heritage varieties of Venetan, subject clitics allow for the structure in (46), but, given the right interpretation, they can also appear before negation (47) or be dropped (48).

In the case of subject clitics preceding negation, their behaviour parallels the one discussed in Section 3.5.3.1 for adverb interpolation:

- (47) Argentinian Venetan
- a. I no savea parlar.
 they.SCL not knew.3SG speak.INF
 ‘They did not know how to speak.’
- b. [CP [TP1 *i_[uR]*] [NegP *no* [TP2 [T *savea*]]...]]

Building on Cyrino (2013), it is claimed that a subject clitic carrying [uR] moves across sentential negation to a higher TP1. I leave the discussion and the motivation of this movement for future research. Sentence (47)a) shows that the subject clitic is clearly not an agreement marker as it appears before preverbal negation just as regular

124 The syntax of subject pronouns in heritage languages

strong pronouns do in Venetan²⁴. The relevant difference with respect to (44) is that the subject clitic in (47)a encodes a [uR] feature that leads in turn to a distribution that parallels that of strong pronouns.

Finally, subject clitics lacking [uR] can be deleted in heritage Venetan, suggesting that a different strategy is at work at PF with respect to Italian Venetan.

(48) Argentinian Venetan

a. No savea parlar.

not knew.3PL speak.INF

‘They could not speak.’

b. [_{CP} [_{TP} ~~SCI~~ [_{NegP} *no* [_{TP2} [_T *savea*]]...]]]

The case represented by (48)a, in which the subject clitic is dropped, is analysed as PF deletion; recall that *pro* is a phonologically null counterpart of a weak pronoun that is deleted under feature-identity with T (48)b). This implies that the subject clitic is deleted only when it lacks [uR], hence it encodes only ϕ -features, which are a subset of the features realised on T.

With respect to the relative order of subject clitics and preverbal negation, it was shown that heritage Venetan allows for a pronominal distribution of subject clitics precisely because of the presence of [uR], an extra discourse-related feature that blocks further operations at PF. In heritage Venetan varieties, subject clitics lacking

²⁴ Recall that Poletto (2000) proposed that subject clitics can occupy different positions with respect to negation:

[_{FP1} invariable SCLs/deictic SCLs [_{NEGP} [_{FP2} number SCLs/person SCLs]]]

However, Poletto (2000) shows that in homeland varieties only a restricted class of invariable and deictic clitics, like the Venetan *a* clitic (Benincà 1994) can appear in the higher position. Heritage Venetan data show that also clitics encoding number and person can appear in this position.

[uR] can be deleted at PF under feature identity, as described in Chapter 2 for the alternation between overt and null subjects in southern Italo-Romance varieties; in Italian Venetan they are overtly realised and they must be adjacent to the verb.

3.5.4. Summary

This section extended the proposal that the distinction of different classes of subject pronouns is not necessary. This is particularly true in the case of subject clitics. It was shown that an analysis in terms of agreement markers is problematic as it does not fully capture the distribution of subject clitics (both in Italian Venetan and in heritage varieties) and is at the same time not desirable in minimalist terms: the analysis of subject clitics as DPs, rather than ϕ Ps is both derivationally and representationally more economic. Some peculiar distributional properties of subject clitics, such as their position with respect to preverbal negation, can be solved by assuming an adjacency constraint of the subject clitic to the verb, representing an alternative to subject deletion in other consistent null subject languages. More specifically, the distributional properties of subject clitics depend on an adjacency requirement that resembles the one described in Baker (1988) with respect to the Condition of Morphological Identification. Adjacency is generally obligatory in Italian Venetan in the case subject clitics lack [uR]. In heritage Venetan, PF deletion is a more common strategy, just as in consistent null subject languages.

3.6. Open questions

Two questions are not discussed in the present study: the realisation of second person singular subject clitics and interrogative inversion.

Second person was excluded from the study because there is no visible change on these subject clitics in heritage varieties; the analysis presented in this study focusses on the structures in which a change was detected. Section 3.6.1 briefly considers the problem of second person.

126 The syntax of subject pronouns in heritage languages

The question of interrogative inversion is not directly relevant for the model of the internal structure of pronouns presented in this study; besides, there is almost no trace of inverted subject clitics in heritage varieties of Venetan (and of Friulian). Section 3.6.2. briefly discusses this issue.

3.6.1. Second person

This study focusses on changes in the system of third person subject clitics. The reason why second person was excluded is very simple: there is no change in its realisation or distribution in heritage varieties²⁵.

That second person subject clitics in northern Italo-Romance varieties have a special role, crucially different from that of third person, was noticed by Vanelli (1998: 48), which proposed the following generalisation: *se una varietà fa un uso costante di almeno un pronome soggetto, questo è quello di seconda persona [if a variety makes constant use of at least one subject pronoun, this is the second person]*. In general, she noticed that all languages spoken in northern Italy have at least one subject clitic form, which is always the second person singular. This generalisation was described in Poletto (2000: 36) in terms of syntactic position: *the lowest subject clitic position is realised after the preverbal negative marker and must be repeated in coordination. It can be occupied only by person clitics, which behave in the same way object clitics do when the inflected verb (...) encodes only a [hearer] feature*. Recall that Poletto (2000) proposes that subject clitics are realised in different functional projections, with respect to the preverbal negative marker. In (49) the portion of structure following the negation is repeated:

(49) (adapted from Poletto 2000: 36)

[NegP [NumP SCI (3rd person) [HearerP **SCI (2nd person)** [Infl]]]]

²⁵ Recall that Venetan lacks subject clitics for first person. In view of this fact, first person could not be included in the study.

In Poletto's approach, second person subject clitics occupy HearerP, the lowest projection available for subject clitics, immediately before Infl. In Section 5, it was proposed that this model can be dispensed with in the approach adopted in this study. However, the idea that second person singular, unlike other subject clitics, parallels the behaviour of object clitics, as well as Vanelli's generalisation, are reflected also in the system of heritage Venetan and Friulian.

Firstly, they are the only subject clitics that are never dropped in heritage Friulian (50)a) and Venetan (50)b):

(50)

a. *(Tu) viodis.

b. *(Te) vedi.

you.SCL see.2SG

'You see.'

With respect to doubling, in Section 3.3 it was shown that it is constrained for third person in Venetan. However, second person subject clitics double strong pronominal subjects without exception in heritage Friulian (51)a) and Venetan (51)b):

(51)

a. Tu *(tu) sintis mior.

b. Ti *(te) senti meio.

you you.SCL hear.2SG better

'You hear better.'

Besides, discourse factors seem not to play any role on the realisation of second person singular. For instance, it is always obligatorily realised in topic continuation in the spontaneous production data discussed in Section 3.4.2:

128 The syntax of subject pronouns in heritage languages

(52) Brazilian Venetan

- a. SHIFT **Te** ciapi i spini del pino.
you.SCL take.2SG the twigs of.the pine tree
- b. CONT Coi spini **te** meti so i pignoni,
with.the twigs you.SCL put.2SG on the pine nuts
- c. CONT dopo **te** i magni
then you.SCL them.OCL eat.2SG

‘You take the twigs of the pine tree. With the twigs you cook the pine nuts (lit.: ‘you put the pine nuts on’), then you eat them.

It is possible to think of the realisation of second person singular again in terms of discourse-related features: second person is very salient in the discourse, hence it does not require an extra discourse-feature in order to be referentially specific. It is plausible to assume that, diachronically, second person singular grammaticalised on the verb, becoming a real agreement marker in the terms of Brandi and Cordin (1989) and Rizzi (1986). This would capture the distribution of second person subject clitics in modern Italian Venetan and also the fact that, unlike third person subject clitics, in heritage Venetan they do not display any pronominal properties. This analysis would not contradict Poletto (2000), who also showed that second person subject clitics belong to a different projection. I leave the discussion of second person for future investigation.

3.6.2. *Interrogative inversion*

Another issue that is not discussed in the present study is subject clitic interrogative inversion. In Venetan and Friulian, as well as other languages of northern Italy, subject clitics appear in preverbal position in declarative sentences (they are proclitic) and in postverbal position in interrogative sentences (they are enclitic). There are phonological and distributional differences between the preverbal and the postverbal subject clitics; compare the paradigm of the verb *to be* in Venetan and Friulian, with respect to proclitic and enclitic subjects:

Table 10. Preverbal and postverbal subject clitics in Friulian and Venetan.

	Friulian		Venetan	
	Preverbal	Postverbal	Preverbal	Postverbal
1sg	i/o soi	soi-o?	so	so-i?
2sg	tu	ses-tu?	te si	si-tu?
3sg	M: al è F: e je	M: is-al? F: is-e?	M: el ze F: la ze	M: ze-lo? F: ze-la?
1pl	i/o sin	sin-o?	semo	semo-i?
2pl	i/o ses	ses-o?	si	si-u?
3pl	a son	son-o?	M: i ze F: le ze	M: ze-i F: ze-le

The two paradigms are quite different; Venetan does not have proclitics for first person singular and plural and for second person plural; besides, enclitics are often different from proclitics. Poletto (2000) proposes that the difference follows from the fact that enclitics are not the same type of elements as proclitics. The inverted subject clitics are analysed as morphological markers of interrogative agreement, realised when the verb moves to CP. An alternative approach is presented in Cardinaletti and Repetti (2008): analysing the phonological processes at work, they propose that proclitic and enclitic forms are the same lexical item. It should be noticed that, also in Italian varieties of Venetan and Friulian, inversion is gradually disappearing and proclitic subjects are normally used also in interrogative contexts (Poletto, 2000; Casalicchio and Frasson, 2018).

However, this gradual process is not attested in heritage varieties of Venetan (and Friulian), in which subject clitic inversion is virtually absent. Consider the following examples of heritage varieties in (53)a)-(56)a), compared to the Italian varieties version in (53)b)-(56)b):

130 The syntax of subject pronouns in heritage languages

(53) Venetan

a. Te si vegnesto solo?
you.SCL be.2SG come alone

b. Si=tu vegnesto solo?
be=2SG.you.SCL come alone

‘Did you come alone?’

(54) Venetan

a. El pol comperar le medicine?
he.SCL can.3SG buy.INF the medicines

b. Po=lo comperar.INF le medicine?
can.3SG=he.SCL buy the medicines

‘Can you (form., lit.: ‘he’) buy the medicines?’

(55) Venetan

a. I guadagna schei?
they.SCL earn.3PL money

b. Guadagne=i schei?
earn.3PL=they.SCL money

‘Do they earn money?’

(56) Friulian

a. Al è già lat vie?
he.SCL be.3SG already gone away

b. Is=al già lat vie?
be.23SG=he.SCL already gone away

‘Has he already left?’

In the heritage examples (53)a)-(56)a), subject clitic inversion should be expected, but it does not take place. It is possible to assume that the peculiar conditions of contact in which these varieties are spoken caused an acceleration of the disappearance of

inverted clitics in heritage varieties, but it is not possible to propose an analysis of change in this case, in the absence of data showing the distribution of inverted subject clitics. Hence, in the study the discussion is limited to proclitic subjects, leaving the investigation of enclitics and the reason behind their absence in heritage varieties for future research.

3.7. Conclusion

This chapter showed that subject clitics in Italian Venetan generally encode only φ -features, while subject clitics in heritage Venetan can optionally encode [uR], hence allowing for a pronominal behaviour. Consistently with the approach on null and overt subjects proposed in Chapter 2, [uR] is associated with the D-layer that is traditionally taken to be lacked by subject clitics (as predicted by Cardinaletti and Starke 1999). Therefore, the behaviour displayed by subject clitics in heritage Venetan (and Friulian, as described in Frasson *et al.*, 2021) cannot be explained if their analysis as markers of φ -agreement (as proposed specifically for subject clitics in Rizzi 1986; Brandi and Cordin 1989; Benincà, 1994; Poletto 1993, 2000) is maintained. It was proposed that subject clitics, just as strong pronouns and null *pro*, are DPs and that the different behaviour displayed by subject clitics in heritage and Italian varieties of Venetan and Friulian depends on the presence of [uR]. The proposal is supported by the fact that an analysis of subject clitics as φ -heads cannot capture several distributional facts in Italian varieties of Friulian and Venetan too. The agreement-like behaviour of subject clitics in these varieties depends on an adjacency requirement that triggers cliticisation of the pronoun at PF, just as in the case of French subject clitics, as proposed by Kayne (1983). This adjacency requirement represents an alternative to the PF deletion under feature identity that was described in Chapter 2: in both cases, the φ -composition of the pronoun is a subset of the feature composition of T. In Italian Venetan and Friulian, the pronoun generally remains overtly realised if it is adjacent to the verb. In heritage Venetan and Friulian, when subject clitics are not assigned [uR], they are generally deleted at PF under feature-

132 The syntax of subject pronouns in heritage languages

identity, just as in other consistent null subject languages (Sheehan 2006, Roberts 2009).

In conclusion, also in the case of subject clitics, there is no need to resort to structural deficiency and classes of pronouns: their behaviour is captured in syntax by the introduction of a discourse-related feature [uR], which explains the different properties displayed by subject clitics in heritage varieties of Friulian and Venetan.

Chapter 4. The position of the antecedent of pronominal, clitic and null subjects

4.1. Introduction

This chapter presents a study of the distribution of different pronominal forms in three varieties of Venetan: Argentinian, Brazilian and Italian Venetan. Specifically, I focus on the distribution of subject clitics and compare it to that of strong pronouns and null subjects. The approach put forward in Chapters 2 and 3 is used to test of the Position of the Antecedent Hypothesis (Carminati, 2002) and to identify the factors that regulate the distribution of different pronominal forms.

The interpretation of subject clitics depends on the presence of a discourse-related feature [uR]. This feature is optionally realised on subject clitics in order to switch the reference to a non-salient discourse antecedent, a possibility that was not investigated in previous studies on subject clitics; switch reference is traditionally thought to be a property of strong subject pronouns, in the terms of Cardinaletti and Starke (1999).

The Position of the Antecedent Hypothesis (1) makes a prediction on the antecedent selection preferences of overt and null subjects:

(1) Position of the Antecedent Hypothesis (PAH; Carminati 2002)

The null pronoun prefers an antecedent which is in the Spec-IP position, while the overt pronoun prefers an antecedent which is not in the Spec-IP position.

Crucial to the PAH is the notion of “preference”, a concept that allows to capture the variation in the realisation of the different pronominal forms. It will be shown that the

134 The syntax of subject pronouns in heritage languages

approach to [uR] accounts for the “preferences” of overt and null subjects without constraining their distribution, unlike what is predicted in the model in which pronouns belong to different classes distinguished by their level of structural deficiency (Cardinaletti and Starke, 1999).

In Chapter 3, it was shown that this approach to the internal structure of pronouns is particularly useful to account for a variety of phenomena related to subject clitics that otherwise could not be explained, especially in heritage varieties. Here it is further shown that testing the PAH on subject clitics can provide a conclusive answer to the proposal made in Chapter 3 on the nature and the interpretive properties of subject clitics. Recall that in the proposed approach the distinction of classes of pronouns, as defined in Cardinaletti and Starke (1999) and, similarly, in Déchaine and Wiltschko (2002) is abandoned and replaced by a simple featural distinction in the D-layers of pronouns: pronouns that encode [uR] in D correspond to strong pronouns (according to Cardinaletti and Starke’s classification) and are generally used to switch reference and obviate (see also Frascarelli 2007); pronouns that do not encode [uR] correspond to weak pronouns and are generally used to refer to salient discourse antecedents and can undergo PF deletion (as in the case of *pro*) or cliticisation (as in the case of subject clitics).

Therefore, testing the PAH on subject clitics allows to provide a final answer to the question regarding the nature of subject clitics: if subject clitics pattern with strong pronouns, in that they prefer a non-subject antecedent which is not in Spec-TP, then they cannot be analysed as agreement markers. The preference for a subject antecedent in Spec-TP, implying a behaviour comparable to that of *pro*, does not contradict the proposal made in this study, however, in that the model captures the possibility for subject clitics to lack [uR] and still be overtly realised under adjacency with finite T²⁶. Recall that the difference between the two possible interpretations of subject clitics depends on the fact that they can, but not have to, encode the discourse-related feature [uR] in the D-layer.

²⁶ As already showed in Chapter (3), the idea of PF cliticisation was already proposed by Kayne (1983) and Rizzi (1986) for French subject pronouns.

The chapter is organised as follows: Section 4.2 presents three studies that, building on Carminati's PAH, show that discourse properties are as important as syntactic properties in the interpretation of different pronominal forms; Section 4.3 introduces the cases under analysis and presents the results of the study: the PAH was tested in three varieties of the same language, spoken in different contact situations: Argentinian Venetan, Brazilian Venetan and Italian Venetan; in Section 4.4 it is shown that the approach to [uR] can explain the different distribution of pronominal forms found in different varieties of Venetan; Section 4.5 discusses the role of contact in shaping the different patterns displayed in the varieties under analysis; Section 4.6 concludes the chapter.

4.2. Pronoun interpretation and the position of the antecedent

4.2.1. Previous studies on pronoun interpretation

The idea that overt and null subjects have different formal properties that are reflected in the way they select their referent received a lot of attention in generative studies. One of the most widely accepted proposals is Montalbetti's (1984) Overt Pronoun Constraint, according to which a null subject can be construed as a bound variable, while an overt subject pronoun cannot. Montalbetti's analysis refers specifically to the antecedence relation between a null or overt pronoun and a quantified subject²⁷:

(2) Spanish (Montalbetti 1984: 82)

- a. **Muchos estudiantes**_j creen que **ellos**^{*j/k} son inteligentes.
many students think.3PL that they be.3PL intelligent
- b. **Muchos estudiantes**_j creen que *pro*_{j/k} son inteligentes.
many students think.3PL that pro be.3PL intelligent

²⁷ This restriction on the interpretation of overt subjects indirectly supports the idea proposed in Chapter 2 that it is the overt subject, and not *pro* that requires some sort of licensing, *contra* Frascarelli (2007).

136 The syntax of subject pronouns in heritage languages

‘Many students think that they are intelligent.’

The overt pronoun in (2)a) cannot be construed as a bound variable, while the null pronoun in (2)b) can.

The Overt Pronoun Constraint, however, does not capture the interpretive contrast between overt and null subjects when the antecedent is not a quantified subject:

(3) Spanish (Montalbetti 1984; 122)

a. **Juan_j** cree que **él_{j/k}** es inteligente.

Juan think.3SG that he be.3SG intelligent

b. **Juan_j** cree que **pro_{j/k}** es inteligente.

Juan think.3SG that pro be.3SG intelligent

‘John thinks that he is intelligent.’

Montalbetti proposes that pragmatic considerations may lead the null subject in the embedded clause in (3)b) to corefer with the subject of the matrix clause; however, also the overt subject in (3)a) allows for this interpretation.

The study of the interpretation of Italian pronouns with respect to non-quantified antecedents carried out in Carminati (2002) follows precisely from syntactic and pragmatic considerations, showing that different pronominal forms have different and specialised functions also in contexts like (2). Carminati’s hypothesis is that Italian null and overt pronouns have distinct antecedent biases in that they select antecedents in different syntactic positions. However, the relevant syntactic positions are also linked to a notion of salience: the preverbal subject (Spec-IP, the highest specifier) is more salient than a postverbal direct or indirect object. Recall that the notion of salience structure proposed in Chapter 2 follows from very similar considerations: the first (and highest) element of a set of ranked discourse antecedents has the highest salience value; lower elements of the set have a lower salience value.

The position of the antecedent of pronominal, clitic and null subjects 137

A null subject will therefore select the most salient among the possible antecedents (the preverbal subject), while an overt pronoun will most likely select a less salient antecedent (the postverbal object):

(4) Italian (Carminati 2002: 363)

- a. **Roberto**_j ha insultato Ugo_k quando **pro**_j era ubriaco.
Roberto have.3SG insulted Ugo when pro was.3SG drunk
- b. Roberto_j ha insultato **Ugo**_k quando **lui**_k era ubriaco.
Roberto have.3SG insulted Ugo when he was.3SG drunk
'Robert insulted Hugh when he was drunk.'

The null subject in (4)a) refers to the subject of the matrix clause (*Roberto*), which has the highest salience value; the overt pronoun in (4)b) refers to the object of the matrix clause (*Ugo*).

This idea is formalised as the Position of the Antecedent Hypothesis (PAH), repeated in (5):

(5) Position of the Antecedent Hypothesis (PAH)

The null pronoun prefers an antecedent which is in the Spec-IP position, while the overt pronoun prefers an antecedent which is not in the Spec-IP position.

Carminati (2002) assumes that Spec-IP is linked to the notion of salience presented in Ariel's (1990) Accessibility Theory: different referring expressions mark different degrees of accessibility, defined in terms of prominence and salience. All referring expressions are therefore organised in an Accessibility hierarchy:

138 The syntax of subject pronouns in heritage languages

(6) Accessibility hierarchy (adapted from Ariel 1990)

Accessibility	'	- Full name
		- Demonstrative
	↑	- Stressed pronoun
		- Unstressed pronoun
	+	- Null

Null subjects, as markers of a high degree of accessibility, retrieve the most accessible or salient antecedent: the preverbal subject; overt subject pronouns, as markers of a lower degree of accessibility, retrieve a less salient antecedent.

A similar proposal is carried out by Mayol (2009) in her study of the distribution of Catalan null and overt subject pronouns. Null subjects in Catalan preferably select the subject of the preceding sentence as an antecedent, while overt pronouns prefer the postverbal non-subject.

(7) Catalan (Mayol 2009: 128)

- a. **La Marta**_j escrivia sovint a la Raquel_k. **pro**_j vivia als Estats Units.
the Marta wrote.3SG often to the Raquel pro lived.3SG in United States
- b. La Marta_j escrivia sovint a **la Raquel**_k. **Ella**_k vivia als Estats Units.
the Marta wrote.3SG often to the Raquel she lived.3SG in United States
'Martha often wrote to Rachel. She lived in the United States.'

In order to explain the role of discourse and pragmatics in antecedent selection, Kaiser and Trueswell (2008) propose a multiple-factor approach, in which multiple syntactic and discourse-related constraints contribute to a referent's salience and to the selection of an antecedent; different referring expressions can exhibit different sensitivity to different factors. In their study on Finnish third person anaphors, Kaiser and Trueswell (2008) show that the pronoun generally selects the syntactic subject as an antecedent, either in preverbal or in postverbal position; the demonstrative, conversely, shows a preference for the antecedent in postverbal

position, generally referring to a newly-introduced referent, but this preference is significant only when the postverbal²⁸ element is an object. In other words, the pronoun is sensitive mainly to the syntactic role of the antecedent, regardless of its pragmatic role, while the demonstrative is sensitive to both syntactic and pragmatic factors.

Likewise, Van Kampen (2012) shows that the distribution of Dutch anaphoric forms is constrained by discourse-related factors:

(8) Dutch (Van Kampen 2012: 68)

a. De detectives_j waarschuwden de secretaresses_k.

the detectives warned.3PL the secretaries

b. **Ze**_{j/k} / **Die**_{*j/k} verlieten het gebouw.

they those left.3PL the building.

‘The detectives warned the secretaries. They left the building.’

(8)b) shows that the antecedent selection of the demonstrative *die* is constrained by discourse-related factors, as it is restricted to the first preceding focus. The subject pronoun *ze* is not subject to such restriction.

What emerges from the various studies just presented (Kaiser and Trueswell 2008; van Kampen 2012; Mayol 2009; see also Samek-Lodovici 1996), is that the preferences in the selection of the antecedent by different pronominal forms may depend on different combinations of syntactic and pragmatic factors. The account of the internal structure of pronouns and the distribution of [uR] captures these interface effects.

²⁸ I discuss the case of postverbal subjects in Venetan in chapter 5.

4.2.2. *The [uR] feature and antecedent selection*

The combination of [uR] and the internal structure of pronouns with the PAH (Carminati 2002) contributes explaining the distribution of different pronominal forms and the selection of different antecedents.

According to the PAH, a null subject prefers a preverbal subject as an antecedent. Given that subjects are generally topics, this position has a high salience value in the sentence. Therefore, it was shown that the preference of embedded *pro* for the matrix subject antecedent is easily explained. Aside from Italian and Catalan, the predictions made by the PAH hold in the Italo-Romance varieties presented in Chapter 2, such as Calabrian:

(9) Calabrian

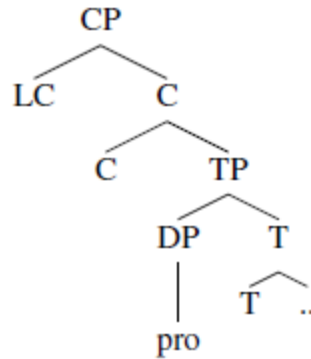
Marcu_j scrivìa sempri a Luca_k quannu **pro_j** stava mmali.

Mark wrote.3SG always to Luke when pro was.3SG sick.

‘Mark always used to write to Luke when he was ill.’

Following Bianchi (2003), it is assumed that subordinate clauses, just as matrix clauses, have a Logophoric Centre in CP. In this case, the discourse referent *Marcu* was selected from the set Δ of possible discourse antecedents. This referent has the highest salience value, so the Logophoric Centre is not assigned an [R] feature, in that the salience structure does not need to be updated in the subordinate. The lack of [R] in the Logophoric Centre does not pose any problems for null *pro*: weak pronouns like *pro* (10) do not have an uninterpretable [uR] feature to be checked; in consistent null subject languages, the pronoun is deleted at PF under feature-identity with finite T, as shown by Sheehan (2006) and Roberts (2009).

(10)



The case of overt pronouns is different, in that, according to the PAH, they prefer a less salient antecedent.

(11) Calabrian

Marcu_j scrivìa sempri a **Luca**_k quannu **iddru**_k stava mmali.

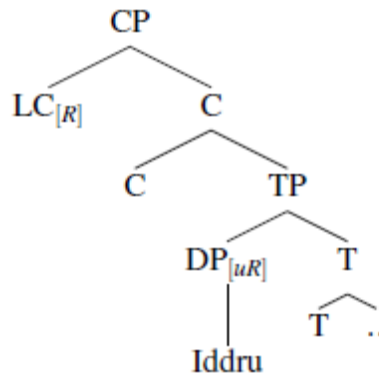
Mark wrote.3SG always to Luke when he was.3SG sick.

‘Mark always used to write to Luke when he was ill.’

In this case, the salience structure needs to be updated, in that reference is switched to the non-salient discourse antecedent *Luca*. The Logophoric Centre in Spec-CP is assigned [R]; a strong pronoun *iddru* is realised in Spec-TP, indicating that it does not refer to the most salient antecedent. In this case, the DP subject encodes a [uR] feature that needs to be checked against the [R] feature on the Logophoric Centre and deleted at LF; this operation was defined as *R-checking*, following Wurmbrand (2010) and Zeijlstra (2012)’s analyses of upward agreement. The presence of [uR], however, makes the ϕ -composition of the pronoun different from that of finite T: as a consequence, PF deletion under feature-identity is impossible.

142 The syntax of subject pronouns in heritage languages

(12)



Languages that display more than one type of overt pronoun follow the same principle. This is the case of Dutch, as shown in Van Kampen (2012). The pronoun *die* has a restricted interpretation, in that it can only refer to the antecedent encoding “new” information, the focus:

(13) Dutch (Van Kampen 2012: 68)

De detectives_j waarschuwden **de secretaresses**_k. **Die**_k verlieten het gebouw.
 the detectives warned.3PL the secretaries they left.3PL the building.
 ‘The detectives warned the secretaries. They left the building.’

Die is a strong pronoun and it displays a parallel behaviour to that of *iddru* in Calabrian. The distribution of strong pronouns, in the present approach, depends on the necessity of switching the reference to a non-salient discourse antecedent. I propose that *die* encodes a [uR] feature²⁹.

Dutch, however, is different from Calabrian in that it is a non-null subject language: where Calabrian allows for PF deletion of a pronoun lacking [uR], Dutch does not. In this case, the pronoun *ze* is realised instead. Interestingly, this type of pronoun has been described as a ‘weak’ pronoun (den Besten, 1977; Kaiser, 2011; see

²⁹ See also Patel-Grosz and Grosz (2017) for an alternative analysis.

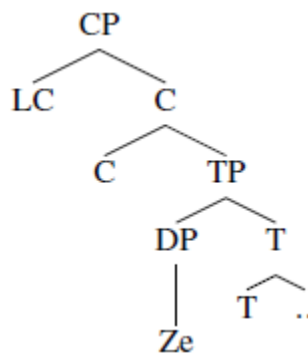
also Haegeman, 1990 for West Flemish) in Cardinaletti and Starke's (1999) model or as a ϕ P in Déchaine and Wiltschko's (2002) terms.

(14) Dutch (Van Kampen 2012: 68)

De detectives_j waarschuwden de secretaresses_k. **Ze**_{j/k} verlieten het gebouw.
 the detectives warned.3PL the secretaries they left.3PL the building.
 'The detectives warned the secretaries. They left the building.'

In the present analysis, *ze* has the same interpretation as subject clitics in Venetan, in that it can, but it does not have to, switch the reference to a non-salient discourse antecedent. In the case in which it does not switch the reference, the pronoun lacks [uR].

(15)



At PF, the pronoun lacking [uR] must be adjacent to the verb in order to be realised, to satisfy Baker's (1988) Condition of Morphological Identification, as showed in Chapter 3. This analysis of subject clitics in Venetan is supported by their interpretation with respect to the predictions made by the PAH, similarly to what happens for Dutch pronouns.

4.2.3. Summary

The studies presented in Section 4.2.1 suggest that, while the interpretation of null and overt pronouns involves a combination of syntactic and discourse-related or pragmatic factors. The role of the different factors may vary across languages, a fact that is easily captured by the introduction of [uR].

So far, only the cases in which the most salient discourse antecedent is in Spec-TP were discussed. Of course, this is not always the case. I will return to this point in Chapter 5.

4.3. The study

The study presented in this section aims to define the role of [uR] in the antecedent selection of subject clitics, with respect to strong pronouns and *pro*. A relevant property of the PAH is the notion of “preference” in the selection of an antecedent. Therefore, despite aiming to provide a tool for the theoretical approach to the encoding of [uR], the description of the data is supported by a statistical analysis following the methodology used in Carminati (2002): One-Way ANOVAs were carried out on the collected data presented below, in order to identify tendencies and significance of the results. The need for a statistical analysis stems from the fact that the interpretation of different types of subject pronouns cannot be defined in terms of grammaticality or ungrammaticality, but rather in terms of preference. The introduction of [uR] captures the concept of preference, as discussed in Section 4.4.3.

4.3.1. Research question and hypotheses

The studies presented in Section 4.2 deal with varieties that display systems with two competing pronominal forms (null and overt pronoun in the case of Italian, Spanish and Catalan; pronoun and demonstrative in the case of Dutch and Finnish). In Chapter 3 it was shown that languages such as Venetan have a more articulated ternary system, including null, clitic and pronominal subjects. It was proposed that subject clitics are

better analysed as pronouns, against the literature on the topic that generally treats subject clitics as agreement markers with a role comparable to that of verbal morphology. The pronominal distribution of subject clitics emerged especially in heritage varieties, in the contexts of topic shift and continuation: subject clitics tend to be realised in the former context but not in the latter and, when realised, they allow for a distribution that is not captured by the analysis of subject clitics as agreement heads, as shown in (17). The position of the subject clitic before the negation signals that they can be separated by the verb and therefore are not realised on T, but rather in Spec-TP, like the subject pronoun in (16). However, subject clitics can be adjacent to the verb in heritage Venetan and they have to be obligatory adjacent to the verb in Italian Venetan (18); finally, subject clitics can be dropped in heritage Venetan in topic continuity (19), while subject clitic drop is not frequent in Italian Venetan.

(16) Venetan (Caxias do Sul, Brazil)

Lori no catea nisuni.
they not found.3PL anyone
'They did not find anyone.'

(17) Venetan (Bento Gonçalves, Brazil)

I no savea come iera.
they.SCL not knew.3PL how was.3SG
'They did not know how it used to be.'

(18) Venetan (Caxias do Sul, Brazil)

No **i** savea mia parlar el portugues.
not they.SCL knew.3SG not speak. INF the Portuguese
'They could not speak Portuguese.'

(19) Venetan (Bento Gonçalves, Brazil)

pro no savea mia parlare el portoghese.
pro not knew.3PL not speak the Portuguese

‘They could not speak Portuguese.’

The general hypothesis is that subject clitics that allow for a more autonomous interpretation, like that in (17) have the same syntactic behaviour as strong pronouns, illustrated in (16): they encode an [uR] feature and are specific enough to switch the reference to a non-salient discourse antecedent. Since the featural composition of the subject clitic and the pronoun is not identical to that of finite T, they do not delete at PF, as showed by Roberts (2009) and Sheehan (2006). The subject clitic in (18) lacks [uR], just as null *pro* in (19): it corefers with the most salient discourse antecedent available. While the pronoun in (16) is deleted at PF under feature-identity with T, the subject clitic in (18) must satisfy an adjacency condition: the subject clitic is morphologically realised only when it is adjacent to finite T.

With respect to the PAH, subject clitics could therefore allow for both an interpretation akin to that of *pro* and to one like that of strong overt pronouns. In the first case, when subject clitics lack [uR], they are expected to corefer with the most salient antecedent in Spec-TP; when they encode [uR], subject clitics are expected to corefer with the less salient antecedent that is not in Spec-TP. This chapter aims to test the interpretation of subject clitics in Venetan, comparing it to the interpretation of null subjects and tonic pronouns with respect to the predictions made for them by the PAH (Carminati 2002).

Given these assumptions, three possible scenarios are predicted for the interpretation of subject clitics:

i. Subject clitics always pattern with null subjects

In this scenario, subject clitics pattern with null subjects in the selection of a preverbal subject antecedent; this supports the idea that the subject clitic needs to be adjacent to the verb. Note that this hypothesis is, in principle, also compliant with the various proposals that have been made regarding the agreement-like nature of subject clitics (see in particular Rizzi 1986; Brandi and Cordin 1989; Benincà 1994; Poletto 1993, 2000) which, however, were already excluded based on the fact that even in homeland varieties such an analysis runs into some problems.

ii. Subject clitics always pattern with overt subjects

In this scenario, subject clitics pattern with subject pronouns in the selection of a postverbal non-subject antecedent; this fact supports the idea that they are not simple agreement markers, contrary to what has been proposed in the literature.

iii. Subject clitics can pattern with either overt or null subjects given the correct interpretation

In this scenario, both (i) and (ii) are captured. In order to be correctly interpreted at the interfaces, subject clitics can pattern with either overt subjects or null subjects. This difference is particularly evident in varieties in which subject clitics allow for two different interpretations, such as heritage Venetan.

In order to test the three scenarios, two contexts used by Carminati (2002) were selected to test the PAH in sentences with canonical preverbal subjects. The contexts and the methodology are described in Section 4.3.2. It will be shown that both interpretations are possible for subject clitics; the distinction will be attributed precisely to the contribution made by the discourse-related [uR] feature. In Chapter 5 a further hypothesis will be put forward, regarding the featural composition of different pronominal forms with respect to discourse features in particular, in order to account for the distribution of different forms in terms of the approach proposed in Chapters 2 and 3. I leave the discussion of this hypothesis for Chapter 5. This chapter focusses on sentences with a canonical preverbal subject in Spec-TP.

In Chapters 2 and 3 it has been shown that heritage Italo-Romance varieties are consistent null subject languages; this is true both for southern Italo-Romance varieties such as Calabrian and for northern Italo-Romance varieties such as Venetan. The case of Venetan is particularly interesting, as it allows for null subjects in the same conditions as consistent null subject languages, but it has a more articulated pronominal system due to the presence of subject clitics. It is possible that the reason for the different conditions on subject clitic drop observed in heritage Venetan (and Friulian, as shown in Frasson *et al.*, 2021) depends precisely on the fact that their pronominal system is more complex than the one of southern varieties, given the

148 The syntax of subject pronouns in heritage languages

presence of an extra form: subject clitics. This study focusses on Venetan and aims to check the distribution of [uR] in subject clitics, in order to allow a better understanding of the change undergone by Argentinian and Brazilian Venetan. Recall that in these varieties subject clitics mainly behave as strong pronouns, while Italian Venetan subject clitics generally display an interpretation akin to that of *pro*.

4.3.2. *The questionnaire*

A total of 65 informants participated in the investigation. All of them were native speakers of Venetan, born in Argentina (n.: 3), Brazil (n.: 25) and Italy (n.: 37). As in previous chapters, I will refer to Argentinian and Brazilian Venetan as heritage Venetan, to distinguish these varieties from Italian Venetan. The details about the informants can be found in Appendix A.

The questionnaire consisted in a preference task with 36 items made up of one proposed sentence and three possible answers. Each sentence was repeated three times, each one with a different pronominal form: a strong subject pronoun, a subject clitic or a null subject. Participants were asked to indicate which interpretation of the proposed sentence they preferred: whether they thought the pronominal form in the subordinate clause stated something about the matrix preverbal subject or about the matrix postverbal (indirect) object. As a third possible option, participants could also judge the proposed sentence as ungrammatical. For each item, only one answer could be chosen.

Four different contexts were tested; in this chapter, the first two contexts are presented, testing the PAH in sentences with canonical preverbal subjects in the matrix clause.

The participants were not given extra clues about the context in which the sentences were uttered, to favour a ‘plain intonation’ reading and not a contrastive

one.³⁰ The complete list of items used in the questionnaire can be found in Appendix B.

The results for the first half of the questionnaire are presented in this chapter, referring to the first two contexts analysed. This first half of the questionnaire included 18 sentences per informant. All 65 informants answered the questionnaire; in total 1170 sentences were analysed in the first two contexts.

The two contexts presented here include items with a preverbal subject in the matrix clause and a subordinate clause introduced by *when* or *that*:

4.3.2.1. Temporal subordinate clauses (when-clauses)

The potential referents of the pronoun are introduced in the matrix clause and the pronoun appears as the subject of the subordinate adverbial clause.

(20) Venetan

- a. Marco_j el scrivea sempre a Luca_k quando che lu_k stea mal.
Mark he.SCL wrote.3SG always to Luke when that he.PRON was.3SG sick
- b. Marco_j el scrivea sempre a Luca_k quando che el_? stea mal.
Mark he.SCL wrote.3SG always to Luke when that he.SCL was.3SG sick
- c. Marco_j el scrivea sempre a Luca_k quando che pro_j stea mal.
Mark he.SCL wrote.3SG always to Luke when that pro was.3SG sick
'Mark always used to write to Luke when he was ill.'

According to the PAH (Carminati 2002), the pronoun *lu* in the embedded clause in (20)a) should display a preference for the indirect object *Luca* in the matrix clause, while *pro* in (20)c) should display a preference for the matrix subject *Marco*. As for subject clitics, in the first of the three scenarios proposed at the beginning of Section

³⁰ Introducing the referents as a set of alternatives may favour an emphatic intonation or give rise to a contrastive interpretation (Cruschina 2011); a study of contrast and emphasis is beyond the goal of this study.

150 The syntax of subject pronouns in heritage languages

4.3.1 (subject clitics pattern with null subjects), the subject clitic *el* in (20)b) should display the same preferences as *pro* in (20)c); in the second scenario (subject clitics pattern with tonic pronouns) the subject clitic in (20)b) should display the same preferences as *lu* in (20)a); in the third scenario (subject clitics can pattern with either strong pronouns or *pro* given the correct interpretation) the subject clitic in (20)b) could be interpreted either with respect to the matrix indirect object *Luca* (as the strong pronoun in (20a)) and to the matrix subject *Marco* (as *pro* in (20)c)). Recall that (20)c) is possible in heritage Venetan, but it should not be possible according to studies on subject clitics in Italian Venetan: subject clitics should be obligatorily realised in a position adjacent to that to the verb in this type of sentences. For this reason, informants had the option to judge the proposed sentences as ungrammatical.

The context in which the sentences could be uttered is deliberately left ambiguous. Besides, in this type of sentences, the pragmatic context does not give any clue as to which the correct antecedent of the pronoun is: it could equally well be the subject of the main clause or the object. This choice is motivated with the need to find to check precisely for antecedent biases of the different types of pronouns.

4.3.2.2. Subordinate clauses introduced by “that” (that-clauses)

In this case too, the potential referents of the pronoun are introduced in the initial main clause and the pronoun appears as the subject of the subordinate clause. Following Carminati (2002), report and belief verbs in the matrix clause³¹ were used:

(21) Venetan

a. La Maria_i la ga dito a la Bruna_k che ela_k ga da crompar el pan.
the Maria she.SCL have.3SG said to the Bruna that she.PRON have.3SG to buy the bread

³¹ Contexts of obligatory control are excluded from the study. In order to avoid cases of obligatory control, all the proposed items have an indicative complement clause with independent tense (Landau, 2004).

- b. La Maria_i la ga dito a la Bruna_k che **la**_i ga da crompar el pan.
 the Maria she.SCL have.3SG said to the Bruna that she.SCL have.3SG to buy the bread
- c. **La Maria**_i la ga dito a la Bruna_k che **pro**_i ga da crompar el pan.
 the Maria she.SCL have.3SG said to the Bruna that pro have.3SG to buy the bread
 ‘Maria told Bruna that she has to buy bread.’

As in the previous context, the pronoun *ela* in the embedded clause in (21)a) should display a preference for the indirect object *Bruna* in the matrix clause, while the embedded *pro* in (21)c) should display a preference for the matrix subject *Maria*. As for the subject clitic *la* in (21)b), in the first scenario, it should display the same preferences as *pro* in (21)c); in the second scenario, it should display the same preferences as *ela* in (21)a); in the third scenario, the subject clitic in (21)b) should be interpretable either with respect to the matrix indirect object *Bruna* (as the pronoun *ela* in (21)a) and to the matrix subject *Maria* (as *pro* in (21)c)). Again, (21)c) is possible in heritage Venetan, while it should not be possible according to studies on northern varieties (Rizzi 1986, Brandi and Cordin 1989); (21)a) could be accepted in Venetan (Benincà 1994) as shown in Chapter 3. Carminati’s (2002) findings suggest that the PAH in *that*-clauses should not hold as strongly as in *when*-clauses. However, Carminati (2002) tested only *that*-clauses in which just one possible antecedent was available for the pronoun in the matrix clause. In the present study, a second possible antecedent is introduced to have an equal level of ambiguity in the two contexts. By doing this, it is possible to ascertain whether there is a difference between the two contexts in equal conditions.

4.3.3. Results

This section presents the results relative to the two contexts discussed in Section 4.3.2: *when*-clauses and *that*-clauses. The study followed the methodology used in Carminati (2002) and more recent studies on the PAH (see in particular Frana 2007, discussed in Chapter 5): one-way ANOVAs (significance: $p = <.05$) were performed in the statistical tool R (R Core Team, 2020), with subjects as random effect to

152 The syntax of subject pronouns in heritage languages

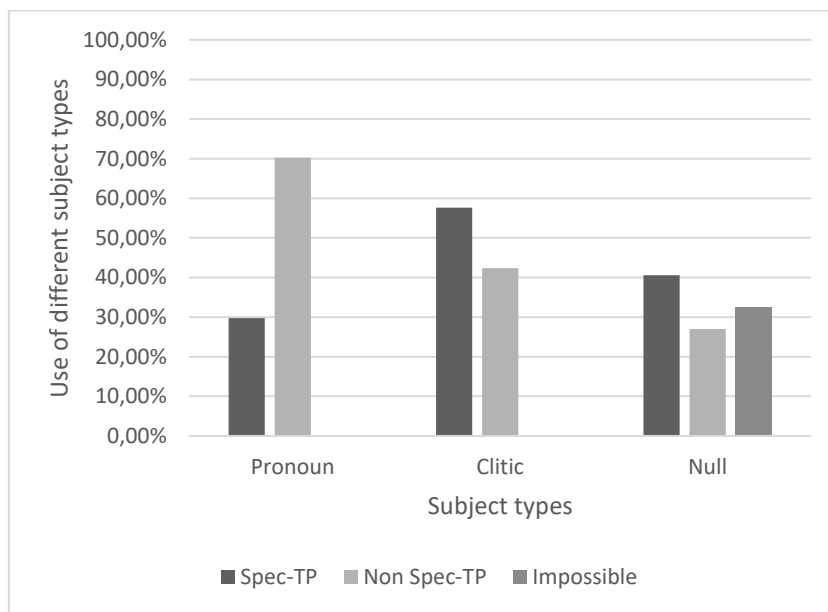
compare the effect of pronoun type (pronominal, clitic or null) on the selection of the antecedent in the two contexts in each of the three groups. The results will show that different interpretations for subject clitics are available in the three contact situations analysed.

4.3.3.1. *When-clauses*

In the case of *when*-clauses, all Venetan varieties data provide strong support for the PAH. However, a significant difference is observed in the interpretation of subject clitics in heritage varieties, with respect to Italian Venetan.

The pattern that emerges in Italian Venetan reflects what is described in the literature, both with respect to the PAH and to subject clitics: *pro* displays a preference for a subject antecedent in Spec-TP, while strong pronouns display a preference for an antecedent that is not in Spec-TP. Finally, subject clitics display a preference for an interpretation akin to that of *pro*. All 37 informants carried out the online questionnaire; each informant had to select the preferred interpretation for 9 sentences with *when*-clauses (3 per pronominal form). In total, 333 sentences were analysed; the results are presented in Figure 1.

Figure 1. When-clauses in Italian Venetan. Number of informants: 37. Number of tested sentences per informant: 9 (3 per pronominal form). Total of analysed sentence: 333.



The pattern displayed by Italian Venetan is supported by the results of the One-way ANOVA. Recall that in the statistical analysis I followed the methodology used in Carminati (2002) and subsequent studies on the PAH (see in particular Frana 2007), by comparing the results for each pronominal form to the other forms in each group of speakers; this was done by means of a One-way ANOVA. As far as Italian Venetan is concerned, results confirmed that the PAH holds in Italian Venetan, as the model rendered a significant difference between the antecedent preference of strong pronouns and *pro* ($F(3.89) = 18.32, p < .05$): while strong pronouns prefer an antecedent that is not in Spec-TP, *pro* prefers the most salient antecedent in Spec-TP. The difference between the antecedent preference of strong pronouns and subject clitics is significant as well ($F(3.88) = 18.93, p < .05$). Crucially, subject clitics on a par with *pro* tend to prefer a higher antecedent: the difference in their antecedent preference is not significant ($F(3.89) = 0.10, p > .05$). The preferred interpretation of strong pronouns, subject clitics and *pro* in Italian Venetan is exemplified in (22):

154 The syntax of subject pronouns in heritage languages

(22) Italian Venetan

La Maria_j la parlea co la Bruna_k quando che **ela_k** / **la_j** / **pro_j** vivea in Italia.
the Maria she.SCL talked with the Bruna when that she.PRON she.SCL pro lived in Italy
'Mary talked to Bruna when she lived in Italy.'

This pattern is consistent with the studies on Venetan (Benincà 1994, Poletto 2000), according to which subject clitics should not behave as tonic pronouns. However, the same studies claim that a subject clitic should not be dropped in the proposed sentences, therefore items without an overt subject clitic should be judged ungrammatical. In order to verify whether subject clitics are indeed obligatory, a further One-way ANOVA was performed to check for items judged ungrammatical. However, the ANOVAs including sentences judged ungrammatical did not render a significant effect for any of the forms (*pro*, accepted vs. not accepted: $F(3.08) = 0.12$, $p = >.05$; tonic and clitic subjects were always accepted). This result shows that most informants accept a sentence without a subject clitic, interpreting it in a way that is perfectly consistent with the predictions made by the PAH. This point is particularly interesting, as it shows that null *pro* is accepted in Italian Venetan as well, although not by all informants.

Moving to heritage Venetan varieties, the study confirmed that the PAH holds in Brazilian Venetan, but subject clitics behave differently than in Italian Venetan. Null *pro* displays a clear preference for a salient antecedent in Spec-TP, while both strong pronouns and subject clitics display a preference for an antecedent that is not in Spec-TP. All 25 informants carried out the online questionnaire; each informant had to select the preferred interpretation for 9 sentences with *when*-clauses (3 per pronominal form). In total, 225 sentences were analysed; the results are presented in Figure 2.

Figure 2. When-clauses in Brazilian Venetan. Number of informants: 25. Number of tested sentences per informant: 9 (3 per pronominal form). Total of analysed sentence: 225.

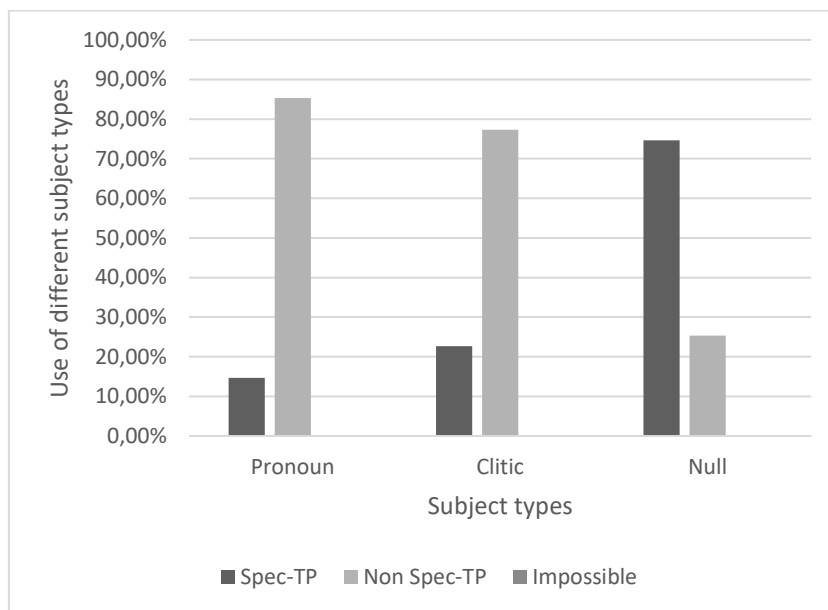


Figure 2 shows that a subject pronoun prefers a low antecedent that is not in the Spec-TP; *pro*, on the other hand, prefers an antecedent in Spec-TP.

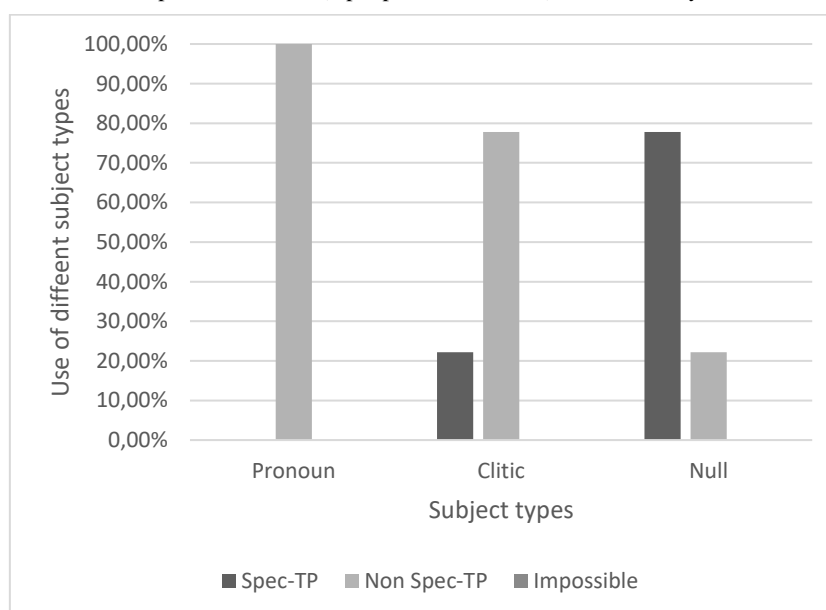
This behaviour is confirmed by the One-way ANOVA. The results showed that the PAH holds in Venetan, in that the antecedent preference of strong pronouns and *pro* are significantly different ($F(3,90) = 84.75, p < .05$). Unlike Italian Venetan, however, also the antecedent preference of subject clitics and *pro* are significantly different in Brazilian Venetan ($F(3,90) = 54.90, p < .05$). The preference of strong pronouns does not significantly differ from that of subject clitics ($F(3,90) = 1.57, p > .05$). In other words, both subject clitics and tonic pronouns behave significantly differently from null subjects in that the two overt forms display a preference towards an antecedent that is not in Spec-TP, while null subjects display a preference towards an antecedent in Spec-TP.

156 The syntax of subject pronouns in heritage languages

None of the Brazilian informants judged “pure” null-subject sentences (with neither a tonic pronoun nor a subject clitic) as ungrammatical, against what is expected on the basis of the observation that subject clitics are as obligatory agreement elements.

The second heritage variety considered in the study is Argentinian Venetan, which displays a pattern parallel to that found in Brazilian Venetan. All 3 informants carried out the online questionnaire; each informant had to select the preferred interpretation for 9 sentences with *when*-clauses (3 per pronominal form). In total, 27 sentences were analysed; the results are presented in Figure 3.

Figure 3. *When-clauses in Argentinian Venetan.* Number of informants: 3. Number of tested sentences per informant: 9 (3 per pronominal form). Total of analysed sentence: 27.



The One-way ANOVA confirmed that the PAH holds in Argentinian Venetan, in that the difference in the antecedent preference of strong pronouns and *pro* is significant ($F(4.49) = 65535, p < .05$); just as in the case of Brazilian Venetan, the difference between subject clitics and *pro* in Argentinian Venetan is significant as well ($F(4.49) = 28, p < .05$); finally, the difference in the preferences of strong

pronouns and subject clitics is not significant ($F(4,49) = 2.28, p = >.05$). In Argentinian Venetan too, both subject clitics and tonic pronouns display a preference towards an antecedent that is not in Spec-TP, while null subjects display a preference towards an antecedent in Spec-TP. As in the case of Brazilian Venetan, it was not necessary to perform a further One-way ANOVA on items judged ungrammatical, in that Argentinian Venetan informants accepted all the proposed items.

In sum, heritage Venetan varieties display the pattern shown in example (23):

(23) Venetan

Marco_i el scrivea sempre a Luca_k quando che **lu_k** / **el_k** / **pro_i** stea mal.
Mark he.SCL wrote always to Luke when that he.PRON he.SCL pro was sick.
'Mark always wrote to Luke when he was sick.'

The PAH strongly holds in heritage Venetan, as shown by the fact that only the null form prefer the antecedent in Spec-TP, while overt forms (both strong and clitic) prefer the antecedent that is not in Spec-TP.

In conclusion, the study showed that the PAH holds in all Venetan varieties. However, subject clitics in Italian Venetan are interpreted on a par with null *pro*, in that prefer an antecedent in Spec-TP. In heritage Venetan, the PAH also holds, but subject clitics pattern with strong pronouns, in that they prefer an antecedent that is not in Spec-TP.

4.3.3.2. *That-clauses*

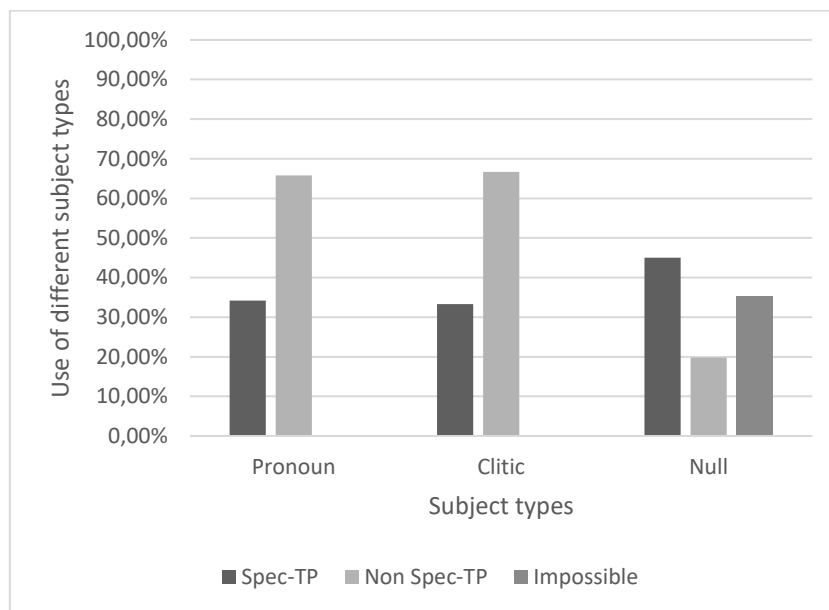
In the case of *that*-clauses, data generally support the predictions made by the PAH. In this case too, a considerable amount of variation is observed in subject clitics in different varieties.

Italian Venetan here display a behaviour akin to that of *when*-clauses, in that strong pronouns prefer an antecedent that is not in Spec-TP, while null subjects prefer an antecedent in Spec-TP. However, unlike the previous context, subject clitics pattern with subject pronouns. All 37 informants carried out the online questionnaire;

158 The syntax of subject pronouns in heritage languages

each informant had to select the preferred interpretation for 9 sentences with *that*-clauses (3 per pronominal form). In total, 333 sentences were analysed; the results are presented in Figure 4.

Figure 4. *That-clauses in Italian Venetan.* Number of informants: 37. Number of tested sentences per informant: 9 (3 per pronominal form). Total of analysed sentence: 333.



The ANOVAs confirmed that the PAH holds in Italian Venetan: the antecedent preference of strong pronouns and *pro* is significantly different ($F(3.89) = 24.33$, $p < .05$). However, unlike the case of *when*-clauses, in the context of *that*-clauses, subject clitics pattern with strong pronouns, in that the difference in their antecedent preference is not significant ($F(3.88) = 0.01$, $p > .05$); the preference of subject clitics, however, is significantly different from that of *pro* ($F(3.89) = 25.80$, $p < .05$). In other words, Italian Venetan allows for a different interpretation of subject clitics, depending on the type of subordinate in which they are realised. Crucially, in the case of complement clauses introduced by *that*, subject clitics prefer an antecedent that is not in Spec-TP, just like strong pronouns.

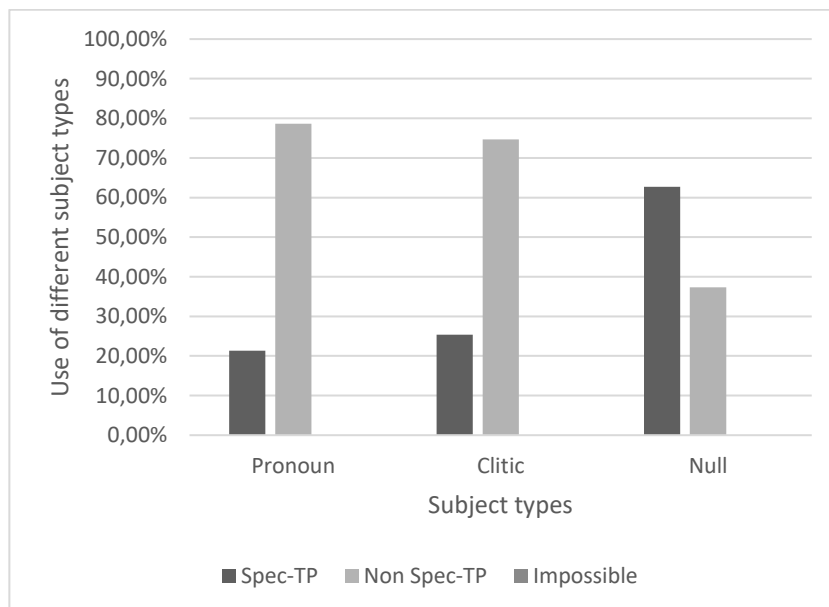
(24) Venetan

La Maria_j la ga dito ala Bruna_k che **ela_k** / **la_k** / **pro_j** ga da crompar el pan.
 the Maria she.SCL has said to.the Bruna that she.PRON she.SCL pro has to buy the bread
 ‘Maria told Bruna that she has to buy bread.’

The ANOVAs including sentences judged ungrammatical did not render a significant effect for any of the forms (null subjects, accepted vs. not accepted: $F(3,08) = 1.42$, $p = >.05$; sentences with strong and clitic subjects were always accepted).

In heritage Venetan varieties, subject clitics display the same preferences in *that*-clauses and in *when*-clauses. Brazilian Venetan, again, displays a regular pattern: subject pronouns prefer an antecedent that is not in Spec-TP, while *pro* prefers an antecedent in Spec-TP. Subject clitics pattern with strong pronouns. All 25 informants carried out the online questionnaire; each informant had to select the preferred interpretation for 9 sentences with *that*-clauses (3 per pronominal form). In total, 225 sentences were analysed; the results are presented in Figure 5.

Figure 5. *That*-clauses in Brazilian Venetan. Number of informants: 225. Number of tested sentences per informant: 9 (3 per pronominal form). Total of analysed sentence: 225.



160 The syntax of subject pronouns in heritage languages

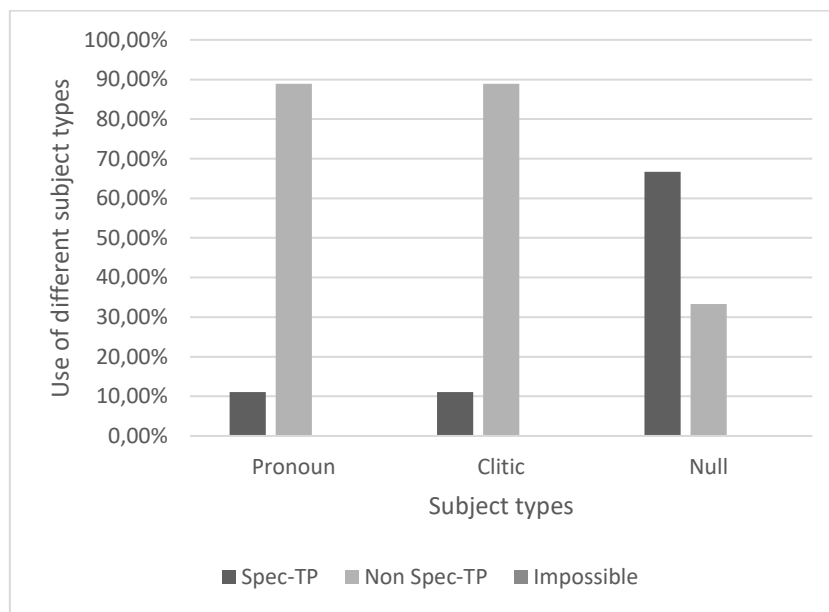
The ANOVAs show that the PAH holds for Brazilian Venetan: the difference in the interpretation of strong pronouns and *pro* is significant ($F(3,90) = 31.46$, $p < .05$). The difference in the interpretation of subject clitics and *pro* is significant too ($F(3,90) = 24.37$, $p < .05$), while that of strong pronouns and subject clitics is not ($F(3,90) = 0.33$, $p > .05$).

In Brazilian Venetan, strong pronouns and subject clitics display a similar behaviour, in that they generally select an antecedent that is not in Spec-TP; null subjects, conversely, behave significantly differently from the two types of overt pronouns in that they prefer an antecedent in Spec-TP.

As in the previous context, none of the Brazilian informants judged “pure” null-subject sentences (with neither a strong pronoun nor a subject clitic) ungrammatical, against what is expected on the basis of the widely accepted analysis of subject clitics as obligatory elements.

Finally, Argentinian Venetan displays the same type of interpretation for the different pronoun types as Brazilian Venetan. All 3 informants carried out the online questionnaire; each informant had to select the preferred interpretation for 9 sentences with *that*-clauses (3 per pronominal form). In total, 27 sentences were analysed; the results are presented in Figure 6.

Figure 6. *That-clauses in Argentinian Venetan.* Number of informants: 3. Number of tested sentences per informant: 9 (3 per pronominal form). Total of analysed sentence: 27.



The ANOVAs confirmed that the PAH holds for *that*-clauses in Argentinian Venetan: the difference in the interpretation of strong pronouns and *pro* is significant ($F(4,49) = 7.69, p < .05$), and so is the difference in the interpretation of subject clitics and *pro* ($F(4,49) = 7.69, p < .05$). The difference in the interpretation of subject clitics and strong pronouns, however, is not significant ($F(4,49) = 0, p > .05$). Also in this case, strong pronouns and subject clitics do not differ significantly in their interpretation: they both select an antecedent that is not in Spec-TP; *pro* prefers an antecedent in SpecTP. The preferences of subject pronoun types in heritage Venetan varieties are shown in (25):

(25) Venetan

Marco_j el ga dito a Luca_k che lu_k / el_k / pro_j ga magnà massa.

Mark he.SCL has told to Luke that he.PRON he.SCL pro has eaten too much
 ‘Mark told Luke that he ate too much.’

162 The syntax of subject pronouns in heritage languages

The ANOVAs including sentences judged ungrammatical was not performed, as all proposed sentences were accepted. This result implies that most informants accept a sentence without a subject clitic, interpreting it in a way that is perfectly consistent with the predictions made by the PAH.

In conclusion, there is no difference between *when*-clauses and *that*-clauses in Italian Venetan as far as the PAH is concerned: the interpretation of strong pronouns is significantly different from that of *pro*. However, subject clitics pattern with *pro* in *when*-clauses, but with strong pronouns in *that*-clauses. In heritage Venetan, the PAH always holds and subject clitics select an antecedent that is not in Spec-TP in both contexts, on a par with subject pronouns.

4.3.3.3. Summary of the results

The results of the study showed that most speakers accept sentences without subject clitics in all contact situations in the two analysed contexts. Therefore, I propose the following generalisation on the realisation of subject clitics:

Subject clitics are not obligatory and verbs can appear without an overt pronominal form if the null subject refers to a salient antecedent.

This generalisation implies that the system of all Venetan varieties analysed include three possible pronominal forms: tonic subject pronouns, subject clitics and null subjects. Recall that, particularly for Italian Venetan, the option of a phonologically null pronoun should not be possible according to previous studies. The main difference between heritage and Italian Venetan, then, is not in the fact that only heritage varieties allow for null subjects, but rather that in the interpretation of subject clitics: while in heritage varieties they always favour an interpretation akin to that of strong pronouns, in Italian Venetan they allow for two different interpretations, depending on the subordinate type in which they are realised. These results are summarised in Table 1:

Table 1. *The PAH in Venetan varieties and the preferences of subject clitics.*

	<i>When</i> -clauses	<i>That</i> -clauses
Italian Venetan	<i>Subject clitic</i> \approx <i>Null</i>	<i>Subject clitic</i> \approx <i>Pronoun</i>
Brazilian Venetan	<i>Subject clitic</i> \approx <i>Pronoun</i>	<i>Subject clitic</i> \approx <i>Pronoun</i>
Argentinian Venetan	<i>Subject clitic</i> \approx <i>Pronoun</i>	<i>Subject clitic</i> \approx <i>Pronoun</i>

Table (1) show that subject clitics are most of the times interpreted on a par with strong pronouns, while their interpretation as null subjects is limited to Italian Venetan and specifically in the context of *when*-clauses. The two possible interpretations of subject clitics in Italian Venetan will be discussed in Section 4.4.

Recall that, in Section 4.3.2, three possible scenarios for the interpretation of subject clitics were predicted:

i. Subject clitics always pattern with null subjects

ii. Subject clitics always pattern with overt subjects

iii. Subject clitics can pattern with either overt and null subjects given the correct interpretation

(i) has to be excluded: it is never the case in the Venetan varieties analysed.

In (ii), subject clitics pattern with tonic subject pronouns in the selection of an antecedent that is not in Spec-TP; this is the case of heritage Venetan.

Finally, (iii) is represented by Italian Venetan. Subject clitics can pattern with either overt subjects or null subjects. I propose that this difference depends on the clause-type.

4.4. Antecedent selection and [uR]

The data presented in Section 4.3 support the proposal on the internal composition of pronouns presented in Chapters 2 and 3. The antecedent preferences that emerge in previous studies on the PAH (Carminati 2002; Kaiser and Trueswell 2008; van Kampen 2012; Mayol 2009) as well as in the present study on Venetan, depend on the

presence of an uninterpretable [uR] feature on the pronoun and an interpretable [R] on a Logophoric Centre in C, which links a pronoun in the subordinate to a possible antecedent; the study on Venetan further suggests that different factors may apply across different varieties and also inside the same variety of a language, depending on the selected form. In this section, this model is applied to the analysed contexts and to the PAH predictions. In Section 4.4.1, it is shown that the notion of [uR] captures the idea of ‘antecedent preference’ and therefore has the power of explaining cross-linguistic and intra-speaker variation in the analysed contexts. In Section 4.4.2, however, I reflect more extensively on the difference displayed in the interpretation of subject clitics in Italian Venetan in the two analysed contexts; therefore, while discourse factors play an important role in defining the preferred interpretation of subject clitics, in some cases the possible interpretation is constrained by syntax.

4.4.1. Antecedent selection properties of overt and null subjects

The PAH holds in all Venetan varieties in the analysed contexts. Recall that the realisation and the interpretation of three types of subjects (strong pronouns, subject clitics and null subjects) were tested in two types of subordinate clauses: temporal clauses introduced by *when* (*when*-clauses) and complement clauses introduced by *that* (*that*-clauses).

As far as null subjects are concerned, in Chapter 2 it was proposed that they lack [uR] on the D-layer; precisely because of this fact, *pro* results from deletion of the pronoun under feature-identity at PF (Sheehan 2006, Roberts 2009). In Venetan, *pro* prefers an antecedent in Spec-TP. Notice that the antecedent in this position is generally linked to the notions of topic and old information, which are compatible with the concept of “highest salience value” discussed in Chapter 2.

(26) Venetan

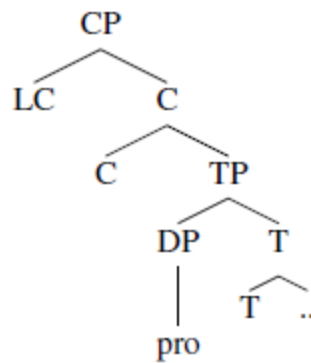
La Maria_j la ciama la Bruna_k quando che **pro_j** riva tardi.

the Maria she.SCL call.3SG the Bruna when that pro arrive.3SG late

‘Maria calls Bruna when she is late.’

Null *pro* does not have an uninterpretable [uR] feature. Therefore, there are no uninterpretable features to be checked and deleted against an interpretable [R] feature on the Logophoric Centre in the embedded CP. In Chapter 2 it was proposed that the Logophoric Centre also lacks [R], in that the most salient discourse antecedent (the one in the matrix Spec-TP) was selected via Context Scanning. The structure is repeated in (27):

(27)



The case of strong pronouns is different, in that they encode a [uR] feature in D. The presence of [uR] implies that the pronoun is used to switch the referent to a non-salient discourse antecedent. In the proposed contexts, such antecedent is the one that is not in Spec-TP in the matrix clause.

(28) Venetan

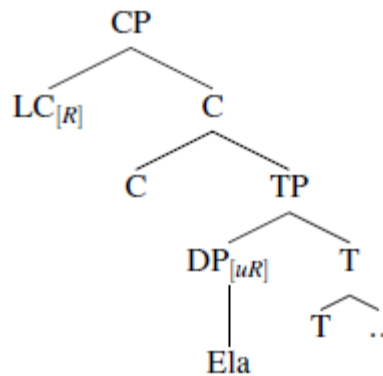
La Maria_j la ciama **la** Bruna_k quando che **ela_k riva tardi.
 the Maria she.SCL call.3SG the Bruna when that she.PRON arrive.3SG late
 ‘Maria calls Bruna when she is late.’**

In this case, PF deletion of the pronoun is not possible, as the ϕ -composition of the pronoun is not identical to that of finite T. [uR] on the pronoun has to be checked against [R] on the Logophoric Centre before transfer to LF, following the definition

166 The syntax of subject pronouns in heritage languages

of Agree proposed in Chapter 2 building on Wurmbrand (2010) and Zeijlstra (2012). Recall that [R] is optionally realised on the Logophoric Centre: the realisation of [R] results from an assignment function at the discourse level, in the case in which reference needs to be switched to a non-salient discourse antecedent. The structure is repeated in (29):

(29)



The presence of [R] on the Logophoric Centre is crucial, in that it deletes [uR] on the pronoun and it allows the derivation to proceed. Only in this case the overt strong pronoun is licensed in consistent null subject languages.

Overall, in the cases of null subjects and pronouns, the predictions made by the PAH are fully captured in the approach adopted in the dissertation. The introduction of [uR] allows to explain the difference at the level of interface between syntax and discourse: the absence of [uR] triggers the selection of the most salient antecedent; the presence of [uR] allows for the selection of a non-salient antecedent. This selection can be defined as preference of a pronominal form for one of the potential antecedents; the notion of preference is also captured in the present model, as showed in Section 4.2. Before moving to the discussion of preference, one last point needs to be addressed: the ambiguous reference of subject clitics.

4.4.2. *The interpretation of subject clitics in subordinate clauses*

In Italian Venetan, subject clitics can select either an antecedent in Spec-TP or an antecedent that is not in Spec-TP. These two positions in the structure are linked to the notion of salience: the referent in Spec-TP is generally the most salient one; the referent that is not in Spec-TP is intended as the less salient or as new information. Crucially, subject clitics have been shown to be able to select both referents as antecedents:

(30) Venetan

La Maria_j la chiama la Bruna_k quando che **la**_{j/k} riva tardi.
the Maria she.SCL call.3SG the Bruna when that she.SCL arrive.3SG late
'Maria calls Bruna when she is late.'

In heritage Venetan the subject clitic *la* prefers the antecedent *Bruna*, which is not in Spec-TP. In this case the subject clitic displays a behaviour akin to that of the strong pronoun in (28): it encodes [uR] and refers a non-salient antecedent, as showed in (29). However, in Italian Venetan, the subject clitic *la* prefers the antecedent *Maria* in Spec-TP. In this case the subject clitic displays a behaviour akin to that of *pro* in (26); in Chapter 3 it was proposed that subject clitics do not get deleted at PF in this case, if they are adjacent to the finite verb in T.

The ambiguous behaviour of subject clitics fact is particularly evident in the comparison of *when*-clauses and *that*-clauses in Italian Venetan; subject clitics display a different behaviour in the two analysed contexts. While the subject clitic *la* in *when*-clauses such as (30) selects the higher antecedent in Spec-TP, the subject clitic *el* in *that*-clauses (31) selects the antecedent that is not in Spec-TP, on a par with strong pronouns:

(31) Venetan

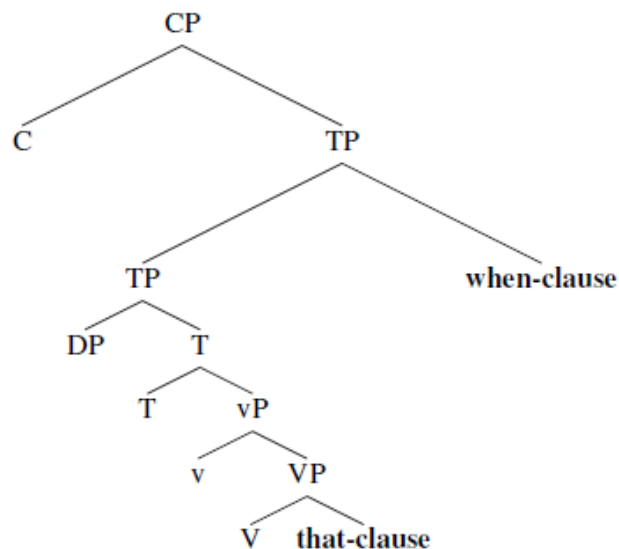
Marco_j el ga dito a Luca_k che **el**_k ga magnà massa.
Mark he.SCL have.3SG told to Luke that he.SCL have.3SG eaten too much

‘Mark told Luke that he ate too much.’

In both heritage and in Italian Venetan, the antecedent selected by *el* is *Luca*, the one that is associated with a lower salience value.

The case of Italian Venetan is particularly interesting, as different interpretations of subject clitics emerge in different syntactic contexts. This difference does not pose any problems for the approach to the internal structure of pronouns; it depends on a structural difference between the two types of subordinate clauses, which has an effect on the type of subject clitic (encoding [uR] or lacking [uR]) allowed in each of them. Such asymmetries were partially discussed in Carminati (2002) with respect to the PAH in Italian: she suggested that different PAH effects may depend on a difference in the attachment level of the two clause types. I assume that this is correct and propose that the constraint in the interpretation of subject clitics in *that*-clauses follows from the idea that complement clauses (like *that*-clauses in the present study) are complements of VP, while adjunct clauses (like temporal clauses in the present studies) are CP adjuncts; in this respect, see for instance Baker’s (1991) analysis of disjoint reference.

(32) (adapted from Carminati, 2002)



I propose that the differences in the interpretation of subject clitics in Italian Venetan follow precisely from this structural difference. Specifically, Italian Venetan constrains the realisation of overt subjects in *that*-clauses to cases of switch reference. In this case, the condition for subject clitics to be realised is the presence of [uR], which implies a switch to a non-salient reference.

The restriction depends again on accessibility. To account for the restriction, it is possible to reduce accessibility to a notion of locality and c-command.

Consider the definition of accessibility in terms of binding given in Reuland and Everaert (2001):

(33) Accessibility (Reuland and Everaert, 2001: 639)

- a. α is accessible to β iff β is in the c-command domain of α , and assignment to β of the index of α would not violate (b).
- b. $*[\tau \dots \delta \dots]$, where τ and δ bear the same index

The definition in (33) builds on Chomsky's (1981) *i*-within-*i* Condition, according to which an anaphoric element cannot occur within a phrase with the same index; therefore, it cannot be applied to the case under analysis, as pronouns are not bound by their antecedent. The present study refers to cases of covaluation, which was defined in Reinhart (2006: 172) as the case in which two elements share the same value but neither binds the other. The notion of 'value' discussed in Reinhart (2006) can be extended to salience values and the definition of accessibility in (33) can be restated in terms of salience structure, to capture the difference in the interpretation of subject clitics in *that*-clauses.

Consider again the notion of salience value presented in Chapter 2. Following Von Heusinger (2000), it was assumed that each sentence has a salience structure that consists of a set $A = \{d_1 \dots d_n\}$ of ranked discourse referents carrying different salience values, where the first element d_1 of the set has the highest salience value. The referent in matrix Spec-TP in (31) corresponds to the element d_1 of the set, so it has the highest salience value. Conversely, the postverbal element in (31)

170 The syntax of subject pronouns in heritage languages

corresponds to an element $d_{\neq i}$, so it has a lower salience value. The embedded clause in (31) also has a salience structure. The subject clitic in embedded Spec-TP in (31) will also carry a salience value: since it corresponds to the element d_i of the set, the subject clitic will carry the highest salience value. This fact implies that the referent in the matrix Spec-TP and the subject clitic in the embedded Spec-TP carry the same salience value³².

Going back to the notion in (33), I propose that accessibility can be restated in terms of salience values as in (34):

(34) Accessibility

- a. an antecedent a is accessible to an overt pronoun p iff p is in the c-command domain of a , and assignment to p of the reference of a would not violate (b).
- b. $*[a \dots p \dots]$, where a and p carry the same salience value.

With respect to (34)a), both the subject and the indirect object in the matrix clause c-command the Logophoric Centre and the subject clitic in the complement clause; however, the realisation of a subject clitic lacking [uR] would violate (34)b), in that it carries the same salience value as the matrix subject. A subject clitic can be licensed in this environment only if it carries an [uR] feature and the Logophoric Centre is assigned [R] via Context Scanning: in this case (34)b) is not violated, as the indirect object in the matrix clause and the subject clitic in the embedded clause carry different salience values. Notice finally that this restriction does not apply to adjunct *when*-clauses, in which the potential antecedents are not local and do not c-command the pronoun in the subordinate. Therefore, a subject clitic lacking [uR] is allowed in this environment.

The situation of heritage Venetan varieties seems to be simplified with respect to that of Italian Venetan, in that the same interpretation is available for subject clitics in both contexts.

³² Recall that [uR] is licensed on the subject clitic precisely when it needs to update the salience structure, in that the referent corresponding to $d_{\neq i}$ in the set needs to become salient.

4.4.3. *A note on the notion of “preference”*

The analysis presented in Section 4.4.2 is developed following the results of the study presented in Section 4.3. Given the amount of data collected, a statistic analysis was carried out on the data; the analysis developed is based on the statistic significance of the results. However, leaving aside statistics, results display a lot of variation: for instance, there are cases in which *pro* corefers with the antecedent that is not in Spec-TP. Conversely, there are cases in which strong pronouns corefer with the most salient antecedent in Spec-TP.

Following Carminati (2002), the predictions made by the PAH are defined in terms of preference. One further advantage of the proposal made in this study on the internal structure of pronouns, is that it can capture the notion of preference in terms of ‘preferred interpretation’ of each pronominal form. This notion is related to the presence or absence of [uR]. As shown in Chapter 2, [uR] is an optional feature: it is generally realised on strong pronouns and it is lacked by *pro*. However, nothing prevents an overt pronoun to be realised without [uR]; in that case it may behave as an overt weak pronoun, in Cardinaletti and Starke’s terms. This is the case of Italian *egli*, as described in Cardinaletti (1997). Italian is a canonical null subject language, in that referential subject pronouns are normally not spelled out unless they encode some extra discourse-related property. However, the pronoun *egli*, despite being overt, has the properties of a weak pronoun, which are, crucially, shared by Venetan subject clitics too, as showed in Chapter 2: it cannot be modified (35)a), coordinated (35)b), dislocated (35)c) or realised in isolation (35)d).

(35) Italian (Cardinaletti, 1997)

- a. *Anche **egli** ha dichiarato la propria responsabilità.
 too he have.3SG declared the own responsibility
 ‘He too has declared his own responsibility.’

172 The syntax of subject pronouns in heritage languages

b. ***Egli** e suo fratello hanno dichiarato la propria responsabilità.

he and his brother have.3PL declared the own responsibility

‘He and his brother have declared their own responsibility.’

c. ***Egli**, Maria non l’ appoggierebbe.

he Mary not him.CL support.3SG

‘As for him, Mary would not support him.’

d. Chi è arrivato? – ***Egli**.

who be.3SG arrived he

‘Who has arrived? – He.’

This evidence is taken by Cardinaletti (1997) to imply that *egli* is a weak pronoun just like *pro*, the only difference between the two being the phonological realisation of the first. In the present approach, Italian *egli* lacks [uR] in the D-layer and it will display the same interpretive properties as *pro*, as far as the PAH is concerned. In fact, even a consistent null subject language like Italian allows for overt subject pronouns in certain environments; it is not the preferred choice, as suggested by the very limited use of *egli* in Italian, but nothing prevents its use. As in the case of Venetan subject clitics lacking [uR], Italian *egli* will simply behave as a “phonological” clitic (Rizzi, 1986), in that the constraints on its realisation apply exclusively at PF, while there is no need to assume its structural deficiency.

As for the cases in which *pro* corefers with the antecedent that is not in Spec-TP, I propose that this is the case in which the Logophoric Centre in embedded C encodes an interpretable [R] feature, but the pronoun in the subordinate Spec-TP lacks [uR]. The interpretation of the embedded clause simply remains ambiguous in this case; the derivation, however, does not crash, as there is no uninterpretable feature to delete on the pronoun.

4.4.4. Summary

The analysis of the structure of pronouns just proposed allows to capture the distribution of different pronominal forms with respect to the predictions made by the PAH. In addition, the optional presence of [uR] on subject clitics explains their ambiguous behaviour, especially in Italian Venetan, in which subject clitics prefer different antecedents in different syntactic configurations. The system of heritage Venetan varieties appears to be more regular and simple than that of Italian Venetan: the preferences of subject clitics are the same in the two tested contexts. At this point, it is necessary to briefly consider the role of language contact in defining the two patterns identified in the data: all Venetan varieties are spoken by bilingual speakers, which have Italian, Portuguese or Spanish as a dominant language.

4.5. The role of the contact languages

The study presented in this chapter allows to formulate two important generalisations on the distribution of null and overt pronouns. The first one regards the availability of both strategies (PF deletion or cliticisation) in Venetan; this was already discussed in Section 4.4. The second generalisation regards the fact that, despite allowing for both strategies, heritage Venetan varieties prefer PF deletion (in that subject clitics are generally interpreted on a par with strong pronouns), while Italian Venetan prefers to overtly realise subject clitics under adjacency (in that subject clitics generally have the interpretive properties of *pro*). This section reflects on this last point, making reference to the role of contact.

All the varieties presented in this study are spoken in contact with another language, which represents the dominant language for all the speakers. These dominant languages are Italian in Italy, Portuguese in Brazil and Spanish in Argentina. Two of these languages (Italian and Spanish) are consistent null subject languages, while Brazilian Portuguese is described as a partial null-subject language. Recall that the difference between consistent and partial null subject languages is defined in Holmberg (2005) in terms of availability of null referential subjects: consistent null

174 The syntax of subject pronouns in heritage languages

subject languages such as Italian and Spanish have null referential subjects, while partial null subject languages such as Brazilian Portuguese lack them. Venetan, as well as the other varieties discussed in Chapters 2 and 3 are consistent null subject languages, regardless of the status of the contact languages with respect to the availability of referential null subjects.

4.5.1. The PAH in contact languages

As far as the PAH is concerned, Italian displays the same pattern of Venetan (excluding subject clitics, which are absent in Italian); Argentinian Spanish, as a consistent null subject language, should ideally follow the same pattern. To check whether this is true, one monolingual speaker of Italian (born and raised in the same area of Italy in which Venetan is spoken) and one of Argentinian Spanish (born and raised in Buenos Aires) were asked to judge sentences parallel to those proposed in the study of the PAH in Venetan.

The judgments of the informants confirm that the PAH holds both in Argentinian Spanish and in Italian:

(36) *when*-clauses

a. Italian

Marco_j scriveva sempre a Luca_k quando lui_k / *pro*_j stava male.

Mark wrote.3SG always to Luke when he pro was.3SG sick

‘Mark always used to write to Luke when he was sick.’

b. Argentinian Spanish

Marcos_j siempre le escribía a Lucas_k cuando él_k / *pro*_j estaba enfermo.

Mark always him.CL wrote.3SG to Luke when he pro was.3SG sick

‘Mark always used to write to Luke when he was sick.’

(37) *that*-clauses

a. Italian

Maria_j ha detto a Lucia_k che **lei_k** / *pro_j* doveva comprare il pane.

Mary have.3SG said to Lucy that she pro had.3SG to buy the bread

‘Mary told Lucy that she had to buy bread.’

b. Argentinian Spanish

María_j le dijo a Lucía_k que **ella_k** / *pro_j* tenía que comprar pan.

Mary her.CL said.3SG to Lucy that she pro have.3SG to buy bread

‘Mary told Lucy that she had to buy bread.’

Examples (36-37) show that the PAH holds in both Argentinian Spanish and Italian: *pro* prefers an antecedent in Spec-TP, while the overt subject pronoun prefers an antecedent that is not in Spec-TP.

As for Brazilian Portuguese, the questionnaire was carried out by a monolingual speaker of the language in Rio Grande do Sul (the same are in which most Venetan speakers are found in Brazil). The judgments evidence that the PAH does not hold in Brazilian Portuguese:

(38) *when*-clauses

Marcos_j sempre escrevia para Lucas_k quando **ele_j** / *pro_j*?? estava doente.

Mark always wrote.3SG for Luke when he pro was.3SG sick

‘Mark always used to write to Luke when he was sick.’

(39) *that*-clauses

Maria_j disse a Lúcia_k que **ela_j** / *pro_j*?? precisava comprar pão.

Mary said.3SG to Lucy that she pro had.3SG to buy bread

‘Mary told Lucy that she had to buy bread.’

In both contexts, an overt pronoun prefers the antecedent in Spec-TP; a subordinate clause with null *pro* in subject position was judged as severely degraded in both

contexts and in any case acceptable only if the selected antecedent is the one in Spec-TP.

A complete analysis on the PAH in the contact languages is beyond the scope of this study; however, judgments from native speakers of Argentinian Spanish, Brazilian Portuguese and Italian evidence two patterns: the first one, in which the PAH holds, is represented by the two consistent null subject languages (Argentinian Spanish and Italian); the second, in which the PAH does not hold (and null subjects are severely degraded) is represented by Brazilian Portuguese, a partial null subject language.

Knowing that Venetan, the focus of this study, is always spoken in contact with one of these three dominant languages, one may expect a transfer effect reminiscent of the one described in Montrul (2004): structures of the dominant L2 language are most likely transferred to heritage L1 language when the input evidence from the two languages is conflicting. In the case under analysis, speakers of Venetan in Brazil should receive conflicting input with respect to subject realisation, since in Brazilian Portuguese (which is the dominant languages for these speaker) null subjects are not allowed in the same conditions as in Venetan and, with respect to the PAH, they may not even be possible at all. Conversely, subject realisation should not be problematic for speakers of Venetan in Argentina and Italy, since their dominant languages (respectively Spanish and Italian) display a distribution of overt and null subjects comparable to that of Venetan. Therefore, Argentinian and Italian Venetan should display the same pattern, while Brazilian Venetan should behave differently with respect to subject realisation; the different behaviour of Brazilian Venetan could derive from structural simplification or convergence resulting from transfer from the dominant language.

However, as showed in Section 4.3, this is not the case: Argentinian and Brazilian Venetan display the same pattern, despite the dominant languages of the speakers (Argentinian Spanish and Brazilian Portuguese) displaying different properties with respect to the realisation of overt referential subjects; Italian Venetan is the one that displays a different pattern instead.

4.5.2. Feature-reassembly in contact: a possible solution

The lack of transfer effects from the dominant languages can be explained applying Lardiere's (2008) idea of feature assembly in second language acquisition to the case of subject realisation in heritage Venetan.

According to Lardiere, it is not an easy task for learners of a second language to acquire all the obligatory and optional conditions and restrictions on the overt expression of a feature. Instead of a parameter-resetting approach, she proposes that cases of this type should be analysed as feature-reassembly.

In the case of Venetan subject clitics, it is plausible that the way features are assembled in heritage varieties is not precisely the same as in Italian Venetan. This refers in particular to the realisation of [uR] on subject clitics, which is significantly more frequent in heritage Venetan than in Italian Venetan, affecting the contexts and syntactic positions in which subject clitics appear. Speakers of Argentinian and Brazilian Venetan do not have any problem in correctly using a subject clitic; however, their grammars allow for subject clitics to encode a [uR] features in a wider number of syntactic environments with respect to what happens in Italian Venetan, as shown by the case of *when*-clauses and *that*-clauses.

Subject clitics have the same structure in all Venetan varieties and they allow for the same featural composition, including [uR] in D. However, building on Lardiere (2008) I suggest the possibility that [uR] is realised in subject clitics in different language-specific ways. This possibility results in different conditions of obligatory, preferred, optional or prohibited environments for subject clitic realisation in the different varieties of the same language.

The difference between the heritage and Italian Venetan systems challenges the idea that heritage languages should be analysed as part of a continuum with their respective homeland varieties (Polinsky 2018). Heritage Venetan varieties display independent discourse-related properties and therefore allow for a different distribution of [uR] on subject clitics. This does not exclude the possibility that contact may play a role in shaping this change, but the notion of transfer from the dominant language seems too strong in this case: there is no "simplification of costly interface

178 The syntax of subject pronouns in heritage languages

processes” in Argentinian and Brazilian Venetan, as it would be expected in Montrul’s analysis.

One point that I leave open for future research is therefore the role of innovation, rather than simplification, in heritage grammars, possibly under the effect of language contact. Ideally, contact can be intended as a fundamental factor, leading to the creation of new conditions applied to existing forms and structures of a language: these conditions may potentially result in more complex or articulated systems.

4.6. Conclusion

This chapter presented a study that aimed to test Carminati’s PAH in Venetan, in particular on subject clitics. The study represents a first application of the model presented in Chapters 2 and 3 and allowed to refine the proposal on the distribution of different pronominal forms in heritage varieties. It was shown that subject clitics in Venetan mainly behave as pronouns and select an antecedent that is not in Spec-TP but in some cases allow also for an interpretation akin to that of null subjects and select an antecedent in Spec-TP.

I proposed that antecedent selection preferences of subject clitics depend again on the presence of [uR]: subject clitics encoding [uR] prefer an antecedent that is not in the matrix Spec-TP, while subject clitics lacking [uR] prefer an antecedent in matrix Spec-TP. It was shown that in Italian Venetan, subject clitics can be of both types. This difference is strictly dependent on the type of subordinate in which the subject clitic is realised; It was also shown that accessibility (redefined in terms of c-command and salience structure of the sentence) is responsible for the constraint on the interpretation of subject clitics in complement *that*-clauses.

Finally, this chapter reflected more extensively on the fact that Venetan, as well as the other languages discussed in previous chapters, are always spoken in contact with other languages. Building on Lardiere (2008), it was proposed that the distribution of [uR] in the three varieties of Venetan depends on feature-reassembly: this proposal allows for the different distributions displayed by the varieties under

The position of the antecedent of pronominal, clitic and null subjects 179

analysis, capturing at the same time the role of innovation, rather than simplification, in heritage languages.

Chapter 5. Decomposing Topic: the Salience Structure Hypothesis

5.1. Introduction

In this chapter, the discussion on the role played by the salience structure in antecedent selection is extended to cases of reference to non-canonical subjects. The present chapter presents an alternative to Carminati's PAH that allows for a better understanding of the way the relationship with the preferred antecedent is established, capturing a wider number of syntactic configurations. It will be shown that subject clitics are sensitive to a number of factors beside the syntactic one, such as the salience structure of the sentence and the argument structure of the verb: the interaction of all these factors leads to fine-grained restrictions on antecedent selection of subject clitics. Strong pronouns, on the other hand, have a distribution that is generally ruled by more pragmatic factors, in that they are always preferably used in specific discourse-related environments, such as switch reference or obviation, no matter the syntactic structure in which the preferred antecedent is used. Such considerations challenge the PAH; specifically, Carminati (2002) cannot account for cases in which there is no preverbal antecedent.

This chapter introduces the Salience Structure Hypothesis, according to which referential ambiguity is solved at the pragmatic level by means of the salience value of potential antecedents, established via Context Scanning. This hypothesis, in turn, provides further support to the approach to different types of pronouns: the realisation of the discourse feature [uR] on pronouns is strictly connected to the distinction between salient and non-salient antecedents, as it is licensed on pronouns only when some update to the salience structure (such as the ones implied in obviation and switch reference) is required.

182 The syntax of subject pronouns in heritage languages

Recall that in the analysis presented here, I consider those discourse-related properties that have an effect on syntax, adopting Miyagawa's (2005, 2010, 2017) Strong Uniformity Principle: discourse-related and information-structural factors are encoded in syntax by means of discourse features that in some languages play a role in syntactic operations.

The chapter is organised as follows: Section 5.2 builds on Frana (2007) and propose the Saliency Structure Hypothesis; Section 5.3 presents the study conducted to test the hypothesis; Section 5.4 analyses the results of the study and show that they fit the model proposed in the previous chapters; in Section 5.5 I return to the discussion of contact, already introduced in Chapter 4, and propose that the data showed in the current chapter provide further support for the notion of feature-reassembly in contact. Section 5.6 concludes the chapter.

5.2. Challenging the PAH: the role of discourse structure in antecedent selection

One of the main assumptions of Carminati's PAH is that the preverbal subject position and the postverbal object position are linked to the information structure of the sentence. The topic is normally found preverbally, in Spec-TP; conversely, the postverbal constituent generally encodes new information or focus. As shown in Chapter 4, the predictions made by the PAH hold strongly in Venetan varieties. However, at this point, it is worth considering antecedent selection in those constructions that are not captured by the PAH or in which the PAH does not hold so strongly. Consider example (1):

(1) Italian (Carminati, 2002: 162)

Dopo che è tomato a casa Lucio, *pro*/*lui* è molto più calmo di prima.
after that be.3SG returned at home Lucio pro he be.3SG much more calm than earlier
'Since Lucio has come back home, he is much calmer than earlier.'

The results of Carminati's (2002) self-reading task showed that *pro* can select the postverbal antecedent, but the reading time for the sentence with a null subject is longer than for the sentence with an overt subject pronoun. Carminati's results simply show that the processing of *pro* in this context is more costly, but they do not specifically answer the question on antecedent preferences or on the factors involved in it. This chapter aims to apply the same methods used in the study presented in Chapter 4 and analyse the preferences of different types of subject pronouns with respect to postverbal antecedents. To do this, the factors involved in antecedent selection and the possible consequences for the model of [uR] need to be defined first.

In Chapter 4, the study by Kaiser and Trueswell (2008) was introduced. They propose a multiple-factor approach: both syntactic and discourse-related constraints contribute to a referent's salience and to the selection of an antecedent, so different types of pronouns can exhibit different sensitivity to different factors.

This approach is supported by Frana (2007), who focusses on the selection of topical antecedents. Building on Carminati's version of the PAH, Frana proposes that *pro* in Italian is preferably linked to a constituent in Spec-TP, as in (2), precisely because Spec-TP is the usual location of the most salient referent in the discourse: the topic. That the preference of null subjects is for topics, and not simply for preverbal subjects, is shown by the fact that *pro* will select a topical antecedent even when the topic no longer corresponds to the element in Spec-TP, as shown in (3).

(2) Italian (Frana 2007: 7)

a. Francesca_j sta cercando casa a Roma.

Francesca be.3SG looking house in Rome

b. Dopo che **Cristina**_k ha mostrato l' appartamento a Francesca_j, **pro**_k ha parlato per mezz'ora.
 after that Cristina has shown the flat to Francesca pro has talked for half hour
 'Francesca is looking for a flat in Rome. After Cristina showed the flat to Francesca, she talked for half an hour.'

184 The syntax of subject pronouns in heritage languages

(3) Italian (Frana 2007: 7)

a. **Francesca**_j sta cercando casa a Roma.

Francesca be.3SG looking house in Rome

b. Dopo che Cristina_k le_j ha mostrato l' appartamento, **pro**_j ha parlato per mezz'ora.

after that Cristina her has shown the flat pro has talked for half hour

'Francesca is looking for a flat in Rome. After Cristina showed her the flat, she talked for half an hour.'

As predicted by the PAH, the null subject in (2)b) selects the preverbal matrix subject *Cristina*, as an antecedent. However, the null subject in (3)b) selects *Francesca*, the topic of the previous sentence, as an antecedent. Such an effect is not predicted by Carminati's PAH. Frana therefore proposes that antecedent selection is driven by the prominence in discourse, rather than the position in the syntactic structure and she puts forward the hypothesis in (4):

(4) Discourse-Prominence Hypothesis of Antecedent Assignment (Frana 2007: 289)

In case of referential ambiguity, the preferred antecedent for null subjects is the most prominent discourse referent available.

The predictions made by the Discourse-Prominence Hypothesis overlap with the PAH when the topical element is found in Spec-TP.

As far as overt pronouns are concerned, Frana (2007) suggests that their preferred antecedent cannot be a topic. She proposes the Anti-Topic Hypothesis (5):

(5) Anti-Topic Hypothesis (Frana 2007: 291)

The preference of overt subjects for lower antecedents decreases when this position correlates with the topic.

The predictions made by the Anti-Topic Hypothesis, however, are only partially supported by Frana's findings: Italian overt subject pronouns can select lower antecedents even when they correspond to the topic.

(6) Italian (Frana 2007: 17)

a. **Tiziana**_j oggi aveva il compito di matematica.

Tiziana today had.3SG the test of math

b. Dopo che Sabina_k le_j ha passato il compito, **lei**_j si è sentita in colpa.

after that Sabina her.CL have.3SG passed the test she REFL.CL be.3SG felt in guilt

‘Today, Tiziana had a math test. After Sabina passed her the test, she felt guilty.’

Frana assumes that the pronoun *lei* in (6b) can select the antecedent *Tiziana* in (6a) as this is the topic of the sentence, contrary to what she suggested with the Anti-Topic Hypothesis. In the next section, it will be shown that both the Discourse Prominence Hypothesis and the Anti-Topic Hypothesis can be collapsed into a single hypothesis on the role of the salience structure.

5.2.1. *The Salience Structure Hypothesis*

The cases discussed in Frana (2007) can be accounted for by the theoretical model presented in the previous chapters. The two hypotheses she entertained, the Discourse Prominence Hypothesis and the Anti-Topic Hypothesis, are captured by the notion of salience structure, the role of the Logophoric Centre and the presence of [uR] on the pronoun in the subordinate clause.

Consider again example (3a)-(3b): the set Δ of potential discourse antecedents here includes *Francesca* as d_1 , the referent with the highest salience value, and *Cristina* as d_2 , with a lower salience value. As predicted by the PAH, *pro* prefers the most salient discourse antecedent³³.

³³ In (3b) the antecedent *Francesca* is referred to by the dative clitic *le*. Frana takes the presence of the clitic as a sign that *Francesca* is the topic of (3b). While I do not argue against this proposal, I have already shown in Chapter 3 that a dative clitic cannot update the salience value of the referent. I therefore assume that there is no way the salience structure of (3b) can be changed with respect to (3a), where *Francesca* corresponds to the referent with the highest salience structure d_1 . The salience structure of (9b) is $\Delta = \{Francesca, Cristina\}$.

In (2)a)-(2)b) too, the set Δ of potential discourse antecedents includes *Francesca* and *Cristina*. However, in this case *Cristina* is d_1 , while *Francesca* is d_2 . This follows from the fact that the same referent obviously cannot be realised more than once in Δ , as it cannot have two different salience values at the same time; the salience structure in (2)b) is updated with respect to (2)a), so *Francesca* is no longer the most salient discourse antecedent. Therefore, also this case is captured by the notion of salience structure: *pro* still preferably selects the most salient antecedent and the different interpretation depends on the fact that the salience structure was updated.

Similarly, in (6)b), Δ includes *Tiziana* and *Sabina*. Here, *Tiziana* is d_1 , while *Sabina* is d_2 . The fact that the pronoun *lei* can refer to *Tiziana* is problematic for Frana's hypothesis, as it implies that the pronoun prefers a topical antecedent. However, this is not necessarily the case in the approach to salience structure presented in Chapter 3. It was proposed that overt pronouns can refer to salient antecedents, a fact that is reflected in their internal structure and in the distribution of [uR]. At the same time, the notion of preference leaves open the possibility for the pronoun to be referentially specific enough to switch the reference to a less salient antecedent: that is the case of strong pronouns, encoding [uR].

There is no need to formulate two separate hypotheses; it is proposed that Frana's Discourse Prominence Hypothesis and Anti-Topic Hypothesis can be collapsed into a single hypothesis, which I call the Salience Structure Hypothesis:

(7) Salience Structure Hypothesis

In case of referential ambiguity:

- a. the preferred antecedent for null subjects is the referent with the highest salience value;
- b. overt subject pronouns can select either referent, depending on the presence or absence of a [uR] feature.

The hypothesis in (7) will be tested by means of the study presented in Section 5.3. Recall that the salience structure of a sentence is relevant for the study of subject realisation in consistent null-subject languages, in that it establishes the contexts in

which the discourse-related feature [uR] is allowed on subject pronouns, allowing or disallowing their overt realisation. Before moving to the study on Venetan subject pronouns, the notion of topicality is discussed and the main proposals presented in the previous chapters with respect to [uR] and the structure of pronouns are briefly repeated.

5.2.1.1. A note on the notion of topicality

So far, topicality was only referred to as in rather vague terms, implicitly assuming that salient antecedents are also topical. While this is correct for the contexts analysed so far, the opposite is not always true: topics are not always salient.

Here, “topic” needs to be intended as a cover term for different notions related to the information structure of the sentence and including a number of possible “topical” interpretations. Topic is generally defined as “what the sentence is about”, the “old information” (Reinhart, 1982; Erteschik-Shir, 2006). Reinhart (1982) integrated the notion of topic as old information with the notion of Common Ground discussed in Chapter 1: the information shared by speaker and addressee (Stalnaker 1974) and the way this information is modified and updated (Krifka 2008). According to Reinhart (1982), the concepts of old and new information present in the Common Ground are associated with entities. Reinhart compares the topic to the title or the heading of a file, under which new information is added and stored. Similarly, the topic constituent identifies an entity or set of entities, to which new information is added. However, this definition does not necessarily imply that the topic refers to old information. This is indeed the case for Familiar/Given Topics (Givón, 1983), which are generally used for topic continuity, but the picture is complicated by other types of topics, such as Aboutness-Shift and Contrastive Topic. As already described in Chapter 2, the Aboutness-Shift Topic has the discourse function of introducing a new topic, thus not specifically referring to old or salient information; Contrastive Topic has the function of inducing alternatives to previously introduced discourse referents (Büring, 1999).

188 The syntax of subject pronouns in heritage languages

The different types of topics were captured in the cartography of topic phrases discussed in the seminal work by Rizzi (1997) and later in Frascarelli and Hinterhölzl (2007). Frascarelli and Hinterhölzl (2007) propose the existence of three dedicated projections in the C-domain, each one hosting a different type of topic: ShiftP hosting Aboutness-Shift Topic, ContrP hosting Contrastive Topic and FamP hosting Familiar/Given Topic:

(8) (adapted from Frascarelli and Hinterhölzl, 2007)
[ForceP [ShiftP [FocP [ContrP [FamP [FinP [IP ...]]]]]]]]

This hierarchy of topic phrases in the C-system correlates with specific discourse functions assigned to each topic type. Frascarelli and Hinterhölzl (2007) show that such distinctions can be identified on the basis of intonational and syntactic evidence.

Working in the same direction, albeit in a feature-based approach, López (2009) argues that topic is not a primitive feature, but a particular amalgam of features that may surface in the grammar of some languages. Following López, the idea that the notion of topic needs to be abandoned and replaced by (or reshaped as) a notion which is able to systematically link the properties of different pronominal forms discussed in this study to their co-reference with an antecedent. Following this idea, it was proposed that an [R] feature needs to be introduced to capture different aspects of the interface between syntax and the information structure of the sentence. This approach does not exclude that more features are needed to capture a wider range of phenomena.

5.2.2. A further application of [uR]

The Saliency Structure Hypothesis proposed in Section 5.2.1 allows to extend the model presented in the previous chapters to the environments described in Frana (2007). The way in which [uR] is licensed on pronouns in these contexts is indeed not different from what was proposed in Chapter 4.

Decomposing Topic: the Saliency Structure Hypothesis 189

Consider again examples (3) and (6), repeated below as (9) and (10). The two sentences have the same saliency structure; however, *pro* in (9) in the main clause refers to the subject of the preposed subordinate clause; the overt pronoun *lei* in (10) can also refer to the previously introduced referent *Tiziana*.

(9) Italian (Frana 2007: 7)

a. **Francesca**_j sta cercando casa a Roma.

Francesca be.3SG looking house in Rome

b. Dopo che Cristina_k le_j ha mostrato l' appartamento, **pro**_j ha parlato per mezz'ora.

after that Cristina her has shown the flat pro has talked for half hour

'Francesca is looking for a flat in Rome. After Cristina showed her the flat, she talked for half an hour.'

(10) Italian (Frana 2007: 17)

a. **Tiziana**_j oggi aveva il compito di matematica.

Tiziana today had.3SG the test of math

b. Dopo che Sabina_k le_j ha passato il compito, **lei**_{j/k} si è sentita in colpa.

after that Sabina her.CL have.3SG passed the test she REFL.CL be.3SG felt in guilt

'Today, Tiziana had a math test. After Sabina passed her the test, she felt guilty.'

The distribution of *pro* and overt pronouns in this context is easily captured by the model put forward in previous chapters. I briefly recap it here.

The pronoun *lei* in (10)b) may or may not encode an uninterpretable [uR] feature and therefore may display preferences similar to those of *pro*; the possibility of having overt pronouns lacking [uR] was already discussed for Italian in Chapter 4. Recall that [uR] is an optional discourse-related feature that makes the pronoun referentially specific enough to switch the reference to a non-salient antecedent. This interpretation is allowed for the pronoun *lei* in (10)b) too, which implies that the pronoun encodes [uR]. In this case, R-checking takes place: in case of switch reference or topic shift, the Logophoric Centre in CP is assigned an interpretable [R] feature via Context

190 The syntax of subject pronouns in heritage languages

Scanning. The [R] feature on the Logophoric Centre R-checks and deletes the [uR] feature on the pronoun.

(11) [R]-checking

An uninterpretable feature [uR] on a syntactic object Y must be checked against an interpretable [R] feature on syntactic object Z c-commanding Y.

As already discussed in Chapter 2, [R]-checking follows from a downward licensing configuration reminiscent of the definition of (Reverse) Agree presented in Wurmbrand (2010) and Zeijlstra (2012).

Therefore, in (10)b), the [uR] feature on the pronoun *lei* is checked and deleted by the [R] feature on the Logophoric Centre and the derivation can proceed as discussed in Section 2.4.3.

For *pro* in (9), in Chapter 2 it was proposed that it does not carry any [uR] feature. This implies that, at the discourse level, there will be no update in the salience structure: *pro* refers to the most salient discourse antecedent. It was also proposed that *pro* is simply a phonological null counterpart of an overt pronoun lacking [uR]. The reason why the pronoun can be null is captured by the notion of *PF-deletion under feature identity* (Sheehan 2006; Roberts 2009), a principle of economy that deletes copies of identical feature bundles at PF. In this case ϕ -features of the pronoun are identical to those of T and therefore the pronoun gets deleted at PF under feature identity, resulting in *pro*.

5.2.3. Summary

In this section, it was shown that the PAH, as proposed by Carminati (2002) is not sufficiently strong to capture configurations different from the canonical one, with a preverbal subject in Spec-TP and a postverbal (indirect) object. Frana (2007) correctly proposes that wider considerations on discourse prominence and topicality need to be taken into account and that it is not possible to reduce antecedence preferences simply to the position of the referents. However, Frana's approach fails to correctly account

for the preferences of overt strong pronouns. Building on her approach, it was proposed the Saliency Structure Hypothesis, which reconciles the different hypotheses she presented with the approach to saliency structure put forward in the previous chapters. The distribution of overt and null forms is therefore captured by the analysis of the internal structure of pronouns and by the distribution of a discourse-related feature [uR], as discussed in Chapter 2.

5.3. The study

The focus of the study is the interpretation of different pronominal forms in Venetan, a language that has three types of subject pronouns: strong pronouns, null subjects and subject clitics. This chapter discusses the second part of the study on antecedent selection in Venetan that was already discussed in Chapter 4. Recall that three varieties of Venetan are considered in the study: Italian Venetan, Brazilian Venetan and Argentinian Venetan.

This part of the study focusses on the selection of non-canonical subjects as antecedents. More specifically, I refer to the selection of preverbal dative experiencers of psych verbs and postverbal subjects of unaccusative verbs by the three types of pronominal subjects available in Venetan.

5.3.1. Research questions and hypotheses

The first goal of this study is to test the Saliency Structure Hypothesis and to compare its predictions to the ones put forward by Carminati (2002) with the PAH. In particular, if the hypothesis is correct, similar preferences should emerge for each type of pronoun in all tested contexts: regardless of the position of the antecedent, the selection depends on the saliency structure on the sentence, which is translated in syntax by means of the distribution of [uR] on pronouns.

The second point regards more specifically the difference between strong pronouns and subject clitics. As it was already shown in Chapter 4, subject clitics allow for a strong pronominal interpretation, which is restricted to a single context in

192 The syntax of subject pronouns in heritage languages

Italian Venetan and generalised in heritage Venetan varieties. The question of whether [uR] behaves exactly the same way in strong pronouns and subject clitics was not addressed in Chapter 4. Therefore, the present chapter aims to show that different pronominal forms make use of the same feature in different ways, in that syntactic and discourse-related factors are involved at different degrees in the distribution of the feature on the two types of overt pronouns. This hypothesis supports Kaiser and Trueswell's (2008) multiple-factor approach.

It was already shown that the realisation of [uR] is optional, a fact that allows to capture the notion of preference, as well as the distribution of subject clitics in the different varieties of Venetan. Here the possible differences in the realisation of [uR] on strong and clitic pronouns are investigated.

In order to test these hypotheses, two types of non-canonical subjects were used: preverbal dative experiencers of psych verbs and postverbal subjects of unaccusative verbs. The tasks and the questionnaire are described in the next section.

5.3.2. The questionnaire

65 informants (Argentina, n=3; Brazil, n=25; Italy, n=37) took part in the study. The details about the informants can be found in Appendix A.

Recall that the questionnaire consisted in a preference task with 36 items made up of one sentence and three possible answers. Participants were asked to indicate which interpretation of the pronominal form in the main clause they preferred with respect to the potential antecedents in a preposed subordinate clause. One of the two potential antecedents was a non-canonical subjects (a preverbal dative experiencer of a psych verb or a postverbal subject of unaccusative verb). Participants could also judge the proposed sentence as ungrammatical. Informants were asked to choose their preferred option, among the two possibilities given.

Three different items per context were used. Each item was manipulated with respect to the pronominal form it included: a strong subject pronoun, a subject clitic or a null subject. As in the case of the data presented in Chapter 4, participants were not given extra clues about the context in which the sentences were uttered, to favour

a ‘plain intonation’ reading and not a contrastive one. The complete list of items used in the questionnaire can be found in Appendix B. The two tested contexts are described in the following section.

5.3.2.1. *Preverbal dative experiencers*

Previous studies on Italian subjects (Belletti and Rizzi 1988, Cardinaletti 1997) showed that a dative experiencer in preverbal position functions as a structural subject, even though the verb displays agreement with the postverbal nominative constituent. Consider the example (12):

(12) Italian (Cardinaletti 1997: 83)
 A Gianni piaceva molto la musica.
 to John liked.3SG a lot the music
 ‘Gianni liked music a lot.’

It is assumed that dative experiencers like *A Gianni* in (12) are not dislocated in the left periphery, but occupy Spec-TP. This assumption is supported by the fact that, in extraction contexts, dative experiencers in Italian cannot be doubled by a clitic (13)³⁴. Conversely, other types of datives, such as goal datives (14) are dislocated in the left periphery, as shown by the fact that they can be doubled by a clitic:

(13) Italian (Cardinaletti, 1997: 84)
 Il periodo in cui a Gianni (*gli) piaceva la musica rock.
 the time in which to John him.CL liked.3SG the music rock
 ‘The time in which John liked rock music.’

³⁴ Consider that there is variation with respect to the acceptability of this example in Italian. Some speakers accept sentences with double experiencers doubled by a clitic.

194 The syntax of subject pronouns in heritage languages

(14) Italian (Cardinaletti, 1997: 84)

Il periodo in cui a Gianni gli studenti gli hanno parlato di questo.
the time in which to John the students him.CL have.3PL talked of this
'The time in which the students talked to John about this.'

This fact supports the idea that the dative experiencer in (13), unlike the goal experiencer in (14), is not dislocated in the left periphery, but rather occupies the preverbal subject position; the peculiar role of dative experiencers of psych verbs makes them particularly relevant in this study. If the Saliency Structure Hypothesis is on the right track, it should be possible to notice similar preferences by the different pronominal forms in the selection of an antecedent.

The hypothesis was tested by means of items consisting of a preposed subordinate clause including a psych verb with a preverbal dative experiencer and a postverbal nominative subject, followed by the main clause, in which one of the three pronominal forms (strong, clitic or null) appeared.

(15) Venetan

Par via che ala Bruna no ghe piaze la Maria, **ela/la/ pro** serca de evitarla.
Since that to.the Bruna not her.CL like.3SG the Maria she she.CL pro try. 3SG to avoid.her.CL
'Since Bruna does not like Maria, she tries to avoid her.'

If saliency structure plays a role in driving antecedent selection, it is plausible that antecedent preferences of different types of subject pronouns will not differ from the ones described in Chapter 4: *pro* should prefer a pre-verbal dative antecedent because of its higher saliency value; an overt subject pronoun should favour the retrieval of the postverbal nominative antecedent in the matrix clause. Notice that in this context, the predictions made by the Saliency Structure Hypothesis should be in line with those of the PAH and the Discourse-Prominence Hypothesis.

5.3.2.2. *Postverbal subjects of unaccusative verbs*

In the second context of non-canonical subjects tested, no overt preverbal antecedent was available to the pronouns. This way it is possible to ascertain that the predictions made by the Saliency Structure Hypothesis are correct: the antecedence preferences of pronouns do not depend on the preverbal or postverbal position of the antecedents, as predicted by Carminati's PAH, but on the saliency structure of the sentence, which is encoded in syntax with [uR].

For this context, unaccusative constructions with postverbal subjects (16) were used:

(16) (Burzio 1986: 21)

Arriveranno molti esperti.

Arrive.FUT-3PL many experts

'Many experts will arrive.'

As observed by Perlmutter (1978) and Burzio (1986), subjects of unaccusative and ergative verbs are internal arguments. In Burzio's analysis of Italian ergative verbs, postverbal subjects like (16) receive a theme or patient theta-role from the verb and are therefore interpreted as 'semantic objects'. Carminati (2002) shows that there are interpretive differences between sentences with preverbal and postverbal subjects, in that the latter generally do not act as the topic of the sentence. Following Lambrecht (1994), she proposes that unaccusative constructions with postverbal subjects are rather interpreted asthetic sentences, reporting or presenting an event, rather than predicating something about a topic.

Hence, the PAH predicts that null subjects should not display a preference for postverbal subjects as antecedents: there is no overt preverbal antecedent and the postverbal antecedent available is not even a topic. Carminati (2002) indeed showed that speakers of Italian hardly choose the postverbal subject as an antecedent for a null subject.

However, contrary to what is expected in Carminati's approach, the Saliency Structure Hypothesis predicts that antecedence preferences depend on the saliency structure of sentences, defined as a set \mathcal{A} of potential discourse antecedents, rather than on the position of the antecedent with respect to the verb. The Saliency Structure Hypothesis does not make any prediction on the topical status of the antecedent. Recall that, in Section 5.2, it was proposed that salient antecedents are generally topical, but not all topical antecedents are salient in the discourse.

To provide the same level of ambiguity as in previously tested contexts, a parallel preference task was used: the pronoun in the main clause has two potential antecedents in a preposed subordinate clause.

(17) Venetan

Quando che ze rivà Luca a casa de Marco, **lu/el/ pro** ga magnà la polenta.
 when that be.3SG arrived Luke at home of Mark he he.CL pro have. 3SG eaten the porridge
 'When Luke arrived at Mark's, he ate porridge.'

The preposed subordinate clause with two possible antecedents is followed by the main clause, in which the pronominal form appears. Notice that both potential antecedents are postverbal, the first being an argument of the unaccusative verb (the postverbal subject *Luca*) and the second one being an adjunct (*Marco* in the PP). By means of this preference task, the preferences of null subjects for postverbal and preverbal antecedents can be compared and the Saliency Structure Hypothesis can be tested. The prediction is that, if the saliency structure plays a role in the selection of an antecedent, the fact that the element with the highest saliency value is not in Spec-TP should not matter for interpretation.

5.3.3. Results

This section presents the results of the study relative to the two contexts discussed in this chapter: preverbal dative experiencers of psych verbs and postverbal subjects of unaccusative verbs. One-way ANOVA (significance: $p = <.05$) with subjects as

random effect was conducted to compare the effect of pronoun type (pronominal, clitic or null) on the selection of the antecedent in the two contexts. Results will show that the Saliency Structure Hypothesis correctly predicts the preferences of different types of pronouns.

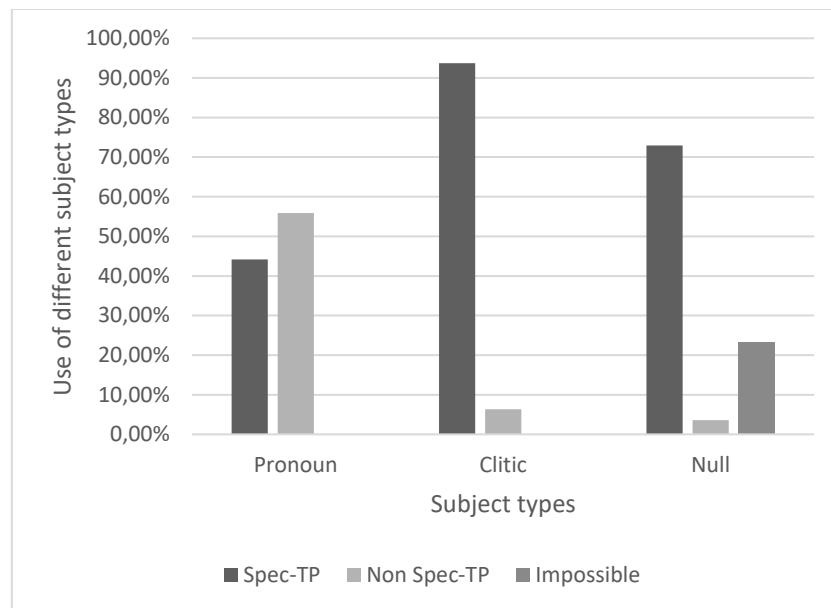
5.3.3.1. Preverbal dative experiencers of psych verbs

In the case of preverbal dative experiencers of psych verbs, the Saliency Structure Hypothesis holds in Venetan. Recall that in this context, the predictions made by the Saliency Structure Hypothesis can be directly compared to those made by Carminati's PAH, as there is a potential preverbal antecedent in Spec-TP and a postverbal one that is not in Spec-TP. As predicted by the PAH, the preference of overt subjects for the postverbal antecedent are not as strong as in the two contexts considered in Chapter 4. The three Venetan varieties are discussed separately.

Italian Venetan shows that strong subject pronouns have a marginally significant preference for the postverbal antecedent. However, subject clitics and null subjects display a clear preference for the dative antecedent in Spec-TP. All 37 informants carried out the online questionnaire; each informant had to select the preferred interpretation for 9 sentences with preverbal dative experiencers (3 per pronominal form). In total, 333 sentences were analysed; the results are presented in Figure 1.

198 The syntax of subject pronouns in heritage languages

Figure 1. Preverbal dative experiencers in Italian Venetan. Number of informants: 37. Number of tested sentences per informant: 9 (3 per pronominal form). Total of analysed sentence: 333.



A One-way ANOVA was performed on the data, as in the case of the data discussed in Chapter 4. Recall that the statistical analysis followed the methodology used, among others, in Carminati (2002) and Frana (2007), by comparing the results for each pronominal form to the other forms in each group of speakers. The model shows a significant difference between the antecedent preference of strong pronouns and *pro* ($F(3,88) = 78.35, p < .05$): while strong pronouns prefer the postverbal antecedent, *pro* prefers the preverbal dative antecedent. The difference between the antecedent preference of strong pronouns and subject clitics is significant as well ($F(3,88) = 88.83, p < .05$). Conversely, subject clitics and *pro* prefer both prefer the preverbal dative antecedent in the matrix clause: the difference in their antecedent preference is not significant ($F(3,88) = 0.23, p > .05$). The preferred interpretation of strong pronouns, subject clitics and *pro* in Italian Venetan is exemplified in (18):

(18) Italian Venetan

Par via che ala Bruna_j no ghe piaze la Maria_k, **ela_k / la_j / pro_j** serca de evitarla.

Since that to.the Bruna not her.CL like.3SG the Maria she she.CL pro try. 3SG to avoid.her.CL

‘Since Bruna does not like Maria, she tries to avoid her.’

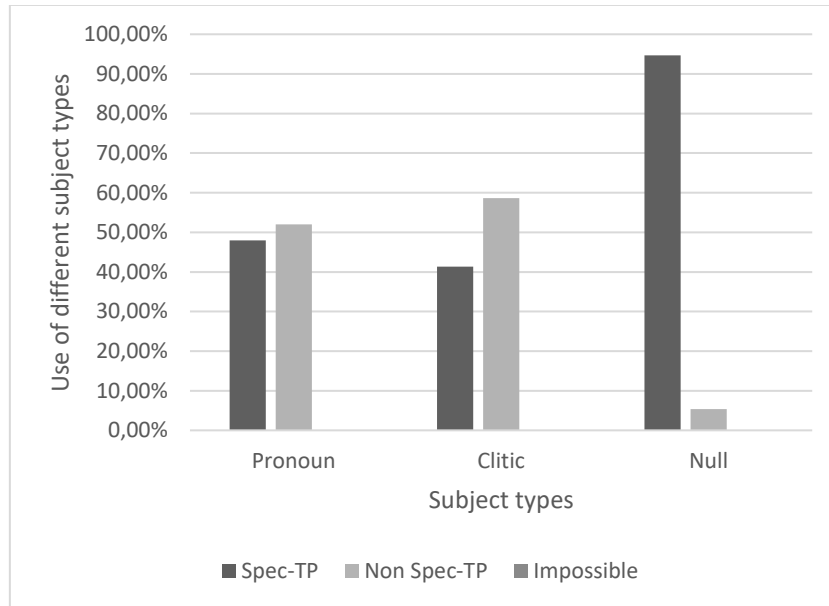
As in the case of the contexts discussed in Chapter 4, also in this case a further One-way ANOVA was performed on items judged ungrammatical in order to verify whether subject clitics are indeed obligatory. Similarly to previous cases, the model did not render a significant effect for any of the forms (*pro*, accepted vs. not accepted: $F(3.08) = 1.25, p = >.05$; strong and clitic subjects were always accepted).

Brazilian and Argentinian Venetan display a similar pattern, which is at the same time different from that of Italian Venetan.

In Brazilian Venetan both overt forms display a marginal preference for the post verbal antecedent, while null *pro* prefers the preverbal dative antecedent. All 25 informants carried out the online questionnaire; each informant had to select the preferred interpretation for 9 sentences with preverbal dative experiencers (3 per pronominal form). In total, 225 sentences were analysed; the results are presented in Figure 2.

200 The syntax of subject pronouns in heritage languages

Figure 2. Preverbal dative experiencers in Brazilian Venetan. Number of informants: 25. Number of tested sentences per informant: 9 (3 per pronominal form). Total of analysed sentence: 225.



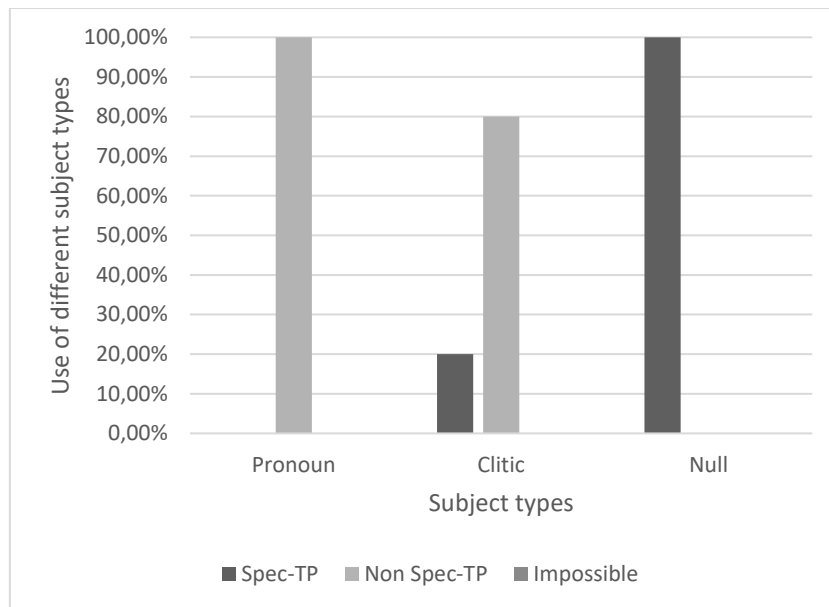
The antecedent preference of strong pronouns and *pro* are significantly different ($F(3,90) = 53.70, p = <.05$). The antecedent preference of subject clitics and *pro* are significantly different too ($F(3,90) = 71.84, p = <.05$). The preference of strong pronouns does not significantly differ from that of subject clitics ($F(3,90) = 0.66, p = >.05$). This pattern implies that subject clitics and strong pronouns behave significantly differently from null subjects in that the two overt forms display a preference towards the postverbal antecedent, while null subjects display a preference towards the preverbal dative antecedent.

None of the Brazilian informants judged “pure” null-subject sentences (with neither a tonic pronoun nor a subject clitic) ungrammatical.

Argentinian Venetan displays the same pattern as Brazilian Venetan. However, the preference of the overt forms for a postverbal antecedent is stronger in Argentinian than in Brazilian Venetan. All 3 informants carried out the online questionnaire; each informant had to select the preferred interpretation for 9 sentences

with preverbal dative experiencers (3 per pronominal form). In total, 27 sentences were analysed; the results are presented in Figure 3.

Figure 3. Preverbal dative experiencers in Argentinian Venetan. Number of informants: 3. Number of tested sentences per informant: 9 (3 per pronominal form). Total of analysed sentence: 27.



The One-way ANOVA confirmed that the antecedent preference of strong pronouns and *pro* is significant ($F(4,49) = 65535$, $p < .05$); the difference between subject clitics and *pro* in Argentinian Venetan is significant as well ($F(4,49) = 16$, $p < .05$); finally, the difference in the preferences of strong pronouns and subject clitics is not significant ($F(4,49) = 4$, $p > .05$). In Argentinian Venetan too, both subject clitics and tonic pronouns display a preference towards the postverbal antecedent in the matrix clause, while *pro* displays a significant preference towards the preverbal dative antecedent. Also in the case of Argentinian Venetan, all informants accepted all the proposed items.

The pattern displayed by heritage Venetan varieties is the one in (19):

202 The syntax of subject pronouns in heritage languages

(19) Heritage Venetan

Anca se a Marco_j no ghe interesa Luca_k, lu_k/el_k/ *pro*_j ghe ga telefonà ieri.
even if to Mark not him.CL interest.3SG Luke he he.CL pro him.CL have.3SG called yesterday
'Even though Mark does not care about Luke at all, he called him yesterday.'

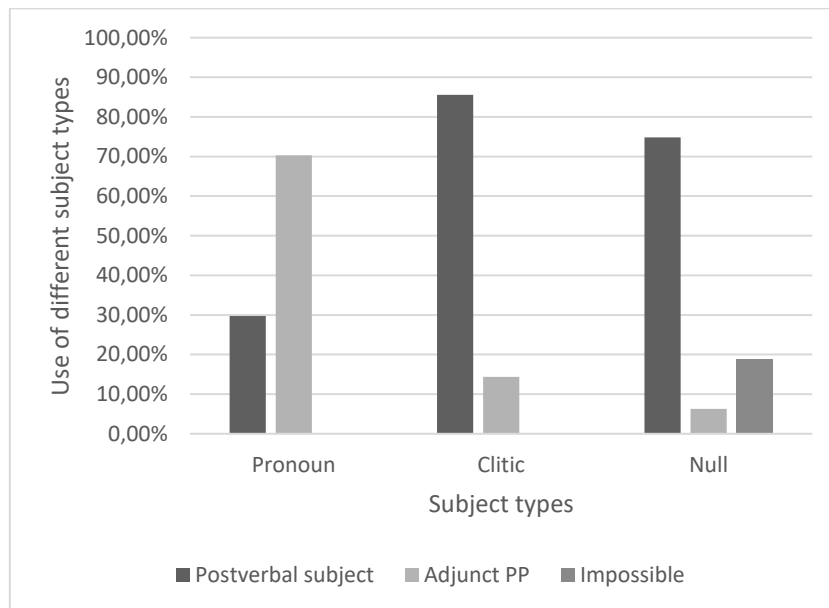
In conclusion, the results for the three varieties of Venetan are in line with the predictions made by the PAH. As far as subject clitics are concerned, in Italian Venetan they prefer the dative antecedent in Spec-TP, while in heritage Venetan varieties they pattern with strong pronouns in the selection of a postverbal antecedent; this difference between Venetan varieties already emerged in Chapter 4.

5.3.3.2. *Unaccusatives constructions with postverbal subjects*

The case of postverbal subjects in unaccusative constructions is crucial to test the Salience Structure Hypothesis. The tested items do not include an overt preverbal antecedent, but two postverbal antecedents: the postverbal subject and an adjunct PP. Therefore, precisely in this context, it should be possible to see the difference between Carminati's PAH (which makes specific reference to the fact that an antecedent needs to be in Spec-TP in order to be selected by *pro*) and the idea that a more general notion of salience structure applies to the selection of an antecedent.

The results of the study on Italian Venetan show that strong pronouns prefer the lower postverbal antecedent in the PP, while *pro* prefers the postverbal subject as an antecedent. As in the previous context, subject clitics pattern with *pro*. All 37 informants carried out the online questionnaire; each informant had to select the preferred interpretation for 9 sentences with unaccusative constructions (3 per pronominal form). In total, 333 sentences were analysed; the results are presented in Figure 4.

Figure 4. Unaccusative constructions in Italian Venetan. Number of informants: 37. Number of tested sentences per informant: 9 (3 per pronominal form). Total of analysed sentence: 333.



The ANOVAs confirmed that the antecedent preference of strong pronouns and *pro* is significantly different ($F(3,89) = 110.65, p < .05$). Subject clitics pattern with *pro*, in that the difference in their antecedent preference is not significant ($F(3,88) = 1.33, p > .05$); the preference of subject clitics is significantly different from that of strong pronouns ($F(3,89) = 103.28, p < .05$).

(20) Italian Venetan

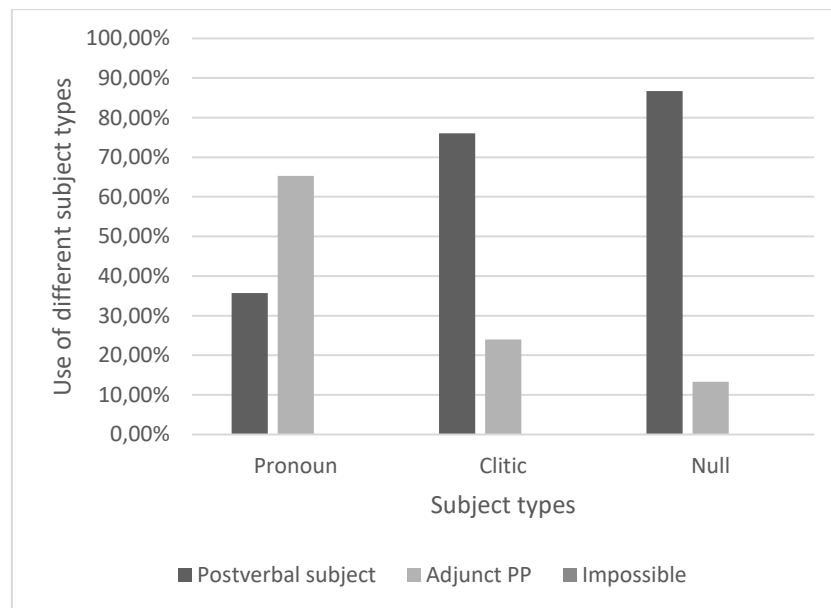
Quando che riva Marco; a casa de Luca, **lu**/**el**/ *pro*; se mete vardar la television.
 when that arrive.3SG Mark at home of Luke he he.CL pro REFL.CL put.3SG watch the television
 ‘When Mark arrives at Luke’s, he starts watching television.’

204 The syntax of subject pronouns in heritage languages

The ANOVAs including items judged ungrammatical by the informants did not render a significant effect (tonic pronouns and subject clitics: always accepted; null subjects, accepted vs. not accepted: $F(3.08) = 0.14, p = >.05$).

In Brazilian Venetan, subject pronouns prefer the antecedent in the adjunct PP, while *pro* prefers the postverbal subject. Unlike all other tested contexts, in the case of unaccusative constructions, subject clitics pattern with *pro*. All 25 informants carried out the online questionnaire; each informant had to select the preferred interpretation for 9 sentences with unaccusative constructions (3 per pronominal form). In total, 225 sentences were analysed; the results are presented in Figure 5.

Figure 5. Unaccusative constructions in Brazilian Venetan. Number of informants: 25. Number of tested sentences per informant: 9 (3 per pronominal form). Total of analysed sentence: 225.

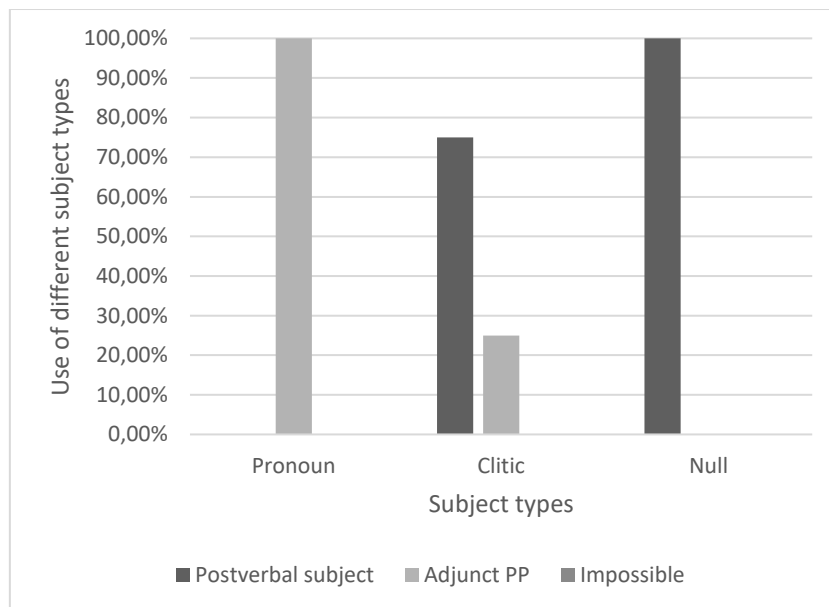


The ANOVAs show that the difference in the interpretation of strong pronouns and *pro* is significant ($F(3.90) = 58.5, p = <.05$). The difference in the interpretation of strong pronouns and subject clitics is significant too ($F(3.90) = 30.91, p = <.05$), while that of subject clitics and *pro* is not ($F(3.90) = 2.82, p = >.05$).

As in the previous context, none of the Brazilian informants judged “pure” null-subject sentences (with neither a strong pronoun nor a subject clitic) ungrammatical, against what is expected on the basis of the widely accepted analysis of subject clitics as obligatory elements.

The same pattern emerges in Argentinian Venetan:

Figure 6. Unaccusative constructions in Argentinian Venetan. Number of informants: 3. Number of tested sentences per informant: 9 (3 per pronominal form). Total of analysed sentence: 27.



The ANOVAs confirmed that the difference in the interpretation of strong pronouns and *pro* is significant ($F(4,49) = 65535$, $p < .05$), and so is the difference in the interpretation of strong pronouns and subject clitics ($F(4,49) = 16.00$, $p < .05$). The difference in the interpretation of subject clitics and *pro*, however, is not significant ($F(4,49) = 4$, $p > .05$). The ANOVAs including sentences judged ungrammatical did not render a significant effect for any of the forms (all proposed sentences were accepted).

The pattern displayed by heritage Venetan varieties is the one in (21):

(21) Heritage Venetan

Quando che vien la Maria_i da so mare_k, **ela_k/la_j** / **pro_j** va subito in leto.
 when that come.3SG the Mary at her mother she she.CL pro go. 3SG immediately in bed
 ‘When Mary comes to her mother’s, she immediately goes to bed.’

In conclusion, in all Venetan varieties, *pro* selects the postverbal subject as the preferred antecedent, while strong pronouns select the lower antecedent contained in a PP. The behaviour of subject clitics is rather different from previous contexts, especially in heritage Venetan varieties: this is the only context in which their preferences pattern with those of *pro*, rather than of strong pronouns.

5.3.3.3. Summary of the results

The data presented in Sections 5.3.3.1 show that the PAH holds in the context of preverbal dative experiencers in all Venetan varieties, although the preference of strong subject pronouns for the postverbal antecedent is not as strong as in the contexts presented in Chapter 4. Conversely, the preference of *pro* for the preverbal dative experiencer is strong. Subject clitics pattern with strong pronouns in heritage Venetan and with null subjects in Italian Venetan.

As far as unaccusative constructions in Section 5.3.3.2 are concerned, recall that the tested items included two potential postverbal antecedent: the postverbal subject of the unaccusative verb and a second referent in an adjunct PP. In all varieties, strong pronouns do not display a preference for the postverbal subject, but for the referent in the adjunct PP. Conversely, in all Venetan varieties, both subject clitics and null subjects prefer the postverbal subject in the matrix clause as an antecedent.

Table 1. The preferences of subject pronouns in Venetan Varieties.

	Dative experiencers	Unaccusatives
Italian Venetan	<i>Subject clitic</i> ≈ <i>Null</i>	<i>Subject clitic</i> ≈ <i>Null</i>
Brazilian Venetan	<i>Subject clitic</i> ≈ <i>Pronoun</i>	<i>Subject clitic</i> ≈ <i>Null</i>

Argentinian Venetan	<i>Subject clitic</i> \approx <i>Pronoun</i>	<i>Subject clitic</i> \approx <i>Null</i>
---------------------	--	---

In conclusion, the preferences of different types of pronouns with respect to preverbal dative experiencers mirror the ones displayed in the contexts presented in Chapter 4. In unaccusative constructions with postverbal subjects the situation is different, especially for heritage varieties: this is the only context in which subject clitics preferences match those of *pro*.

5.4. The saliency structure and [uR]: an extension of the model

The results presented in Section 5.3.3. showed that, as already pointed out by Frana (2007), Carminati's (2002) PAH makes too strong predictions on the link between syntactic positions and antecedent selections. This hypothesis is supported by the fact that, when there is no overt potential antecedent in Spec-TP and both potential antecedents are found postverbally, as in the case of the tested unaccusative constructions, the same predictions hold. Strong pronouns prefer the potential antecedent in the adjunct PP, while null subjects prefer the postverbal subject.

(22) Venetan

Quando che vien la Maria_i da so mare_k, **ela_k / pro_j** va subito in leto.
 when that comes the Mary at her mother she pro goes immediately in bed
 'When Mary comes to her mother's, she immediately goes to bed.'

To capture cases such as (23), Frana proposes the DPH, according to which the preferred antecedent for null subjects is the most prominent discourse referent available.

Section 5.2 extended this reasoning and proposed the Saliency Structure Hypothesis to capture both the preferences of null and overt subject pronouns. Its application to the cases under analysis is discussed in Section 5.4.1.

A second point addressed here is the behaviour of subject clitics, especially in the case of unaccusative constructions. This is the only context in which subject

208 The syntax of subject pronouns in heritage languages

clitics display preferences that mirror those of *pro* in all varieties, including the heritage ones, in which subject clitics otherwise generally prefer an interpretation akin to that of strong pronouns. The availability of both interpretations for subject clitics and their different behaviour with respect to strong pronouns, answers the question on the distribution of [uR] in different types of pronouns: both strong and clitic pronouns make use of [uR], but not exactly in the same way. This is discussed in Section 5.4.2.

5.4.1. *The syntactic encoding of [uR] and salience structure*

The Salience Structure Hypothesis is able to capture example (22), repeated here as (23):

(23) Venetan

Quando che vien la Maria_j da so mare_k, **ela_k / pro_j** va subito in leto.
when that comes the Mary at her mother she pro goes immediately in bed
'When Mary comes to her mother's, she immediately goes to bed.'

In this sentence, both potential antecedents for the subject pronoun in the main clause are found postverbally in a preposed subordinate clause. The preferences of the pronoun *ela* and of *pro* in this context mirror exactly the preferences displayed by the same elements in the other tested contexts, in which a preverbal and a postverbal potential antecedent were found, regardless of their grammatical function, as it was shown for preverbal dative experiences of psych verbs (24):

(24) Venetan

Par via che ala Bruna_j no ghe piaze la Maria_k, **ela_k / pro_j** serca de evitarla.
Since that to.the Bruna not her.CL like the Maria she pro tries to avoid.her.CL
'Since Bruna does not like Maria, she tries to avoid her.'

While the preferences of the pronouns in (24) are consistent with the predictions of the PAH, the ones in (23) are not: there is no available antecedent in Spec-TP; both

pro and the strong pronoun select one of the two potential postverbal antecedents. Interestingly, *pro* still prefers the most salient antecedent, while the strong pronoun prefers the less salient one.

In other words, the relevant factor here is again the saliency value of the antecedent, and not the position of the antecedent, as claimed by Carminati. Antecedent preferences, in the case in which one antecedent is in Spec-TP and the other one is not, follow from more general considerations on the saliency structure of the sentence: the antecedent in Spec-TP has a higher saliency value, while the one not in Spec-TP has a lower saliency value. Saliency values, crucially, do not distinguish between syntactic positions in the sentence.

In Chapter 2 it was proposed, following Von Stechow (2000), that each sentence has a saliency structure that consists of a set of ranked discourse referents carrying different saliency values. Recall that saliency is defined as a context-dependent choice function (Reinhart 1997), by which a referring expression refers to an element d_x of a set $\mathcal{A} = \{d_1 \dots d_n\}$ of possible discourse antecedents associated with the same descriptive content. The function does not constrain the antecedent choice to syntactic positions; it simply orders the potential antecedents starting from the one with the highest saliency value (d_1) to the ones with lower saliency values ($d_{\neq 1}$).

The set \mathcal{A} of possible discourse antecedents in (23) includes two referents: *Maria*, the referent with the highest saliency value (d_1), and *so mare*, the referent with the lowest saliency value (d_2). Recall also that the antecedence relationship between one member of the set and the pronominal form in matrix Spec-TP is mediated by a Logophoric Centre in the matrix CP; at the discourse level, the Logophoric Centre is hence responsible for the selection of an element from the set of possible antecedents; the selection of the first element d_1 in the set implies that there is no update in the saliency structure; in this case, the Logophoric Centre does not license a referentially specific interpretation and the subject will corefer with the most salient discourse antecedent. The selection of another element $d_{\neq 1}$ of the set implies that the saliency value of this element needs to be updated and the Logophoric Centre licenses a referentially specific interpretation, in that reference has to be switched.

210 The syntax of subject pronouns in heritage languages

At the structural level, derivation proceeds similarly to what was shown in Chapter 4. Null *pro* results from deletion of the pronoun under feature-identity at PF (Sheehan 2006, Roberts 2009), as it lacks [uR] on the D-layer; therefore, *pro* will display a preference for an antecedent with a high salience value. Unlike what was proposed in Chapter 4, following Carminati (2002), it is the salience value that defines the antecedent preference of *pro*, and not topicality.

As shown in Chapter 2, *pro* does not have an uninterpretable [uR] feature and the Logophoric Centre also lacks [R], in that the most salient discourse antecedent was selected via Context Scanning. Therefore, [*R*]-checking does not take place.

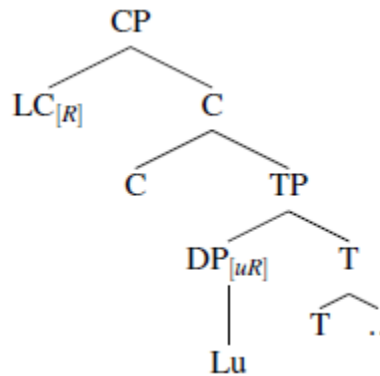
Strong pronouns encode a [uR] feature in D: the pronoun is used to switch the referent to a non-salient discourse antecedent. Recall that in the unaccusative constructions tested, this antecedent is contained in an adjunct PP.

(25) Venetan

Quando che riva Marco; a casa de Luca_k, lu_k se mete vardar la television.
when that arrives Mark at home of Luke he REFL.CL puts watch the television
'When Mark arrives at Luke's, he starts watching television.'

In this case, the φ -composition of the pronoun is not identical to that of finite T, so the pronoun does not get deleted at PF. [uR] on the pronoun has to be checked against [R] on the Logophoric Centre before transfer to LF. The realisation of [R] on the Logophoric Centre results from an assignment function at the discourse level, in the case in which reference needs to be switched to a non-salient discourse antecedent. The structure is repeated in (26):

(26)



As in previous contexts, regardless of the position of the antecedent, [uR] captures the distribution of overt and null subjects: the absence of [uR] triggers the selection of the most salient antecedent; the presence of [uR] allows for the selection of a non-salient antecedent.

The Saliency Structure Hypothesis is therefore correct: it captures the distribution of different pronominal forms with respect to postverbal antecedents in unaccusative constructions. Such cases cannot be captured by Carminati's PAH, whose predictions are too strong. In other words, the PAH appears to follow from more general considerations on the saliency structure of the sentence. Frana's DPH correctly focusses on the fact that *pro* does not display a preference for a subject, but rather for a prominent antecedent. However, the saliency structure of the sentence needs to be taken into account to be able to formulate generalisations also on other types of pronouns.

5.4.2. Subject clitics and non-argument antecedents

In Chapter 4 it was shown that subject clitics in heritage Venetan varieties generally display a preference for the antecedent with the lowest saliency value, on par with strong subjects. This is true for temporal clauses (27) and complement clauses (28) as well as for the context of preverbal dative antecedents (29) discussed in the present

212 The syntax of subject pronouns in heritage languages

chapter. The only exception is the context of unaccusative constructions (30): in this case, the subject clitic in the matrix clause prefers the postverbal subject of the unaccusative verb in a preposed subordinate clause.

(27) Heritage Venetan

Marco; el scrivea sempre a **Luca**_k quando che **el**_k stava mal.

Mark he.SCL wrote always to Luke when that he.SCL was sick.

‘Mark always wrote to Luke when he was sick.’

(28) Heritage Venetan

La Maria; la ga dito **ala Bruna**_k che **la**_k ga da crompar el pan.

the Maria she.SCL has said to.the Bruna that she.SCL has to buy the bread

‘Maria told Bruna that she has to buy bread.’

(29) Heritage Venetan

Par via che ala Bruna; no ghe piaze **la Maria**_k, **la**_k serca de evitarla.

Since that to.the Bruna not her.CL like the Maria she.SCL tries to avoid.her.CL

‘Since Bruna does not like Maria, she tries to avoid her.’

(30) Heritage Venetan

Quando che vien **la Maria**_j da so mare_k, **la**_j va subito in leto.

when that comes the Mary at her mother she.CL goes immediately in bed

‘When Mary comes to her mother’s, she immediately goes to bed.’

At a first glance, it may seem that in all four contexts the subject clitic prefers the same antecedent: the one that is not in Spec-TP. However, the postverbal subject *Maria* in (30) does not have the same salience value as the postverbal antecedents in ((27)-(29)). While in ((27)-(29)) the postverbal antecedent represents d_2 , the discourse antecedent with the lowest salience value, in (30) it represents d_1 , the discourse antecedent with the highest salience value.

A second point that needs to be taken into account is the theta-structure of the preposed subordinate clause in (30): unlike all other types of antecedents tested in this study, the second available antecedent in (30), *so mare*, is not an argument of the verb, but an adjunct. I tentatively propose that the subject clitic does not display a preference for this antecedent, precisely because this is not an argument of the verb³⁵.

A similar constraint in the reference to an antecedent in an adjunct was noticed independently by Samek-Lodovici (1996) for Italian *by*-phrases in passive constructions:

(31) Italian

Dopo che la mostra è stata visitata da Gianni, *pro*_{#j} / *lui*_j è andato all' università.
 after that the exhibition is been visited by John *pro* he is gone at.the university
 'After the exhibition was visited by John, he went to university.'

In (31), the PP *da Gianni* can only function as an antecedent for the overt pronoun *lui* and not for null *pro*.

I propose that the same constraint holds for subject clitics: the argumental status of the antecedent is also a relevant factor in the selection. In the case of (30), subject clitics select the antecedent with the highest salience value, unlike all other contexts, because this is the only argument antecedent available. It is possible that this fact follows precisely from the fact the PP is allowed optionally in the proposed items; in this respect, for example, Reinhart (2000) claims that the optionality of the PP implies that the referent is not associated with a mental state in that context: in this sense, it is possible to assume that adjuncts are less accessible (in Ariel's terms) than arguments at a discourse level.

³⁵ A similar restriction in antecedent selection was described in Haegeman (2019) with respect to *pro*-drop in some written registers of English. Her data showed that, in journalistic prose, subject pronouns can be dropped in certain limited contexts, defined again in terms of accessibility (Ariel, 1990). However, *pro*-drop is banned when there is only one potential accessible antecedent contained in an adjunct PP.

214 The syntax of subject pronouns in heritage languages

At this point, however, the question remains of whether this behaviour of subject clitics is relevant at a syntactic level too. Recall that the aforementioned constraint holds for subject clitics (and *pro*), but not for strong pronouns, as already shown in Samek-Lodovici (1996). Consider example (32):

(32) Venetan

Quando che vien la Maria_j da so mare_k, **ela**_k va subito in leto.
when that comes the Mary at her mother she goes immediately in bed
'When Mary comes to her mother's, she immediately goes to bed.'

The strong pronoun *ela* prefers the antecedent *so mare*, which has a lower salience value and, crucially, is not an argument of the matrix verb.

In other words, it seems that [uR] is allowed only on strong pronouns in this context. Subject clitics, conversely, generally lack [uR], leading to coreference with the most salient antecedent.

As a tentative answer to the question of how this difference is encoded in syntax, I propose that strong pronouns here behave as true logophors, as defined in Reuland (2006) in that their antecedent can occur at any depth of embedding. As far as the restriction on the realisation of [uR] on subject clitics, it is plausible to assume that the antecedent in the PP-adjunct is too deeply embedded for the subject clitic to switch reference to it; this conclusion was already reached by Levin and Rappaport Hovav (2005), who claimed that embedded constituents receive less prominent syntactic realisations. Besides, the PP-adjunct is not in an A-position: the behaviour of subject clitics in this context is reminiscent of Reinhart's (1983) and Heim and Kratzer's (1998) proposal that a pronoun can only be bound by a c-commanding A-position. A more fine-grained examination of this restriction on the realisation of [uR] on subject clitics is a matter of future research.

The present data cannot provide a conclusive answer to the difference in the interpretation of strong pronouns and subject clitics in this context. For the time being, it is only possible to assume that the different interpretation of the subject clitic in (30) and the strong pronoun in (32) depends on the argument structure of the verb, as

proposed by Reinhart (2002), on purely processing factors, as shown for example in Arnold (2001)³⁶ or, alternatively, as a combination of the two: Ramchand (2013) claims that argument structure and processing are often competing factors in determining syntactic behaviours; some languages tend to pay more attention to the cognitive import of referentiality and saliency, while in other languages argument structure plays a major role.

In any event, this difference does not pose any problem for the model of [uR]: being a feature that strictly depends on information structure and on discourse-factors, it is reasonable to allow for the possibility that more refined pragmatic and semantic factors influence the interpretation of subject clitics in this context. Also, I leave open the possibility that more features are involved in the antecedent selection of strong pronouns.

In conclusion, the context of unaccusative constructions with two potential postverbal antecedents³⁷ confirms that both strong and clitic pronouns make use of [uR]; however, it is also clear that the two forms do not make use of [uR] exactly in the same way and, as a result, the distribution of the feature of the two types of

³⁶ In the literature, there is evidence that the distinction between argument and non-argument is relevant for anaphora resolution at the level of language processing. In this respect, Arnold (2001) shows that pronouns generally prefer complements of PPs as antecedents when they are also arguments; conversely, when complements of PPs are not arguments, pronouns hardly select them as antecedents. She proposes that “*language processing involves unconscious hypotheses about where the discourse is going and that these hypotheses influence the activation of discourse referents. The listener’s predictions about the discourse flow are neither conscious nor categorical. Rather, certain referents are activated probabilistically, for a short period of time, as various kinds of information become available. Activation is influenced by many sources of information, including the thematic roles of referents in the preceding clause.*” While it was possible to explain the different interpretations in the contexts presented in Chapter 4 in syntactic terms as a restriction on the licensing of [uR] in complement clauses, the role played by syntax in subject clitics preferences with respect to postverbal subjects of unaccusative constructions is less clear.

³⁷ Recall that the sentences are not pronounced with special intonation or emphasis on the pronouns, in order to avoid a contrastive or emphatic interpretation.

pronouns is not exactly the same. This difference has been attributed to the fact that the non-argumental status of the less salient antecedent makes it more difficult for the subject clitic to select it as a preferred antecedent. Finally, this last part of the study confirmed that multiple factors are involved in the process of antecedent selection. As previously showed by Kaiser and Trueswell (2008), syntactic role, information structure, semantics of the verb, word order and processing all contribute in different, language-specific ways to reference resolution.

5.5. Further consequences for the analysis of contact-induced change

This last section, goes back to what was discussed in Chapter 4 with respect to contact and feature-reassembly. I built on Lardiere (2008) and suggested that [uR] is licensed on subject clitics following language-specific configurations. Therefore, the obligatory, preferred, optional or prohibited environments for the realisation of [uR] on subject clitics vary across different varieties of Venetan.

The same difference emerges in the contexts presented in this chapter, especially for unaccusative constructions. However, in this case, heritage varieties are the ones that allow for a different distribution of [uR] on subject clitics. Recall that Argentinian and Brazilian Venetan speakers generally tend to interpret subject clitics on par with strong pronouns and prefer the less salient potential antecedent; however, when both antecedents are postverbal and one of them is the complement of a PP, subject clitics prefer the most salient antecedent, similarly to *pro*.

In other words, the system of heritage Venetan varieties does not display a simplification of the system, but rather a complexification, in that subject clitics in Italian Venetan do not display the same sensitivity to the argument status of the antecedent.

5.5.1. Innovation and complexification in heritage languages

Most studies on heritage languages focussed on changes that led to reduction of complexity, convergence with the dominant language or levelling with other varieties of the heritage language (see Aalberse and Muysken, 2013, for an overview). Fewer studies (Trudgill, 2011; Shin, 2014; Andriani *et al.*, 2022) discussed the complexification of the heritage system as a possible outcome of language contact. In Chapter 1, complexification was defined as the case in which additional constraints are added to existing structures in heritage languages; this is what happens in heritage Venetan with respect to subject clitics. There are two possible explanations for the complexification of subject clitics in heritage Venetan: the generation of the speakers and more general considerations on referential ambiguity.

With respect to the generation, Heine and Kuteva (2003) claim that there is a limited number of attested cases of complexification in heritage languages because this process cannot be completed in the span of one generation: many generations are necessary to complete it. One possible factor influencing the system of heritage Venetan subject clitics is therefore the generation of the informants: recall that, particularly in the case of Brazilian Venetan, the vast majority of the informants are third or fourth generation heritage speakers. It is therefore reasonable to expect that heritage Venetan displays a completed change in contact, which is not possible to observe in first-generation heritage speakers.

The second factor that can explain the emergence of a more complex system is language-internal and it depends on the referential ambiguity, as already shown in Shin (2014). The items proposed in the questionnaire were intentionally left contextually ambiguous to check whether pronouns display a bias for a given antecedent regardless of the context in which the sentence is uttered³⁸. The same level of referential ambiguity (with two potential antecedents in the immediately preceding discourse) was maintained for all tested contexts. Shin (2014) studied the realisation

³⁸ This property is defined by Carminati (2002) as “global ambiguity”, in that the context does not give the informant any clue as to which the correct interpretation of the pronoun is. In this study, global ambiguity was applied to all proposed items.

218 The syntax of subject pronouns in heritage languages

of subject pronouns in heritage Spanish speakers and showed that complexification may emerge precisely when there are more potential antecedents for the pronoun in the immediate previous discourse. The results of the study presented in Section 5.3 showed that this is also the case of heritage Venetan.

The behaviour of subject clitics in heritage Venetan provides a further piece of evidence in the direction of an analysis of heritage languages as independent systems that develop their own constraints for the way features are assembled, as already proposed by Lardiere (2008). One important point in Lardiere's analysis consists in the fact that some features (normally uninterpretable ones) are assembled and realised idiosyncratically in each language and this obviously plays a role in determining the amount of variation displayed by different languages with respect to a phenomenon. In previous chapters the 'optionality' of [uR] is considered as a fact that follows from its discourse-related nature. At a syntactic level, it was also proposed that [uR] participates in syntactic operations in a fashion that does not exactly parallel that of morphosyntactic features (*contra* Miyagawa, 2005). This was shown, for instance, by the fact that [R]-checking is implemented by means of a (Reverse) Agree operation of the type described in Wurmbrand (2010) and Zeijlstra (2012)³⁹. Feature-reassembly is therefore favoured by the idiosyncratic conditions on the realisation of [uR] and on the fact that its encoding in syntax does not exactly parallel that of purely morphosyntactic features.

5.6. Conclusion

In this chapter, it was shown that the PAH is challenged by the cases in which the discourse role of the antecedents of pronominal forms does not match their syntactic position.

Building on multiple-factor approaches to antecedent selection (Kaiser and Trueswell, 2008; Frana, 2007), it was proposed that preferences of all types of

³⁹ A similar proposal on "downward" anaphoric agreement was presented in Diercks, van Koppen and Putnam (2020).

pronouns in different constructions can be reconciled by the Saliency Structure Hypothesis. The core idea of this hypothesis is based on the notion of saliency structure and saliency values presented in the previous chapters. Strong pronouns generally prefer an antecedent with a low saliency value, while *pro* generally prefers the antecedent with the highest saliency value. As far as the proposal made in the previous chapter on the encoding of a discourse-related [uR] feature on pronouns, the Saliency Structure Hypothesis correctly predicts that the feature will be assigned to overt strong pronouns and not to *pro*.

Subject clitics display a more ambiguous behaviour, especially in heritage Venetan varieties. In these varieties, subject clitics generally prefer an antecedent with a low saliency value, except for the case in which such antecedent is not an argument of the verb. In the same context, strong subject pronouns display a preference for the antecedent with a low saliency value, even though it is not an argument of the verb. I proposed that, unlike the contexts discussed in Chapter 4, in which subject clitic preferences were attributed to a difference in the syntactic structure, in the case of unaccusative constructions discussed in the present chapter, the difference can be attributed to a difference in the theta-structure of the verb or, alternatively, on an extra processing load in the interpretation of such antecedence relationship. Therefore, the interpretation of subject clitics depends on a fine-grained system of syntactic, discourse and semantics factors. Conversely, strong pronouns behave as logophors, in that their antecedent selection is mainly ruled by discourse-related factors, regardless of the syntactic configuration.

220 The syntax of subject pronouns in heritage languages

Chapter 6. Conclusion

6.1. Summary of the thesis

6.1.1. Chapter overview

The aim of this study is to investigate syntactic change in situations of language contact. The languages included in the study qualify as heritage languages and are spoken in Italy and by communities of descendants of Italian immigrants in Argentina, Belgium, Brazil and Canada. Rothman (2009) defines a heritage variety as a language that is not a dominant language of the larger national society, spoken mainly at home in a familiar context.

All heritage speakers are either sequential or simultaneous bilinguals (Polinsky 2018) and the dominant language of the larger national society may affect their heritage language in different ways and at different levels. Therefore, according to Rothman (2009), heritage speakers have some command of their native heritage language acquired naturalistically, but their competence differs from that of native monolinguals as a consequence of language contact.

The idea that heritage competence in the native language differs from monolingual competence often led linguists to conclude that heritage speakers are unbalanced bilinguals, since they use the heritage language only in very restricted situations and circumstances. Their weaker language is in fact their native language, while the stronger language is the dominant language of the society (Polinsky 2018).

This study focusses on discourse-related features involved in the distribution of different types of subject pronouns. I investigate how different types of subject pronouns interact with syntax and with information structure and what happens to such features when languages get in contact with others. The present approach builds on Miyagawa (2005), who proposes that agreement with discourse-related features is

222 The syntax of subject pronouns in heritage languages

associated with a discourse feature. Following Aboh (2010), it is proposed that discourse features are assigned to lexical items already in the numeration. The discourse features encoded in these lexical items display a very specific syntactic behaviour and are therefore comparable to other types of formal features. This study focusses on the role of discourse features at the DP-level and their effect on the interpretation of subject pronouns.

In Chapter 1 I introduce the discussion on language contact and heritage varieties. After presenting the case of heritage Italo-Romance languages, the role of information structure and discourse features in shaping the pronominal system of null-subject languages is discussed.

Chapter 2 presents an approach to subject pronouns in null-subject languages. The main hypothesis is that all subject pronouns have the same internal structure and differences in their interpretation depend on a discourse feature realised in the D-head. This feature is defined as [R] (referential): when subject pronouns encode this feature, they are overt and referentially specific enough to obviate or switch reference; when subject pronouns lack this feature, they refer to the most salient discourse antecedent and are normally not phonologically realised. The hypothesis is tested on heritage southern Italo-Romance varieties and, partially, on Venetan. I start by discussing the internal structure of subject pronouns, referring specifically to strong subject pronouns and to null weak subjects. I propose that both 'strong' and 'weak' pronouns have the structure of a DP; the two types of pronouns are distinguished by an [uR] feature encoded by D in strong pronouns and lacked in weak pronouns. The [uR] feature makes the pronoun referentially specific enough to switch the reference to a non-salient discourse antecedent. For weak pronouns, the lack of [uR] implies that the pronoun will corefer with the most salient discourse antecedent. In consistent null subject languages, weak subject pronouns can undergo PF deletion under ϕ -feature identity with finite T, resulting in a phonologically null *pro* (Holmberg 2005, Sheehan 2006, Roberts 2009). The difference between strong pronouns and weak (null) pronouns is represented in (1):

- (1)
- a. [DP_[uR] she [ϕ P [NP]]]
 - b. [DP *pro* [ϕ P [NP]]]

This proposal captures the distribution of null and overt subjects in heritage Italo-Romance varieties. These languages are well-behaved null subject languages, as shown by the fact that the distribution of null and overt subjects depends on a precise interaction of agreement and discourse factors. Particularly in the case of first person (the most salient in discourse), ϕ -features can identify a null subject. This is true also for third person, but the reference is sometimes less accessible; in those cases, an overt pronoun needs to be realised. Also in heritage varieties, the difference between the overt and the null element is the presence of [uR] in the former.

This approach combines traditional views on the interaction of discourse and syntax in the realisation of null subjects (Frascarelli 2007; Sorace et al. 2009; Sigurdsson 2011, 2014) to the analysis of null subjects presented in Holmberg (2005), Sheehan (2006) and Roberts (2009).

Chapter 3 extends the approach adopted in this study to Venetan, a northern Italo-Romance variety that allows for overt pronouns even when they lack [uR]. This is the case of subject clitics. Most works on subject clitics claim that they are agreement markers rather than real pronouns; however, analysing subject clitics from the perspective of information structure, it is possible to notice that their behaviour is clearly pronominal. The hypothesis is that Venetan subject clitics can be analysed as pronouns that may or may not be referentially specific enough to obviate or switch reference. Specifically, when they encode [uR], they behave as regular overt pronouns; when they lack [uR], they are phonologically realised counterparts of null *pro*. It is shown that subject clitics in Italian Venetan generally encode only ϕ -features, while subject clitics in heritage Venetan encode [uR] more frequently, allowing for a pronominal behaviour. Consistently with the approach on null and overt subjects proposed in Chapter 2, [uR] is associated with the D-layer that is traditionally taken to be lacked by subject clitics (as predicted, for instance, by Cardinaletti and Starke 1999). Therefore, the behaviour displayed by subject clitics in heritage Venetan

224 The syntax of subject pronouns in heritage languages

cannot be explained if their analysis as markers of φ -agreement (as proposed specifically for subject clitics in Rizzi 1986; Brandi and Cordin 1989; Benincà, 1994; Poletto 1993, 2000) is maintained. It is proposed that subject clitics, just as strong pronouns and null *pro*, are DPs and that the different behaviour displayed by subject clitics in heritage and Italian varieties of Venetan depends on the presence of [uR]. The proposal is supported by the fact that an analysis of subject clitics as φ -heads cannot capture several distributional facts in Italian varieties of Venetan too. The agreement-like behaviour of subject clitics in these varieties depends on an adjacency requirement that triggers cliticisation of the pronoun at PF, just as in the case of French subject clitics, as proposed by Kayne (1983). This adjacency requirement represents an alternative to the PF deletion under feature identity that was described in Chapter 2: in both cases, the φ -composition of the pronoun is a subset of the feature composition of T. In Italian Venetan, the pronoun generally remains overtly realised if it is adjacent to the verb. In heritage Venetan, when subject clitics are not assigned [uR], they are generally deleted at PF under feature-identity, just as in other consistent null subject languages (Sheehan 2006, Roberts 2009).

In Chapter 4, the Position of the Antecedent Hypothesis ('PAH'; Carminati 2002) is tested on the three types of Venetan subject pronoun: strong, clitic and null.

(2) Position of the Antecedent Hypothesis (PAH; Carminati 2002)

The null pronoun prefers an antecedent which is in the Spec-IP position, while the overt pronoun prefers an antecedent which is not in the Spec-IP position.

The present study involves three Venetan varieties: Italian Venetan, Argentinian Venetan and Brazilian Venetan. The results of the study show that the antecedent preferences of subject clitics in heritage Venetan varieties match those of overt pronouns. Besides, the predictions made by the PAH are captured by the model proposed in Chapters 2 and 3, supporting the idea that at least some discourse-related features display clear syntactic effects. This study on Venetan pronouns represents a first application of the model put forward in Chapters 2 and 3 and allows to refine the proposal on the distribution of different pronominal forms in heritage varieties. It is

shown that, with respect to the predictions made by the PAH, subject clitics in Venetan are generally interpreted on par with strong pronouns and select an antecedent that is not in Spec-TP; in some cases they allow also for an interpretation akin to that of null subjects and select an antecedent in Spec-TP. Antecedent selection preferences of subject clitics depend on the presence of [uR]: subject clitics encoding [uR] prefer an antecedent that is not in the matrix Spec-TP, while subject clitics lacking [uR] prefer an antecedent in matrix Spec-TP. In Italian Venetan, subject clitics can be of both types. This difference is strictly dependent on the type of subordinate in which the subject clitic is realised; accessibility (defined in terms of c-command and salience structure of the sentence) is responsible for the constraint on the interpretation of subject clitics in complement *that*-clauses. Finally, I reflected more extensively on the fact that Venetan, as well as the other languages discussed in previous chapters, are always spoken in contact with other languages. Building on Lardiere (2008), it was proposed that the distribution of [uR] in the three varieties of Venetan depends on feature-reassembly: this proposal allows for the different distributions displayed by the varieties under analysis, capturing at the same time the role of innovation, rather than simplification, in heritage languages.

In Chapter 5 the role of salience in determining antecedent selection by different types of pronouns is discussed. I present a study on antecedent selection in Venetan, showing that the PAH is too strong and that the salience structure of a sentence is the main factor determining the correct interpretation of a pronoun at the discourse level. In particular, the PAH is challenged by the cases in which the discourse role of the antecedents of pronominal forms does not match their syntactic position. It is shown that subject pronouns are sensitive to syntactic and discourse factors alike: the interaction of these factors leads to precise conditions of antecedent selection. In contact situations, these conditions are not lost, but rearranged and may lead to a complexification of the system. Building on multiple-factor approaches to antecedent selection (Kaiser and Trueswell, 2008; Frana, 2007), it is proposed that preferences of all types of pronouns in different constructions can be reconciled by the Salience Structure Hypothesis.

(3) Saliency Structure Hypothesis

In case of referential ambiguity:

- a. the preferred antecedent for null subjects is the referent with the highest saliency value;
- b. the preferred antecedent for overt subjects is the referent with the lowest saliency value.

The core idea of this hypothesis is based on the notion of saliency structure and saliency values presented in the previous chapters. Strong pronouns generally prefer an antecedent with a low saliency value, while *pro* generally prefers the antecedent with the highest saliency value. As far as the proposal made in the previous chapter on the encoding of a discourse-related [uR] feature on pronouns, the Saliency Structure Hypothesis correctly predicts that the feature will be assigned to overt strong pronouns and not to *pro*. Subject clitics display a more ambiguous behaviour, especially in heritage Venetan varieties. In these varieties, subject clitics generally prefer an antecedent with a low saliency value, except for the case in which such antecedent is not an argument of the verb. In the same context, strong subject pronouns display a preference for the antecedent with a low saliency value, even though it is not an argument of the verb. I propose that, unlike the contexts discussed in Chapter 4, in which subject clitic preferences were attributed to a difference in the syntactic structure, in the case of unaccusative constructions discussed in the present chapter, the difference can be attributed to a difference in the theta-structure of the verb or, alternatively, on an extra processing load in the interpretation of such antecedence relationship. Therefore, the interpretation of subject clitics depends on a fine-grained system of syntactic, discourse and semantics factors. Conversely, strong pronouns behave as logophors, in that their antecedent selection is mainly ruled by discourse-related factors, regardless of the syntactic configuration.

6.2. Suggestions for future research

The analysis presented in this dissertation addresses several issues related to the syntax of subject pronouns. However, the data collected for this study evidenced a number of further directions that can be taken in order to further support the model proposed here. They are summarised in this section.

6.2.1. *Beyond subjects: the salience of object pronouns*

While the approach presented in this study partially challenges the widely accepted tripartite model of structural deficiency discussed in Cardinaletti and Starke (1999), that model has the merit of identifying common properties in subject and object pronouns belonging to the same class. In other words, they showed that structural deficiency can be identified in pronouns regardless of their role in the sentence.

The approach adopted here too can be extended to object pronouns⁴⁰. This fact implies that, just like subject pronouns, all object pronouns are all DPs. Differences in their behaviour depend on their internal featural composition and, more specifically, on the presence of discourse features. In the case of Argentinian Friulian, for instance, the strong pronoun *je* (4a) and the object clitic *la* (4b) are both pronominal, as also proposed in D'Alessandro (2022). Notice that a null object (4c) is not accepted in Argentinian Friulian.

(4) Argentinian Friulian

a. Tu cognosis je.

you know.2SG her

b. Tu la cognosis.

you her.CL know.2SG

⁴⁰ See Pescarini (2021) for an analysis of object pronominal clitics that dispenses with classes of pronouns and structural deficiency.

228 The syntax of subject pronouns in heritage languages

c. *Tu cognosis (je).

you know.2SG her

‘You know her’.

Null objects are rare in Romance languages; Argentinian Friulian is not an exception in this respect. One possible explanation for the unacceptability of a null object is the generally lower salience value of objects compared to subjects. It was shown that null pronouns lack a [uR] feature and refer to the most salient discourse referent; the ungrammaticality of (4)c) may therefore depend on the fact that all object pronouns in Friulian encode a [uR], resulting in the unavailability of null objects. However, the situation in Brazilian Venetan is more complex. Brazilian Venetan too allows for strong (5)a) and clitic (5)b) objects. Null objects (5)c) are not accepted; however, a sentence with a null object is marginally acceptable if the adverb *belche* is added (5)d).

(5) Brazilian Venetan

a. Te cognosi ela.

you know.2SG her

b. Te la cognosi.

you her.CL know.2SG

c. *Te cognosi (ela).

you know.2SG her

‘You know her’.

d. ?Te belche cognosi (ela).

you already know.2SG her

‘You know her.’

The difference between (5)c) and (5)d) and the role played by the adverb in making (5)d) more acceptable are a matter of future research.

6.2.2. *Subject resumption in relative clauses*

The distribution of resumptive pronominal pronouns in Brazilian Venetan presents some differences with respect from the varieties of Italian Venetan described in Benincà (1994).

The difference concerns the realisation of resumptive pronominal subjects in appositive and restrictive relative clauses. The two types of relative clauses have different semantic properties that emerge in the way they modify the head noun: appositive relative clauses provide some further specification about the head, while restrictive relative clauses restrict the reference of the head noun.

The two types of relative clauses are distinguished by intonation, but languages may employ syntactic strategies to distinguish them too. In Italian Venetan, for instance, appositive relative clauses feature a subject clitic agreeing with the relative head noun (6a), while restrictive relative clauses do not (6b). In Brazilian Venetan, however, both appositive (7a) and restrictive relative clauses (7b) are realised without a subject clitic.

(6) Italian Venetan

a. Le tose, che le ga diese ani, le va scuola.
 the girls which they have ten years they go school
 ‘The girls, which are ten years old, go to school.’

b. Le tose che ga diese ani, le va scuola.
 the girls that have ten years they go school
 ‘Girls that are ten years old go to school.’

(7) Brazilian Venetan

a. Le tose, che ga diese ani, va scuola.
 the girls which have ten years go school
 ‘The girls, which are ten years old, go to school.’

230 The syntax of subject pronouns in heritage languages

- b. Le tose che ga diese ani, va scuola.
the girls that have ten years go school
'Girls that are ten years old go to school.'

The difference may depend on different factors.

First, as it was already discussed in Chapter 3, Brazilian Venetan drops subject clitics more often than Italian Venetan: the use of the resumptive pronoun could then be associated with the general conditions on *pro*-drop in the two varieties.

Another possible explanation is a structural difference at the level of attachment of different types of relative clauses in the two varieties. Several studies (Jackendoff 1977, De Vries 2006, Bianchi 1999, Alexiadou *et al.* 2000) showed that restrictive relative clauses are complements of N, while appositive relative clauses are complements of D. Brazilian Venetan may have undergone simplification in this respect and that only one attachment site is available for the two types of relative clause: the complement of N. This, in turns, means that the two constructions display a very similar syntactic behaviour, being differentiated only at the intonational level.

6.2.3. *The problem of intra-speaker variation*

Most speakers Brazilian Venetan speakers use two different forms for third person of the verb *to be*.

(8) Brazilian Venetan

- a. Ze vegnesto me pare.
is come my father
- b. L'è vegnesto me pare.
is come my father
'My father came.'

Speakers use the two forms are used interchangeably. Notice that in Italian Venetan, the form *ze* is grammaticalized as third person form of the verb *to be* from the 18th century (Benincà 2007), while the nature of *l'è* is more ambiguous. According to Poletto (1993), *l'* functions as an 'auxiliary clitic', a specific form realized exclusively with 3rd person singular present form of the verb 'to be'. Crucially, *l'* is not a subject clitic, as shown by the fact that it is only realised with the auxiliary form (9)a); when there is no auxiliary form, the clitic does not appear (9)b).

(9) Brazilian Venetan

a. Me pare l'è rivà casa.
 my father is arrived home
 'My father arrived at home.'

b. Me pare riva casa.
 my father arrives home
 'My father arrives at home.'

The fact that in the same syntactic context the clitic appears only with the auxiliary form, shows that its realization is independent of the subject, but depends on the auxiliary itself (Garzonio and Poletto 2011). Besides, the use of *l'è* is geographically more restricted than *ze*; it is limited to certain northern and eastern varieties of Venetan that, in turn, do not display *ze*. In parallel, central Venetan varieties display *ze* but not *l'è*.

In Brazilian Venetan conversely, both forms are possible; *l'* does not play any syntactic role, as it is in free distribution with *ze*. In other words, there is intra-speaker variation, in that the two forms represent semantic/functional equivalents of the other form. Following Tortora (2014), it is possible that each variant is the reflex of a different grammar: speakers allowing both variants are bi-dialectal. This possibility is supported by the fact that first generation immigrants to Brazil used to speak slightly different local varieties of Venetan; it is therefore plausible that elements from different Venetan varieties entered Brazilian Venetan grammar,

232 The syntax of subject pronouns in heritage languages

remaining available as free alternatives. This fact is particularly relevant in the study of language contact, as it provides evidence against the levelling hypothesis, the reduction of difference between varieties of the same language in contact situation (Kerswill 2003). Levelling implies that different varieties of the same language converge into a new variety in which forms that are locally or dialectally marked disappear. This is not the case of Brazilian Venetan, in which elements belonging to different varieties of the language remain available to all speakers.

The current data did not evidence any difference in the interpretation and the use of the two copular forms in Brazilian Venetan. I leave the investigation of possible differences in the interpretation (following what was shown in Chapter 3 for Štivorian, for instance) for future research.

6.3. General conclusion

This study addressed several theoretical questions.

The general question addressed throughout the dissertation regards the **role of language contact** in the distribution of overt and null subjects in heritage languages. I proposed that heritage languages undergo feature-reassembly, which allows for different distribution and syntactic properties of subject pronouns. This proposal captured the role of innovation in heritage languages. Innovation may lead to complexification of the system.

The first one is concerned with the **internal structure of pronouns**. I proposed that all pronominal forms have the same syntactic structure, consisting of three layers: D, ϕ and N. It was shown that strong pronouns (in the terms of Cardinaletti and Starke 1999: pronouns that can have their own referential range, hence are not dependent on a salient discourse antecedent) encode [uR] in D. In the case of null subjects (weak pronominal elements, according to Cardinaletti and Starke, 1999), the D layer is still realised, but it lacks the [uR] feature.

The second question regards the **nature of subject clitics** in Venetan, a northern Italo-Romance variety. It was shown that subject clitics are not simple agreement markers, as proposed in previous studies. They are instead regular subject

pronouns. Building on previous studies on discourse-related properties of subject clitics, it was shown that their analysis as markers of ϕ -agreement leaves too many open issues on their distribution and realisation also in Italian varieties of the languages. Therefore, subject clitics are better analysed as pronouns; while they sometimes must be always adjacent to the verb (as per Baker's condition on morphological identification), in heritage Venetan they display a more autonomous behaviour in that they can be separated from the verb. This difference depends on the realisation of a [uR] feature in heritage Venetan subject clitics.

The third question regarded the **position of the antecedent** of different types of subject pronoun. It was shown that the interpretation of subject clitics depends on the presence of [uR]. Specifically in the case of subject clitics, this feature is optionally realised to switch the reference to a non-salient discourse antecedent, a possibility that was not investigated in previous studies on subject clitics; switch reference is traditionally thought to be a property of strong subject pronouns, in the terms of Cardinaletti and Starke (1999).

Finally, the question of how **extra-syntactic factors** are involved in the distribution of different types of pronouns is addressed. It is proposed the Saliency Structure Hypothesis, according to which referential ambiguity is solved at the pragmatic level by means of the saliency value of potential antecedents, established via Context Scanning. This hypothesis provides further support to the fact that the realisation of [uR] on pronouns is strictly connected to the distinction between salient and non-salient antecedents, as it is licensed on pronouns only when some update to the saliency structure is required.

234 The syntax of subject pronouns in heritage languages

References

- Aalberse, Suzanne and Pieter Muysken. 2013. Position paper: Perspectives on Heritage Languages. Radboud University, Faculty of Arts & Centre for Language Studies.
- Aboh, Enoch. 2010. Information structuring begins with the numeration. *Iberia: An International Journal of Theoretical Linguistics*, 2(1). 12-42.
- Adger, David and Gillian Ramchand. 2005. Merge and Move: Wh-Dependencies Revisited. *Linguistic Inquiry*, 36(2). 161-193.
- Aikhenvald, Alexandra. 2002. *Language contact in Amazonia*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra. 2006. Grammars in contact. A cross-linguistic perspective. In Alexandra Aikhenvald and Robert M. W. Dixon (eds), *Grammars in contact: A cross-linguistic typology*. Oxford: Oxford University Press. 2-66.
- Alexiadou, Artemis, Paul Law, André Meinunger, Chris Wilder. 2001. Introduction. In Artemis Alexiadou, Paul Law, André Meinunger, Chris Wilder (eds), *The syntax of relative clauses*. Amsterdam-Philadelphia: John Benjamins Publishing Company. 1-52.
- Andriani, L., R. D'Alessandro, A. Frasson, B. van Osch, S. Terenghi, L. Sorgini. 2022. *Adding the microdimension to the study of language change in contact. Three Italo-Romance case studies*. *Glossa: A Journal of General Linguistics*, 7(1).
- Ariel, Mira. 1990. *Accessing Noun Phrase Antecedents*. London: Routledge.
- Arnold, Jennifer. 2001. The Effect of Thematic Roles on Pronoun Use and Frequency of Reference Continuation. *Discourse Processes*, 31-2. 137-162.
- Baker, Mark. 1988. *Incorporation: A Theory of Grammatical Function Changing*. University of Chicago Press.

236 The syntax of subject pronouns in heritage languages

Baker, Mark. 1991. On some subject/object non-asymmetries in Mohawk. *Natural Language and Linguistic Theory*, 9. 537 – 576.

Belletti, Adriana and Luigi Rizzi. 1988. Psych verbs and theta-theory. *Natural Language and Linguistic Theory*, 6. 291-352.

Benincà, Paola. 1994. *La variazione sintattica*. Il Mulino, Bologna.

Benincà, Paola. 2007. Clitici e ausiliari: gh'ò, z'è. In Delia Bentley and Adam Ledgeway (eds), *Sui dialetti italo-romanzi. Saggi in onore di Nigel Vincent. Supplement to 'The Italianist' 1/27*. 27-47.

Benincà, Paola and Cecilia Poletto. 2004. Topic, focus and V2: defining the CP sublayers. In Luigi Rizzi (ed), *The Structure of IP and CP. The Cartography of Syntactic Structures. Vol. 2*. Oxford: Oxford University Press. 52-75.

Bentley, Delia, Francesco Ciconte, Silvio Cruschina. 2015. *Existentials and Locatives in Romance Dialects of Italy*. Oxford: Oxford University Press.

Bentley, Delia and Silvio Cruschina. 2018. Non-Canonical Postverbal Subjects. *Special issue of Italian Journal of Linguistics*, 30-2. 3-10.

Berruto, Gaetano. 1987. *Sociolinguistica dell'italiano contemporaneo*. Roma: Carocci.

Bianchi, Valentina. 2002. Headed relative clauses in generative syntax – Part I. *Glott International*, 6-7. 197-204.

Bianchi, Valentina. 2003. On finiteness as logophoric anchoring. In *Temps et point de vue/Tense and Point of View*. Paris: Université Paris X - Nanterre. 213-246.

Bobaljik, Jonathan. 1994. What does adjacency do. *MIT Working Papers in Linguistics*, 22. 1-32.

Boswijk, Vincent and Matt Coler. 2020. What is Saliency? *Open Linguistics*, 6-1. 713-722.

Brandi, Luciana and Patrizia Cordin. 1981. Dialetti e Italiano: un confronto sul Parametro del Soggetto Nullo. *Rivista di Grammatica Generativa*, 6. 33–87.

Brandi, Luciana and Patrizia Cordin. 1981. Two Italian Dialects and the Null Subject Parameter. In Osvaldo A. Jaeggli and Kenneth J. Safir (eds), *The Null Subject Parameter*. Dordrecht: Springer. 111-142.

- Brunetti, Lisa. 2009. On the semantic and contextual factors that determine topic selection in Italian and Spanish. *The Linguistic Review*, 26-2/3. 261-289.
- Büring, Daniel. 1999. *Topic*. In Peter Bosch and Rob van der Sand (eds), *Focus – Linguistic Cognitive and Computational Perspectives*. Cambridge, Cambridge University Press. 142-165.
- Burzio, Luigi. 1986. *Italian Syntax: A Government-Binding Approach*. Dordrecht: Redel.
- Cardinaletti, Anna. 1997. Subjects and Clause Structure. In Liliane Haegeman (ed.), *The New Comparative Syntax*. London: Longman. 33-63.
- Cardinaletti, Anna and Michal Starke 1999. The typology of structural deficiency: On the three grammatical classes. In Hans van Riemsdijk (ed), *Clitics in the Languages of Europe*. Berlin: Mouton de Gruyter. 145–233.
- Carminati, Maria Nella. 2002. *The processing of Italian subject pronouns*. Doctoral Dissertation, University of Massachusetts.
- Carstens, Vicki. 2003. Rethinking complementizer agreement: agree with a case-checked goal. *Linguistic Inquiry* 34. 393-412.
- Carvalho, Ana and Michael Child. 2011. Subject pronoun expression in a variety of Spanish in contact with Portuguese. In Jim Michnowicz and Robin Dodsworth (eds), *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*. Somerville: Cascadilla Proceedings Project. 14-25.
- Casalicchio, Jan and Alberto Frasson. 2018. Cambiamenti nell'uso dei soggetti clitici veneti: il ruolo del contatto con l'italiano. In Annamaria Chilà e Alessandro De Angelis (eds), *Capitoli di morfosintassi delle varietà romanze d'Italia: teoria e dati empirici*. Palermo: Centro di studi filologici e linguistici siciliani.
- Casalicchio, Jan and Alberto Frasson. 2019. The syntax of spatial anchoring: logophoricity in a heritage Italo-Romance variety. Presented at Going Romance XXXIII, Leiden.
- Casalicchio, Jan and Vania Masutti. 2015. A Syntactic Analysis of the Subject Clitic "a" in the Friulian Variety of Campone. *Isogloss - Open Journal of Romance Linguistics. Special Issue*. 103-132.
- Chafe, Wallace. 1987. Cognitive Constraints on Information Flow. In Russell Tomlin (ed), *Coherence and Grounding in Discourse*. Amsterdam: John Benjamins. 21-51.

238 The syntax of subject pronouns in heritage languages

Chierchia, Gennaro. 1992. Anaphora and dynamic binding. *Linguistics and philosophy*, 15. 111-183.

Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. Doctoral Dissertation, University of Massachusetts.

Heim, Irene. 1990. E-type pronouns and donkey anaphora. *Linguistics and philosophy*, 13. 137-177.

Chomsky, Noam. 1981. *Lectures of government and binding*. Dordrecht: Foris.

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

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed), *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press. 1-52.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero, Maria Luisa Zubizarreta (eds), *Foundational issues in Linguistic Theory*. Cambridge: The MIT Press. 133-166.

Cognola Federica and Jan Casalicchio. 2018. *Null Subjects in Generative Grammar: A Synchronic and Diachronic Perspective*. Oxford: Oxford University Press.

Cole, Melvyn. 2009. Null subjects: a reanalysis of the data. *Linguistics*, 47-3. 559-587.

Cole, Melvyn. 2010. Thematic null subjects and accessibility. *Studia Linguistica*. 64-3. 271-320.

D'Alessandro, Roberta. 2004. Impersonal *si* constructions. How semantics determines Agreement. In Matthew Wolf and Keir Moulton (eds), *Proceeding of NELS 34*. GLSA. 61-73.

D'Alessandro, Roberta. 2007. *Impersonal si constructions. Agreement and Interpretation*. Berlin/New York: Mouton de Gruyter.

D'Alessandro, Roberta. 2015. *Null Subjects*. In Antonio Fábregas, Jaume Mateu and Michael Putnam (eds), *Contemporary Linguistic Parameters*. London: Bloomsbury Press. 201-226.

D'Alessandro, Roberta. 2022. Crossing domains: topic marking and doubling in Romance. In Boneh, Nora, Daniel Harbour, Ora Matushansky and Roy (eds), *Building on Babel's Rubble*. PUV.

- D'Alessandro, Roberta, David Natvig, and Michael Putnam. 2021. Addressing challenges in formal research on moribund heritage languages: A path forward. *Front. Psychol.* 12: 700126.
- Déchaine, Rose-Marie and Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry*, 33-3. 409-442.
- De Mauro, Tullio. 1970. *Storia linguistica dell'Italia unita*. Roma: Laterza.
- Diercks, Michael , Marjo van Koppen, Michael Putnam. 2020. Agree probes down: Anaphoric feature valuation and phase reference. In Peter W. Smith, Johannes Mursell, Katharina Hartmann (eds), *Agree to agree. Agreement in the Minimalist Programme*. Berlin: Language Science Press. 347 – 389.
- Duarte, Maria Eugênia. 1995. *A perda do princípio evite pronome no português brasileiro*. Doctoral Dissertation. University of Campinas.
- Erteschik-Shir, Nomi. 2006. The architecture of topic and focus. In Valéria Molnár and Susanne Winkler (eds), *The Architecture of Focus*. Berlin: Mouton de Gruyter. 33-58.
- Ferguson, Charles. (1959). Diglossia. *Word*, 15. 325-340.
- Flores, Cristina, and Esther Rinke. 2020. Factors constraining subject expression in European Portuguese spoken in Hamburg. A bi-generational corpus investigation. *Discours. Revue de linguistique, psycholinguistique et informatique*, 26. 3-31.
- Frana, Ilaria. 2007. The role of discourse prominence in the resolution of referential ambiguities. Evidence from co-reference in Italian. *University of Massachusetts Occasional Papers in Linguistics*, 33. 1-24.
- Francez, Itamar. 2007. *Existential propositions*. Doctoral Dissertation. Stanford University.
- Frascarelli, Mara. 2007. Subjects, topics and the interpretation of referential pro. *Natural Language & Linguistic Theory*, 25. 691-734.
- Frascarelli, Mara, and Roland Hinterhölzl 2007. Types of topics in German and Italian. In Kerstin Schwabe and Susanne Winkler (eds), *On Information Structure, Meaning and Form*. Amsterdam: John Benjamins. 87–116.

240 The syntax of subject pronouns in heritage languages

Frasson, Alberto. 2021. Clitics are not enough: on agreement and null subjects in Brazilian Venetan. *Glossa: a journal of general linguistics* 6-1: 86. 1-17.

Frasson, Alberto, Roberta D'Alessandro, Brechje van Osch. 2021. Subject clitics in microcontact. A case study from heritage Friulian in Argentina and Brazil. *Heritage Language Journal*, 18-1. 1-36.

Garzonio, Jacopo And Cecilia Poletto 2011. I clitici di ausiliare nelle varietà piemontesi. *Studi sui dialetti del Piemonte, Quaderni di lavoro ASIt*. 117-134.

Gasque de Souza, Karoline, Da Silveira Chaves Lurian and Leonor Simioni. 2018. Sujeitos nulos no português de Poblado Uruguay. *PAPIA*, 28(1). 7-24.

Givón, Thomas. 1983. *Topic continuity in discourse: a quantitative crosslanguage study*. Philadelphia: PA John Benjamins.

Grassi Corrado, Alberto Sobrero, Tullio Telmon. 2006. *Introduzione alla dialettologia italiana*. Roma-Bari: Laterza.

Groenendijk, Jeroen and Martin Stokhof. 1991a. Two theories of dynamic semantics. In Jan van Eijck (ed), *Logics in AI. JELIA 1990. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence)*. Berlin-Heidelberg: Springer. 55-64.

Groenendijk, Jeroen and Martin Stokhof. 1991b. Dynamic Predicate Logic. *Linguistics and Philosophy*, 14-1. 39-100.

Haegeman, Liliane. 2019. Register-Based Subject Omission in English and its Implication for the Syntax of Adjuncts. *Anglophonia*, 28.

Heim, Irene and Angelika Kratzer. 1998. *Semantics in generative linguistics*. Oxford: Blackwell Publishers.

Heine, Bernd and Tania Kuteva. 2003. On contact-induced grammaticalization. *Studies in Language*, 27-3. 529-572.

von Heusinger, Klaus. 2000. Anaphora, antecedents, and accessibility. *Theoretical Linguistics*, 26. 75-93.

Holmberg, Anders. 2005. Is There a Little Pro? Evidence from Finnish. *Linguistic Inquiry*, 36-4. 533-564.

- Holmberg, Anders. 2009. Null Subject Parameters. In In Theresa Biberauer, Anders Holmberg, Ian Roberts, Michelle Sheehan (eds), *Parametric variation: Null subjects in minimalist theory*. Cambridge: Cambridge University Press. 88–124.
- Huang, J. 1984. *On the distribution and reference of empty pronouns*. *Linguistic Inquiry*, 15-4: 531-574.
- Hulk, A. and N. Müller. 2000. Bilingual first language acquisition at the interface between syntax and pragmatics. In *Bilingualism: Language and Cognition* 3.3: 227–44.
- Jackendoff, Ray. 1977. *X syntax: A study of phrase structure*. Cambridge: MIT Press.
- Jiménez-Fernández, Ángel. 2010. Discourse-agreement features, phasal C and the edge: A minimalist approach. *Diacrítica—Ciências da Linguagem*, 24. 23-46.
- Jiménez-Fernández, Ángel. 2016. When discourse met null subjects. *Borealis: An International Journal of Hispanic Linguistics*, 5-2. 173-189.
- Jiménez-Fernández, Ángel and Vassilios Spyropoulos. 2013. Feature inheritance, vp phases and the information structure of small clauses. *Studia Linguistica*, 67-2. 185-224.
- Kaiser, Elsi and John Trueswell. 2008. Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes*, 23. 709-748.
- Kamp, Hans. 1981. A theory of truth and semantic representation. in Jeroen Groenendijk (ed), *Formal Methods in the Study of Language*. Amsterdam: Mathematisch Centrum.
- van Kampen, Jacqueline. 2012. Obviation properties of the d-pronoun in Dutch. *Linguistics in the Netherlands 2012*. 68–81.
- Kato, Mary. 1999. Strong and weak pronominals in the null subject parameter. *Probus*, 11. 1-37.
- Kato, Mary. 2000. The partial pro-drop nature and the restricted VS order in Brazilian Portuguese. In Mary Kato and Esmeralda Negrão (eds), *Brazilian Portuguese and the null subject parameter*. Frankfurt: Vervuert. 223-258.
- Kato, Mary and Maria Eugênia Duarte. 2014. Restrições na distribuição de sujeitos

242 The syntax of subject pronouns in heritage languages

nulos no Português Brasileiro. *Veredas*, 18-1. 1-22.

Kerswill, Paul. 2003. Dialect levelling and geographical diffusion in British English. In David Britain and Jenny Cheshire (eds), *Social dialectology. In honour of Peter Trudgill*. Amsterdam: Benjamins. 223-243.

Koppen, Marjo van. 2012. The distribution of phi-features in pronouns. *Natural Language & Linguistic Theory*, 30-1, 135-177.

Krifka, Manfred. 2008. Basic notions of Information Structure. *Acta Linguistica Hungarica*, 55:3-4. 243-276.

Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representation of discourse referents*. Cambridge: Cambridge University Press.

Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural language and linguistic theory*, 22. 811-877.

Lardiere, Donna. 2008. Feature-Assembly in Second Language Acquisition, in Juana Liceras, Helmut Zobl, Helen Goodluck (eds), *The role of formal features in second language acquisition*. New York: Lawrence Erlbaum Associates. 106–140.

Levin, Beth and Malka Rappaport Hovav. 2005. *Argument realization*. Cambridge, Cambridge University Press.

Longobardi, Giuseppe. 1994. Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form. *Linguistic Inquiry*, 25-4. 609-665.

López, Luis. 2009. *A derivational syntax for information structure*. Oxford: Oxford University Press.

Loporcaro, Michele. 2009. *Profilo linguistico dei dialetti italiani*. Bari: Laterza.

Maiden, Martin and Mair Parry. 1997. *The dialects of Italy*. London: Routledge.

Manzini, Maria Rita and Leonardo Maria Savoia 2005. *I dialetti italiani e romanci. Morfosintassi generativa*. Alessandria: Edizioni dell'Orso.

Marantz, Alec. 1988. Clitics, Morphological Merger, and the Mapping to Phonological Structure. In Michael Hammond and Michael Noonan (eds), *Theoretical Morphology*. San Diego: Academic Press. 253–270.

- Mayol, Laia. 2009. *Pronouns in Catalan: information, discourse And strategy*. Doctoral Dissertation. University of Pennsylvania.
- Miyagawa, Shigeru. 2005. *On the EPP*, in Martha McGinnis and Norvin Richards (eds), *Perspectives on Phases*, Cambridge: MIT Working Papers in Linguistics, 49. 201-236.
- Miyagawa, Shigeru. 2010. *Why Agree? Why Move? Unifying Agreement-Based and Discourse Configurational Languages*. Cambridge: MIT Press.
- Miyagawa, Shigeru. 2017. *Agreement beyond phi*. Cambridge: MIT Press.
- Montalbetti, Mario. 1984. *After Binding: On the Interpretation of Pronouns*. Doctoral Dissertation. Massachusetts Institute of Technology.
- Montrul, Silvina. 2004. Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and cognition*, 7-2. 125-142.
- Montrul, Silvina. 2008. *Incomplete Acquisition in Bilingualism. Re-examining the Age Factor*. John Benjamins, Amsterdam.
- Nagy, Naomi, Nina Aghdasi, Derek Denis, Alexandra Motut. 2011. Null Subjects in Heritage Languages: Contact Effects in a Cross-linguistic Context. *University of Pennsylvania Working Papers in Linguistics*, 17-2. 134-144.
- Negrão, Esmeralda and Ana Müller 1996. As mudanças no sistema pronominal do português brasileiro: substituição ou especialização de formas? *D.E.L.T.A.*, 12-1. 125-152.
- Patel-Grosz, Pritty and Patrick Grosz. 2017. Revisiting pronominal typology. *Linguistic Inquiry*, 48-2. 259–297.
- Pellegrini, Giovan Battista. 1977. *Carta dei dialetti d'Italia*. Pisa: Pacini.
- Perlmutter, David. 1971. *Deep and surface constraints in syntax*. New York: Holt, Rinehart and Winston.
- Perlmutter, David. 1978. Impersonal Passives and the Unaccusative Hypothesis. *Berkeley Linguistic Society IV*. 157-189.

244 The syntax of subject pronouns in heritage languages

Pescarini, Diego. 2020. An emergentist view on functional classes. In Ludovico Franco and Paolo Lorusso (eds) *Linguistic Variation: Structure and Interpretation*. Berlin-Boston: De Gruyter Mouton. 531-560.

Pescarini, Diego. 2021. *Object Clitics*. Oxford: Oxford University Press.

Platzack, Christer. 2001. The Vulnerable C-domain. *Brain and Language* 77-3. 364-377.

Pinto, Manuela. 1997. *Licensing and interpretation of inverted subjects in Italian*. Doctoral Dissertation. Utrechts Instituut voor Linguïstiek.

Poletto, Cecilia. 1993. *La sintassi del soggetto nei dialetti italiani settentrionali*. Padova: Unipress.

Poletto, Cecilia. 2000. *The higher functional field: evidence from Northern Italian Dialects*. Oxford: Oxford University Press.

Polinsky, Maria. 2018. *Heritage Languages and their Speakers*. Cambridge: Cambridge University Press.

Ramchand, Gillian. 2013. Argument Structure and Argument Structure Alternations. In Marcel den Dikken (ed), *Cambridge Handbook of Generative Syntax*. Cambridge: Cambridge University Press. 265-321.

Reinhart, Tanya. 1982. Pragmatics and linguistics: an analysis of sentence topics. *Philosophica*, 27. 53-94.

Reinhart, Tanya. 1983. Coreference and bound anaphora: a restatement of the anaphora questions. *Linguistics and Philosophy*, 6. 47-88.

Reinhart, Tanya. 1992. Wh-in-situ: An Apparent Paradox. in Paul Dekker and Martin Stokhof (eds), *Proceedings of the Eight Amsterdam Colloquium*.

Reinhart, Tanya. 2002. The theta-system – an overview. *Theoretical Linguistics*, 28. 229-290.

Renzi, Lorenzo and Laura Vanelli. 1983. I pronomi soggetto in alcune varietà romanze. In Paola Benincà (ed), *Scritti linguistici in onore di Giovan Battista Pellegrini*. Pisa: Pacini. 121-145.

- Reuland, Eric. 2006. Logophoricity. In Eric Reuland and Martin Everaert (eds), *The Blackwell Companion to Syntax, Volume 1*. 1-20.
- Rinke, Esther and Cristina Flores. 2018. Another look at the interpretation of overt and null pronominal subjects in bilingual language acquisition: Heritage Portuguese in contact with German and Spanish. *Glossa: a journal of general linguistics*, 3-1. 1-68.
- Ritter, Elizabeth and Martina Wiltschko. 2014. The composition of INFL. *Nat Lang Linguist Theory*, 32. 1331–1386.
- Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Foris, Dordrecht.
- Rizzi, Luigi. 1986. On the Status of Subject Clitics in Romance. In Osvaldo A. Jaeggli and Carmen Silva-Corvalan (eds), *Studies in Romance Linguistics*. Dordrecht: Foris. 391-419.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. in Liliane Haegeman (ed), *Elements of Grammar*. Dordrecht: Kluwer. 281–337.
- Rizzi, Luigi and Ur Shlonsky. 2007. Strategies of subject extraction. In Hans-Martin Gärtner and Uli Sauerland (eds), *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*. Berlin: Mouton de Gruyter. 115-160.
- Roberts, Ian. 2009. A deletion analysis of null subjects. In Theresa Biberauer, Anders Holmberg, Ian Roberts, Michelle Sheehan (eds) *Parametric variation: Null subjects in minimalist theory*. Cambridge: Cambridge University Press. 88–124.
- Roberts, Ian. 2010. *Agreement and Head Movement. Clitics, incorporation and defective goals*. Cambridge: MIT Press.
- Rothman Jason. 2009. Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism*, 13-2. 155-163.
- Samek-Lodovici, Vieri. 1996. *Constraints on subjects: an Optimality Theoretic analysis*. Doctoral Dissertation, Rutgers University.
- Sheehan, Michelle. 2006. *The EPP and null subjects in Romance*. Doctoral Dissertation, Newcastle University.
- Sheehan, Michelle. 2010. 'Free' inversion in Romance and the Null Subject Parameter. In Theresa Biberauer, Anders Holmberg, Ian Roberts, Michelle Sheehan

246 The syntax of subject pronouns in heritage languages

(eds), *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press. 231-262.

Sheehan, Michelle. 2016. Subjects, null subjects and expletives in Romance. In Susann Fischer and Christoph Gabriel (eds), *Manuals of Romance Linguistics (MRL): Grammatical Interfaces*. Berlin: Mouton de Gruyter. 329-362.

Shin, Naomi. 2014. Grammatical complexification in Spanish in New York: 3sg pronoun expression and verbal ambiguity. *Language Variation and Change*, 26-3. 303-330.

Sigurðsson, Halldor. 2011. Conditions on argument drop. *Linguistic Inquiry*, 42. 267-304.

Sigurðsson, Halldor. 2014. Context-linked grammar. *Language Sciences*, 46. 175188.

Slabakova, Roumyana. 2014. The bottleneck of second language acquisition. *Foreign Language Teaching and Research*, 46-4. 543-559.

Sorace, Antonella. 2011. Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism*, 1-1. 1-33.

Sorace, Antonella and Francesca Filiaci. 2006. Anaphora resolution in near-native speakers of Italian. *Second Language Research*, 22-3. 339-368.

Sorace, Antonella and Ludovica Serratrice. 2009. Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism*, 13-2. 195-210.

Sorace, Antonella, Ludovica Serratrice, Francesca Filiaci, Michela Baldo. 2009. Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua*, 119-3. 460-477.

Sundaresan, Sandhya. 2012. *Context and (co)reference in the syntax and its interfaces*. Doctoral Dissertation, University of Stuttgart.

Speas, Margaret. 2004. Evidentiality, logophoricity, and the syntactic representation of pragmatic features. *Lingua*, 114. 255-276.

Sportiche, Dominique. 1996. Clitic Constructions. In Johan Rooryck and Laurie Zaring (eds), *Phrase Structure and the Lexicon. Studies in Natural Language and Linguistic Theory*, vol 33. Dordrecht: Springer. 213-276.

- Stalnaker, Robert. 1974. Pragmatic presuppositions. In Milton K. Munitz and Peter K. Unger (eds), *Semantics and Philosophy*. New York: New York University Press. 197-214.
- Tamburelli, Marco. 2010. The vanishing languages of Italy: Diglossia, bilingualism, and shift. in Piotr Chruszczewski (ed), *Philologica Wratislaviensia: Acta et Studia. Proceedings of the Conference Languages in Contact 2010*.
- Tamburelli, Marco. 2014. Uncovering the 'hidden' multilingualism of Europe: an Italian case study. *Journal of Multilingual and Multicultural Development*. 35-3. 252-270.
- Taraldsen, Knut Tarald. 1980. On the NIC, vacuous application and the that-trace filter. Indiana University Linguistics Club.
- Telmon, Tullio. 1992. *Le minoranze linguistiche in Italia*. Alessandria: Edizioni dell'Orso.
- Tortora, Cristina. 1997. *The syntax and semantics of the weak locative*. Doctoral dissertation, University of Delaware.
- Tortora, Cristina. 2014. Addressing the Problem of Intra-speaker Variation for Parametric Theory. In Raffaella Zanuttini and Laurence Horn (eds), *Micro-Syntactic Variation in North American English*. Oxford: Oxford University Press. 294-323.
- Trudgill, Peter. 2011. *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.
- Tsimpli, Ianthi, Antonella Sorace, Caroline Heycock, Francesca Filiaci. 2004. First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism*, 8-3. 257-277.
- Vanelli, Laura. 1998. *I dialetti italiani settentrionali nel panorama romanzo. Studi di sintassi e morfologia*. Roma: Bulzoni.
- Vries, Mark de. 2006. The Syntax of Appositive Relativization: On Specifying Coordination, False Free Relatives, and Promotion. *Linguistic Inquiry*, 37-2. 229-270.
- Weiß, Helmut. 2015. When the subject follows the object. On a curiosity in the syntax of personal pronouns in some German dialects. *Journal of Comparative German Linguistics*, 18. 65-92.

248 The syntax of subject pronouns in heritage languages

Wurmbrand, Susanne. 2010. Parasitic morphology in Germanic: consequences for the theory of feature checking. Ms, University of Connecticut.

Zeijlstra, Hedde. 2012. There is only one way to agree. *The Linguistic Review*, 29-3. 491-539.

Appendix A - Informants

1. Informants for Chapters 2 and 3

Table 1. Varieties per country.

Country	Variety	Informants
Argentina	Calabrian	1
	Eastern Abruzzese	1
	Venetan	2
Belgium	Neapolitan	1
	Sicilian	1
Brazil	Calabrian	2
	Venetan	22
Canada	Venetan	1
Italy ⁴¹	Venetan	6
Total		37

Table 2. Sociolinguistic information

N.	Country	City	Variety	Age	Gender	Education
1	Argentina	Buenos Aires	Calabrian	51-70	M	Primary
2	Argentina	La Plata	E. Abruzzese	71-90	M	University
3	Argentina	Rosario	Venetan	51-70	M	Primary
4	Argentina	Rosario	Venetan	71-90	M	Primary
5	Belgium	Brussels	Sicilian	71-90	F	High
6	Belgium	La Louvière	Neapolitan	31-50	F	Middle

⁴¹ Data from Italy come from the Microcontact Atlas (<https://microcontact.hum.uu.nl/#contributions>). These data include short interviews with native speakers of Central Venetan.

250 The syntax of subject pronouns in heritage languages

7	Brazil	Porto Alegre	Calabrian	71-90	M	High
8	Brazil	Porto Alegre	Calabrian	51-70	M	University
9	Brazil	Vale Veneto	Venetan	71-90	F	Primary
10	Brazil	Vale Veneto	Venetan	+91	M	Primary
11	Brazil	Ivorá	Venetan	51-70	M	University
12	Brazil	Nova Palma	Venetan	71-90	M	Primary
13	Brazil	Montebelo	Venetan	51-70	F	Primary
14	Brazil	Montebelo	Venetan	51-70	M	High
15	Brazil	Montebelo	Venetan	31-50	F	University
16	Brazil	Bento G.	Venetan	51-70	M	High
17	Brazil	Bento G.	Venetan	51-70	M	Middle
18	Brazil	Bento G.	Venetan	51-70	M	High
19	Brazil	Bento G.	Venetan	71-90	F	Primary
20	Brazil	Flores d. C.	Venetan	51-70	M	University
21	Brazil	Flores d. C.	Venetan	51-70	M	High
22	Brazil	Caxias do Sul	Venetan	71-90	F	Primary
23	Brazil	Caxias do Sul	Venetan	71-90	F	Primary
24	Brazil	Caxias do Sul	Venetan	31-50	F	University
25	Brazil	Caxias do Sul	Venetan	51-70	M	University
26	Brazil	Caxias do Sul	Venetan	51-70	F	High
27	Brazil	Silveira M.	Venetan	71-90	M	University
28	Brazil	Silveira M.	Venetan	71-90	M	Primary
29	Brazil	Porto Alegre	Venetan	71-90	M	Primary
30	Brazil	Porto Alegre	Venetan	51-70	M	University
31	Brazil	Porto Alegre	Venetan	71-90	F	Primary
32	Canada	Quebec City	Venetan	51-70	F	University
33	Italy	Preganziol	Venetan	71-90	M	Primary
34	Italy	Breda di P.	Venetan	71-90	F	Primary
35	Italy	Treviso	Venetan	51-70	M	High
36	Italy	Conegliano	Venetan	71-90	F	Primary
37	Italy	Vittorio V.	Venetan	71-90	F	Primary

Figure 1. Informants by level of education

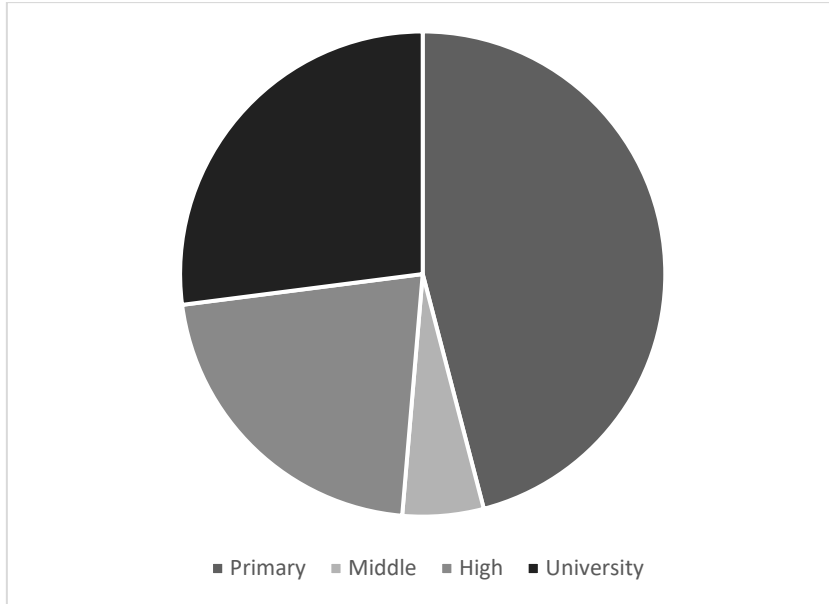


Figure 2. Informants by age group

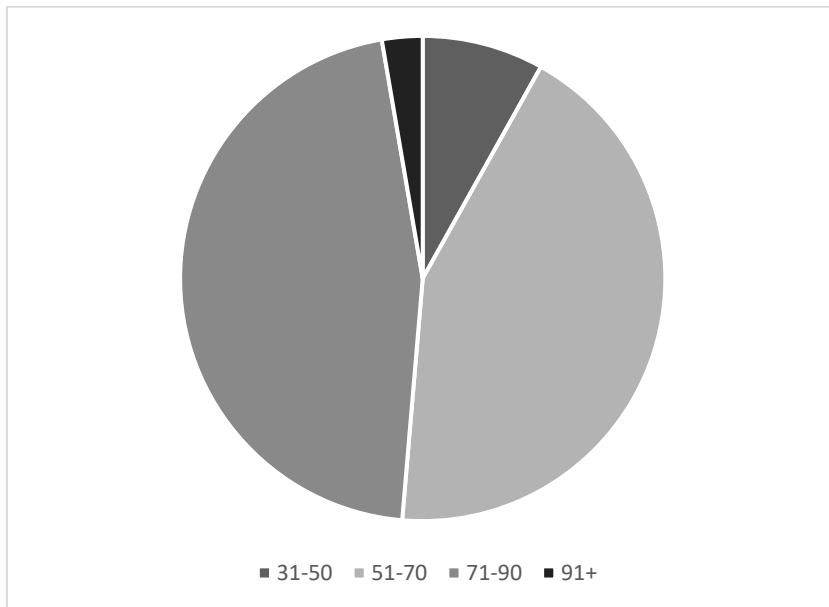
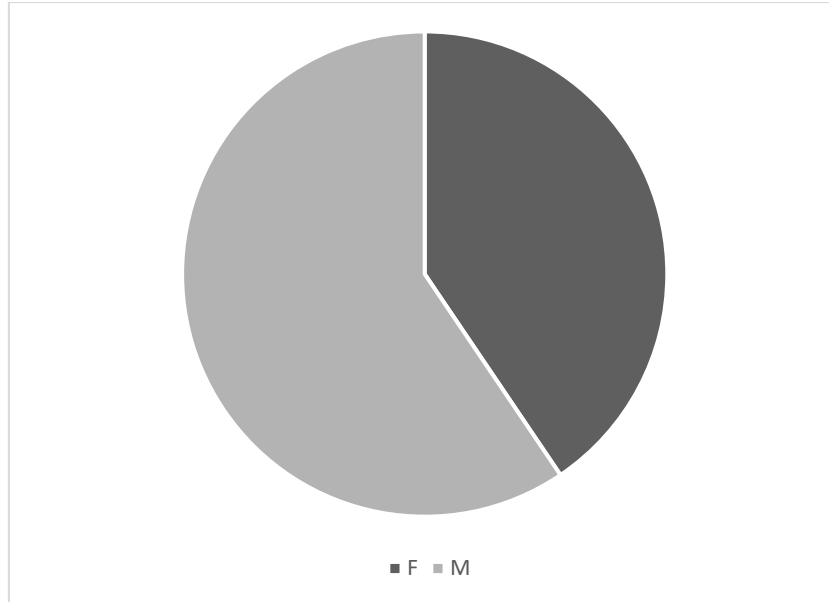


Figure 3. Informants by gender



2. Informants for Chapters 4 and 5

Table 4. Sociolinguistic information

N.	Country	City	Variety	Age	Gender	Language use
1	Argentina	Mendoza	Venetan	30-50	F	Sometimes
2	Argentina	Mendoza	Venetan	50-70	M	Sometimes
3	Argentina	Mendoza	Venetan	30-50	M	Sometimes
4	Brazil	Bento G.	Venetan	18-30	M	Sometimes
5	Brazil	Caxias do Sul	Venetan	30-50	M	Sometimes
6	Brazil	Flores d. C.	Venetan	30-50	F	Always
7	Brazil	Flores d. C.	Venetan	50-70	M	Always
8	Brazil	Caxias do Sul	Venetan	+70	M	Sometimes
9	Brazil	Flores d. C.	Venetan	50-70	F	Sometimes
10	Brazil	Caxias do Sul	Venetan	50-70	F	Sometimes
11	Brazil	Caxias do Sul	Venetan	50-70	F	Sometimes
12	Brazil	Ivorá	Venetan	30-50	M	Sometimes
13	Brazil	Bento G.	Venetan	18-30	M	Always

Appendix A 253

14	Brazil	Cambara d. S.	Venetan	50-70	M	Sometimes
15	Brazil	Carlos B.	Venetan	30-50	M	Sometimes
16	Brazil	Ivorá	Venetan	50-70	M	Sometimes
17	Brazil	Ivorá	Venetan	50-70	F	Sometimes
18	Brazil	Lajeado	Venetan	50-70	M	Always
19	Brazil	Ibiaçá	Venetan	30-50	M	Always
20	Brazil	Silveira M.	Venetan	30-50	M	Always
21	Brazil	Nova Prata	Venetan	50-70	F	Sometimes
22	Brazil	Ivorá	Venetan	+70	M	Sometimes
23	Brazil	Garibaldi	Venetan	+70	M	Always
24	Brazil	Nova Palma	Venetan	+70	M	Always
25	Brazil	Carlos B.	Venetan	50-70	F	Sometimes
26	Brazil	Ivorá	Venetan	30-50	M	Always
27	Brazil	Caxias do Sul	Venetan	50-70	F	Sometimes
28	Brazil	Bento G.	Venetan	50-70	F	Sometimes
29	Italy	Lugo	Venetan	30-50	M	Sometimes
30	Italy	Treviso	Venetan	30-50	F	Always
31	Italy	Treviso	Venetan	50-70	F	Always
32	Italy	Treviso	Venetan	50-70	M	Always
33	Italy	Rovigo	Venetan	18-30	F	Always
34	Italy	Preganziol	Venetan	18-30	F	Sometimes
35	Italy	Treviso	Venetan	50-70	F	Always
36	Italy	Spresiano	Venetan	50-70	F	Always
37	Italy	Preganziol	Venetan	50-70	F	Always
38	Italy	Preganziol	Venetan	50-70	M	Always
39	Italy	Treviso	Venetan	30-50	M	Always
40	Italy	Casale s. S.	Venetan	30-50	M	Sometimes
41	Italy	Casale s. S.	Venetan	50-70	M	Sometimes
42	Italy	Paese	Venetan	50-70	M	Sometimes
43	Italy	Preganziol	Venetan	+70	M	Sometimes
44	Italy	Treviso	Venetan	30-50	F	Always
45	Italy	Treviso	Venetan	50-70	F	Sometimes
46	Italy	Paese	Venetan	30-50	M	Sometimes

254 The syntax of subject pronouns in heritage languages

47	Italy	Preganziol	Venetan	50-70	F	Sometimes
48	Italy	Volpago	Venetan	30-50	F	Sometimes
49	Italy	Ponzano V.	Venetan	50-70	M	Always
50	Italy	Treviso	Venetan	18-30	M	Always
51	Italy	Treviso	Venetan	30-50	F	Sometimes
52	Italy	Treviso	Venetan	18-30	F	Always
53	Italy	Treviso	Venetan	50-70	F	Always
54	Italy	Treviso	Venetan	50-70	M	Sometimes
55	Italy	Treviso	Venetan	50-70	M	Always
56	Italy	Treviso	Venetan	50-70	F	Sometimes
57	Italy	Treviso	Venetan	50-70	F	Always
58	Italy	Treviso	Venetan	50-70	F	Always
59	Italy	Treviso	Venetan	50-70	F	Always
60	Italy	Treviso	Venetan	50-70	M	Always
61	Italy	Treviso	Venetan	30-50	F	Sometimes
62	Italy	Treviso	Venetan	50-70	M	Always
63	Italy	Villorba	Venetan	30-50	M	Always
64	Italy	Treviso	Venetan	50-70	F	Always
65	Italy	Treviso	Venetan	+70	M	Always

Figure 4. Informants by language usage per country: “When do you use Venetan?”

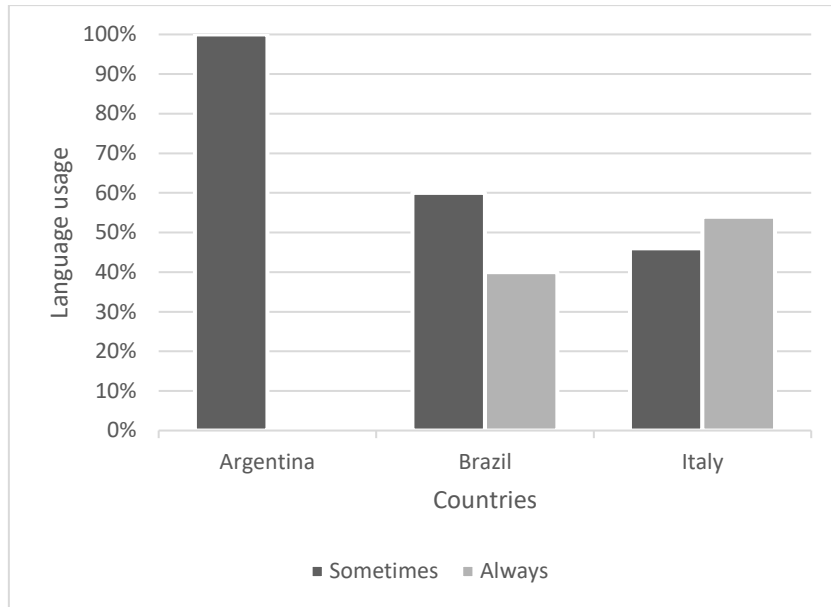


Figure 5. Informants by gender per country

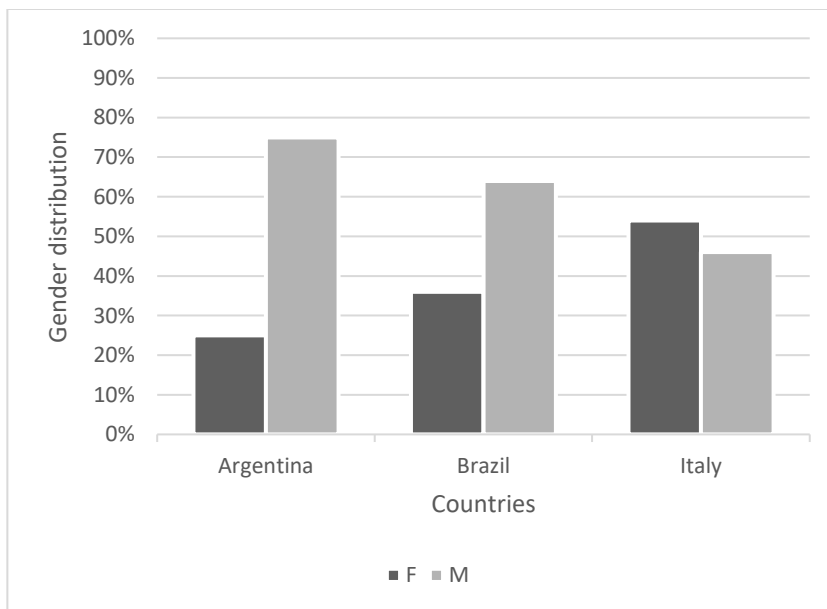
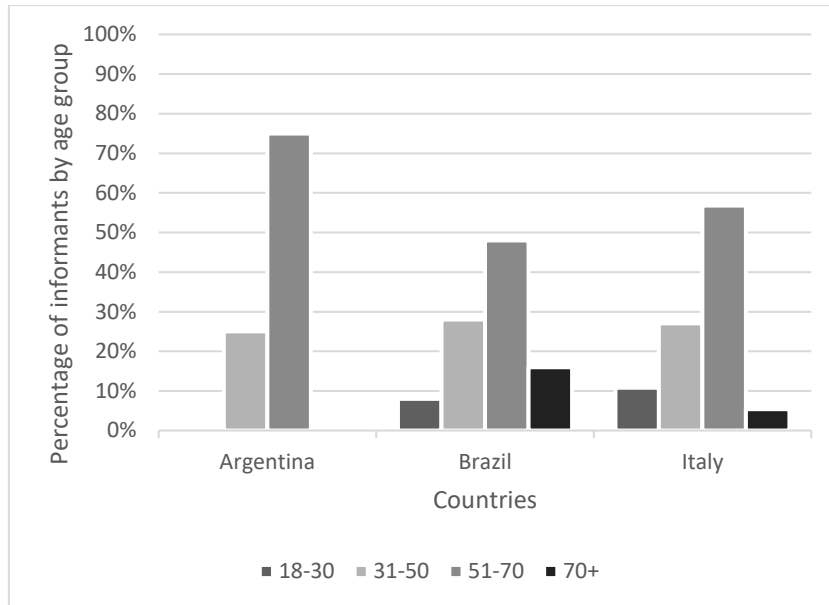


Figure 6. Informants by age group per country



Appendix B – Tasks and questionnaires

1. The exploratory questionnaire

The questionnaire was administered during the fieldworks that took place in 2019 in Argentina, Brazil and Canada. The questionnaire was rather exploratory, in that it had the goal of providing a better understanding of all the phenomena analysed in the context of the ERC Microcontact project: subject clitics ('SCI'), differential object marking ('DOM'), null subjects ('*pro*'), person-case constraint ('PCC'), auxiliary selection ('AUX') and demonstratives ('DEM'). As such, the questionnaire did not have the goal of providing conclusive results, but to identify issues related to the phenomena. The identified issues were then tackled specifically in the second study (Section 3). In this Appendix, only the items targeting subject clitics and *pro* will be reported.

Items and instructions were recorded by a native speaker of Venetan. Each informant was therefore provided with auditive stimuli. All the answers were recorded using Zoom H1n handy recorders. All recordings are stored on Utrecht University servers.

Forced choice task

Instruction: Sielga la fraze che ghe par pì iusta tra le do che ghe fao scoltar deso.

[Now you will listen to two sentences. Choose the one that sounds better to you]

1. [SCI] a. No'l me ga visto.
 not he.SCL me.OCL have.3SG seen
- b. El no me ga visto.
 he.SCL not me. OCL have.3SG seen

258 The syntax of subject pronouns in heritage languages

‘He did not see me.’

2. [SCI] a. El sempre dizea cusita.
he.SCL always said.3SG this
b. El dizea sempre cusita.
he.SCL said.3SG always this

‘He always said this.’

3. [*pro*] a. Mi te go visto ieri.
I you.OCL have.1SG seen yesterday
b. Te go visto ieri.
you.OCL have.1SG seen yesterday

‘I have seen you yesterday.’

4. [*pro*] a. Qua mi poso fumar, se voio.
here I can.1SG smoke.INF if want.1SG
b. Qua mi poso fumar, se voio.
here can.1SG smoke.INF if want.1SG

‘I can smoke here, if I want.’

5. [SCI] a. Marco, ieri, ga magnà masa.
Mark yesterday have.3SG eaten too much
b. Marco, ieri, el ga magnà masa.
Mark yesterday el.SCL have.3SG eaten too much

‘Mark, yesterday, ate too much.’

6. [SCI] a. La Maria ga cromptà el pan.
the Mary have.3SG bought the bread
b. La Maria la ga cromptà el pan.
the Mary she.SCL have.3SG bought the bread

‘Mary bought bread.’

7. [SCI] a. El magna e el beve.
he.SCL eat.3SG and he.SCL drink.3SG
b. El magna e beve.
he.SCL eat.3SG and drink.3SG

‘He eats and drinks.’

8. [SCI] a. Go parlà co le tose che le ze rivade ancuò.
 have.1SG spoken with the girls that they.SCL be.3PL arrived today

b. Go parlà co le tose che ze rivà ancuò.
 have.1SG spoken with the girls that be.3PL arrived today

‘I spoke to the girls that arrived today.’

9. [*pro*] a. Mi magno e mi bevo.

I eat.1SG and I drink.1SG

b. Magno e bevo.

eat.1SG and drink.1SG

‘I eat and drink.’

10. [SCI] a. Un biso, ieri, lo ga copà.

a snake yesterday it.OCL have.3SG killed

b. Un biso, ieri, el lo ga copà.

a snake yesterday he.SCL it.OCL have.3SG killed

‘Yesterday he killed a snake.’

11. [SCI] a. El ze zà ndà via?

he.SCL be.3SG already gone away

b. Zelo zà ndà via?

be.3SG.he.SCL already gone away

‘Has he already left?’

12. [*pro*] a. Te pensi che ti te starà promosso.
 you.SCL think.2SG that you you.SCL be.FUT.2SG promoted

b. Te pensi che te starà promosso ti.
 you.SCL think.2SG that you.SCL be.FUT.2SG promoted you

‘You think you will be promoted.’

13. [*pro*] a. In Italia se magna la pasta.

in Italy SE eat.3SG the pasta

b. In Italia magna la pasta.

in Italy eat.3SG the pasta

‘In Italy they eat pasta.’

260 The syntax of subject pronouns in heritage languages

14. [SCI] a. Varda! Ze drio piovar tanto!

look.IMP be.3SG behind rain.INF much

b. Varda! A ze drio piovar tanto!

look.IMP it.CL be.3SG behind rain.INF much

‘Look! It is raining a lot!’

15. [*pro*] a. Stanote mi go sognà che mi go copà un biso.

last.night I have.1SG dreamed that I have.1SG killed a snake

b. Stanote go sognà che go copà un biso.

last.night have.1SG dreamed that have.1SG killed a snake

‘Last night I dreamed that I killed a snake.’

16. [*pro*] a. Cuà mi poso fumar se mi voio.

here I can.1SG smoke.INF if I want.1SG

b. Cuà poso fumar se voio.

here can.1SG smoke.INF if want.1SG

‘I can smoke here, if I want.’

2. Analysis of spontaneous speech

Aside from the questionnaire presented in Section 1, informants that were interviewed during the 2019 fieldwork were asked to perform a spontaneous speech task. They were asked to tell a short story or episode from their childhood. The interviews were recorded using Zoom H1n handy recorders. All recordings are stored on Utrecht University servers.

The short stories told by the informants were transcribed to create corpora of spoken heritage Italo-Romance languages. All the transcriptions are stored on Utrecht University servers.

In Chapters 2 and 3 of this dissertation I used data from the heritage Venetan corpus (18397 words) and the southern Italo-Romance corpus (1853 words). All complete sentences (1308 for Venetan; 192 for southern Italo-Romance varieties) were subsequently transcribed in an Excel sheet, following the methodology already used in Frasson *et al.* (2021). All sentences were coded for the following variables:

PERSON (1, 2, 3, 4, 5, 6), GENRE (masculine, feminine), TYPE OF CONSTRUCTION (referential, generic, existential, impersonal), TYPE OF VERB (transitive, unergative, unaccusative, passive, copula), TYPE OF SENTENCE (matrix, embedded) TYPE OF PRONOUN (strong, clitic, null), POSITION (preverbal, postverbal), TOPIC (shift, continuation), MENTION (first, second, reintroduction). Some additional variables were added specifically for sentences including subject clitics, in order to identify issues with their behaviour: CLITIC REALISATION (yes, no), EXPECTED CLITIC (yes, no).

3. The questionnaire on subject pronouns

The questionnaire was administered online in spring 2020. It included 36 written stimuli, translated in Venetan by a native speaker. The questionnaire was shared on Google Forms with 65 speakers of Venetan in Argentina, Brazil and Italy. Each item consisted in one sentence with three possible answers; informants could choose one of the possible answers. The questionnaire was anonymous; the answers were saved as an Excel sheet and stored on Utrecht University servers.

Spanish instructions (for informants in Argentina): Muchas gracias por participar en este estudio! Este cuestionario se destina a hablantes de las variedades venetas. En el cuestionario, tienes que dar tu opinion sobre algunas frases en veneto. El cuestionario dura aproximadamente 10 minutos. Tus datos personales permanecerán confidenciales: el investigador nunca revelará tus datos personales.

Vas a leer algunas frases en veneto. Por cada frase, elige la respuesta que describe mejor la situación. Algunas frases parecerán raras en la variedad de veneto que hablas. En este caso, seleccione simplemente la opción que te parece menos rara.

Portuguese instructions (for informants in Brazil): Obrigado por concordar em participar deste estudo! O questionário é destinado a falantes do vêneto no Brasil. No questionário você tem que dar a sua opinião sobre algumas frases em vêneto. O questionário dura aproximadamente 10 minutos. Todas as suas informações serão

262 The syntax of subject pronouns in heritage languages

mantidas confidenciais. O pesquisador não irá compartilhar as suas respostas individuais com ninguém além da equipe da pesquisa.

Você vai ler algumas frases em vêneto. Para cada frase, seleccione a resposta que melhor descreve a situação. Algumas frases podem parecer estranhas ou erradas no vêneto que você fala. Neste caso, seleccione apenas a opção menos estranha.

Italian instructions (for informants in Italy): Grazie per aver accettato di partecipare a questo studio! Questo questionario è destinato ai parlanti del veneto. Nel questionario ti sarà chiesto di dare la tua opinione su alcune frasi in veneto. Il questionario dura circa 10 minuti. I tuoi dati personali rimarranno confidenziali: il ricercatore non divulgherà mai i tuoi dati personali.

Leggerai alcune frasi in veneto. Per ogni frase, seleziona la risposta che descrive meglio la situazione. Certe frasi potranno sembrarti strane nella tua varietà di veneto, per il modo in cui sono scritte o per le parole che vengono utilizzate. In questo caso, scegli semplicemente l'opzione che ti sembra meno strana.

[Thank you for accepting to take part in this study! This questionnaire is intended to speakers of Venetan. In the questionnaire you will be asked to give your opinion about some sentences in Venetan. The questionnaire lasts around 10 minutes. Your personal data will remain confidential: the researcher will not share your personal data. You will read some sentences in Venetan. For each sentence, select the answer that describes best the situation. Some sentences may sound strange in your variety of Venetan, because of the way they are written or for the used words. In this case, simply select the option that sounds less strange.]

1. Marco ghe scrivea sempre a Luca quando che el stava mal.
Mark him.IOCL wrote.3SG always to Luke when that he.SCL was.3SG sick
'Mark always wrote to Luke when he was sick.'
- a. Marco stava mal.
Mark was sick.
'Mark was sick.'

b. Luca stava mal.

Luke was sick

‘Luke was sick.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

2. Par via che ala Bruna no ghe piaze la Maria, serca de evitarla.

since that to.the Bruna not her.IOCL like the Mary try.3SG to avoid.INF.her.OCL

‘Since Bruna does not like Mary, she tries to avoid her.’

a. La Bruna serca de evitar la Maria.

the Bruna try.3SG to avoid.INF the Mary

‘Bruna tries to avoid Mary.’

b. La Maria serca de evitar la Bruna.

the Mary try.3SG to avoid.INF the Bruna

‘Mary tries to avoid Bruna.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

3. La Maria la chiama la Bruna quando che la riva tardi.

the Mary she.SCL call.3SG the Bruna when that she.SCL arrive.3SG late

‘Maria calls Bruna when she is late.’

a. La Maria riva tardi.

the Mary arrive.3SG late

‘Mary arrives late.’

b. La Bruna riva tardi.

the Bruna arrive.3SG late

‘Bruna arrives late.’

c. No se capisse.

not SE understand.3SG

264 The syntax of subject pronouns in heritage languages

‘I do not understand.’

4. Quando che riva me fradel a casa de Marco, se mete vardar la television.
when that arrive.3SG my brother to house of Mark SE put.3SG watch.INF the television

‘When my brother arrives at Mark’s place, he starts watching television.’

a. Me fradel se mete vardar la television.

my brother SE put.3SG watch.INF the television

‘My brother starts watching television.’

b. Marco se mete vardar la television.

Mark SE put.3SG watch.INF the television

‘Mark starts watching television.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

5. Par via che ala Bruna no ghe piaze la Maria, ela serca de evitarla.

since that to.the Bruna not her.IOCL like the Mary she try.3SG to avoid.INF.her.OCL

‘Since Bruna does not like Mary, she tries to avoid her.’

a. La Bruna serca de evitar la Maria.

the Bruna try.3SG to avoid.INF the Mary

‘Bruna tries to avoid Mary.’

b. La Maria serca de evitar la Bruna.

the Mary try.3SG to avoid.INF the Bruna

‘Mary tries to avoid Bruna.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

6. Quando che vien la Maria a casa de so mama, va subito in leto.

when that come.3SG the Mary to house of her mom go.3SG immediately in bed

‘When Mary comes to her mom’s place, she immediately goes to bed.’

- a. La Maria va subito in letto.
 the Mary go.3SG immediately in bed
 ‘Mary immediately goes to bed.’
- b. So mare va subito in letto.
 her mom go.3SG immediately in bed
 ‘Her mom immediately goes to bed.’
- c. No se capisse.
 not SE understand.3SG
 ‘I do not understand.’

7. Anca se a Marco no ghe interesa Luca, ghe ga telefonà ieri.
 even if to Mark not him.IOCL interest.3SG Luke him.IOCL have.3SG called yesterday
 ‘Even if Mark does not care about Luke, yesterday he called him.’

- a. Marco ga telefonà a Luca.
 Mark have.3SG called to Luke
 ‘Mark called Luke.’
- b. Luca ga telefonà a Marco.
 Luke have.3SG called to Mark
 ‘Luke called Mark.’
- c. No se capisse.
 not SE understand.3SG
 ‘I do not understand.’

8. La Maria ghe ga dito ala Bruna che la ga da comprar el pan.
 the Mary him.IOCL have.3SG said to.the Bruna that she.SCL have.3SG to buy.INF the bread
 ‘Mary told Bruna that she has to buy bread.’

- a. La Maria ga da comprar el pan.
 the Mary have.3SG to buy.INF the bread
 ‘Mary has to buy bread.’
- b. La Bruna ga da comprar el pan.
 the Bruna have.3SG to buy.INF the bread

266 The syntax of subject pronouns in heritage languages

‘Bruna has to buy bread.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

9. La Maria la parlava co la Bruna quando che stava in Italia.

the Mary she.SCL spoke.3SG with the Bruna when that was.3SG in Italy

‘Mary spoke to Bruna when she was in Italy.’

a. La Maria stava in Italia.

the Mary was.3SG in Italy

‘Mary was in Italy.’

b. La Bruna stava in Italia.

the Bruna was.3SG in Italy

‘Bruna was in Italy.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

10. Marco ghe ga dito a Luca che el ga magnà tanto.

Mark him.IOCL have.3SG said to Luke that he.SCL have.3SG eaten much

‘Mark told Luke that he ate a lot.’

a. Marco ga magnà tanto.

Mark have.3SG eaten much

‘Mark ate too a lot.’

b. Luca ga magnà tanto.

Luke have.3SG eaten much

‘Luke ate a lot.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

11. Marco ghe ga dito a Luca che gavea razon.

Mark him.IOCL have.3SG said to Luke that had.3SG reason

‘Mark told Luke that he was right.’

a. Marco gavea razon.

Mark had.3SG reason

‘Mark was right.’

b. Luca gavea razon.

Luke had.3SG reason

‘Luke was right.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

12. Marco ghe scrivea sempre a Luca quando che lù stava mal.

Mark him.IOCL wrote.3SG always to Luke when that he was.3SG sick

‘Mark always wrote to Luke when he was sick.’

a. Marco stava mal.

Mark was sick.

‘Mark was sick.’

b. Luca stava mal.

Luke was sick

‘Luke was sick.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

13. Quando che ze rivà Luca a casa de Marco, lù ga magnà la polenta.

when that be.3SG arrived Luke to house of Mark he have.3SG eaten the porridge

‘When Luke arrived at Mark’s place, he ate porridge.’

a. Luca ga magnà la polenta.

Luke have.3SG eaten the porridge

268 The syntax of subject pronouns in heritage languages

‘Luke ate porridge.’

b. Marco ga magnà la polenta.

Mark have.3SG eaten the porridge

‘Mark ate porridge.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

14. Par via che ala Maria no ghe interessa la Bruna, ghe risponde mal.
since that to.the Mary not her.IOCL interest the Bruna her.IOCL answer.3SG bad

‘Since Mary does not care about Bruna, she talks back at her.’

a. La Maria ghe risponde mal ala Bruna.

the Mary her.IOCL answer.3SG bad to.the Bruna

‘Mary talks back to Bruna.’

b. La Bruna ghe risponde mal ala Maria.

the Bruna her.IOCL answer.3SG bad to.the Mary

‘Bruna talks back to Mary.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

15. Marco ghe ga dito a Luca che lù ga magnà tanto.

Mark him.IOCL have.3SG said to Luca that he have.3SG eaten much

‘Mark told Luca that he ate a lot.’

a. Marco ga magnà tanto.

Mark have.3SG eaten much

‘Mark ate too a lot.’

b. Luca ga magnà tanto.

Luke have.3SG eaten much

‘Luke ate a lot.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

16. Quando che vien la Maria a casa de so mama, ela va subito in leto.
 when that come.3SG the Mary to house of her mom she go.3SG immediately in bed

‘When Mary comes to her mom’s place, she immediately goes to bed.’

a. La Maria va subito in leto.

the Mary go.3SG immediately in bed

‘Mary immediately goes to bed.’

b. So mare va subito in leto.

her mom go.3SG immediately in bed

‘Her mom immediately goes to bed.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

17. Par via che ala Bruna no ghe piaze la Maria, la serca de evitarla.

since that to.the Bruna not her.IOCL like the Mary she.SCL try.3SG to avoid.INF.her.OCL

‘Since Bruna does not like Mary, she tries to avoid her.’

a. La Bruna serca de evitar la Maria.

the Bruna try.3SG to avoid.INF the Mary

‘Bruna tries to avoid Mary.’

b. La Maria serca de evitar la Bruna.

the Mary try.3SG to avoid.INF the Bruna

‘Mary tries to avoid Bruna.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

18. Quando che ze rivà Luca a casa de Marco, ga magnà la polenta.

when that be.3SG arrived Luke to house of Mark have.3SG eaten the porridge

270 The syntax of subject pronouns in heritage languages

‘When Luke arrived at Mark’s place, he ate porridge.’

a. Luca ga magnà la polenta.
Luke have.3SG eaten the porridge

‘Luke ate porridge.’

b. Marco ga magnà la polenta.
Mark have.3SG eaten the porridge

‘Mark ate porridge.’

c. No se capisse.
not SE understand.3SG

‘I do not understand.’

19. Anca se a Marco no ghe interesa Luca, lù ghe ga telefonà ieri.
even if to Mark not him.IOCL interest.3SG Luke he him.IOCL have.3SG called yesterday

‘Even if Mark does not care about Luke, yesterday he called him.’

a. Marco ga telefonà a Luca.
Mark have.3SG called to Luke

‘Mark called Luke.’

b. Luca ga telefonà a Marco.
Luke have.3SG called to Mark

‘Luke called Mark.’

c. No se capisse.
not SE understand.3SG

‘I do not understand.’

20. Par via che ala Maria no ghe interesa la Bruna, la ghe risponde mal.
since that to.the Mary not her.IOCL interest the Bruna she.SCL her.IOCL answer.3SG bad

‘Since Mary does not care about Bruna, she talks back at her.’

a. La Maria ghe risponde mal ala Bruna.
the Mary her.IOCL answer.3SG bad to.the Bruna

‘Mary talks back to Bruna.’

b. La Bruna ghe risponde mal ala Maria.

the Bruna her.IOCL answer.3SG bad to.the Mary
'Bruna talks back to Mary.'

c. No se capisse.
not SE understand.3SG
'I do not understand.'

21. La Maria la chiama la Bruna quando che ela riva tardi.
the Mary she.SCL call.3SG the Bruna when that she arrive.3SG late
'Maria calls Bruna when she is late.'

a. La Maria riva tardi.
the Mary arrive.3SG late
'Mary arrives late.'

b. La Bruna riva tardi.
the Bruna arrive.3SG late
'Bruna arrives late.'

c. No se capisse.
not SE understand.3SG
'I do not understand.'

22. Marco ghe ga dito a Luca che ga magnà tanto.
Mark him.IOCL have.3SG said to Luke that have.3SG eaten much
'Mark told Luke that he ate a lot.'

a. Marco ga magnà tanto.
Mark have.3SG eaten much
'Mark ate too a lot.'

b. Luca ga magnà tanto.
Luke have.3SG eaten much
'Luke ate a lot.'

c. No se capisse.
not SE understand.3SG
'I do not understand.'

272 The syntax of subject pronouns in heritage languages

23. Quando che ze rivà Luca a casa de Marco, el ga magnà la polenta.
when that be.3SG arrived Luca to house of Mark he.SCL have.3SG eaten the porridge

‘When Luke arrived at Mark’s place, he ate porridge.’

a. Luca ga magnà la polenta.

Luke have.3SG eaten the porridge

‘Luke ate porridge.’

b. Marco ga magnà la polenta.

Mark have.3SG eaten the porridge

‘Mark ate porridge.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

24. La Maria ghe ga dito ala Bruna che ela ga da comprar el pan.
the Mary him.IOCL have.3SG said to.the Bruna that she have.3SG to buy.INF the bread

‘Mary told Bruna that she has to buy bread.’

a. La Maria ga da comprar el pan.

the Mary have.3SG to buy.INF the bread

‘Mary has to buy bread.’

b. La Bruna ga da comprar el pan.

the Bruna have.3SG to buy.INF the bread

‘Bruna has to buy bread.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

25. La Maria la parlava co la Bruna quando che ela stava in Italia.
the Mary she.SCL spoke.3SG with the Bruna when that she was.3SG in Italy

‘Mary spoke to Bruna when she was in Italy.’

a. La Maria stava in Italia.

the Mary was.3SG in Italy

‘Mary was in Italy.’

b. La Bruna stava in Italia.

the Bruna was.3SG in Italy

‘Bruna was in Italy.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

26. Marco ghe scrivea sempre a Luca quando che stava mal.

Mark him.IOCL wrote.3SG always to Luke when that was.3SG sick

‘Mark always wrote to Luke when he was sick.’

a. Marco stava mal.

Mark was sick.

‘Mark was sick.’

b. Luca stava mal.

Luke was sick

‘Luke was sick.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

27. Quando che vien la Maria a casa de so mama, la va subito in leto.

when that come.3SG the Mary to house of her mom she.SCL go.3SG immediately in bed

‘When Mary comes to her mom’s place, she immediately goes to bed.’

a. La Maria va subito in leto.

the Mary go.3SG immediately in bed

‘Mary immediately goes to bed.’

b. So mare va subito in leto.

274 The syntax of subject pronouns in heritage languages

her mom go.3SG immediately in bed
'Her mom immediately goes to bed.'

c. No se capisse.

not SE understand.3SG
'I do not understand.'

28. Anca se a Marco no ghe interesa Luca, el ghe ga telefonà ieri.
even if toMark not him.IOCL interest.3SG Luke he.SCL him.IOCL have.3SG called yesterday
'Even if Mark does not care about Luke, yesterday he called him.'

a. Marco ga telefonà a Luca.
Mark have.3SG called to Luke
'Mark called Luke.'

b. Luca ga telefonà a Marco.
Luke have.3SG called to Mark
'Luke called Mark.'

c. No se capisse.

not SE understand.3SG
'I do not understand.'

29. La Maria ghe ga dito ala Bruna che ga da comprar el pan.
the Mary him.IOCL have.3SG said to.the Bruna that have.3SG to buy.INF the bread
'Mary told Bruna that she has to buy bread.'

a. La Maria ga da comprar el pan.
the Mary have.3SG to buy.INF the bread
'Mary has to buy bread.'

b. La Bruna ga da comprar el pan.
the Bruna have.3SG to buy.INF the bread
'Bruna has to buy bread.'

c. No se capisse.

not SE understand.3SG
'I do not understand.'

30. La Maria la parlava co la Bruna quando che la stava in Italia.
 the Mary she.SCL spoke.3SG with the Bruna when that she.SCL was.3SG in Italy
 ‘Mary spoke to Bruna when she was in Italy.’

a. La Maria stava in Italia.
 the Mary was.3SG in Italy
 ‘Mary was in Italy.’

b. La Bruna stava in Italia.
 the Bruna was.3SG in Italy
 ‘Bruna was in Italy.’

c. No se capisse.
 not SE understand.3SG
 ‘I do not understand.’

31. Quando che riva me fradel a casa de Marco, lù se mete vardar la television.
 when that arrive.3SG my brother to house of Mark he SE put.3SG watch.INF the television
 ‘When my brother arrives at Mark’s place, he starts watching television.’

a. Me fradel se mete vardar la television.
 my brother SE put.3SG watch.INF the television
 ‘My brother starts watching television.’

b. Marco se mete vardar la television.
 Mark SE put.3SG watch.INF the television
 ‘Mark starts watching television.’

c. No se capisse.
 not SE understand.3SG
 ‘I do not understand.’

32. Marco ghe ga dito a Luca che el gavea razon.
 Mark him.IOCL have.3SG said to Luke that he.SCL had.3SG reason

276 The syntax of subject pronouns in heritage languages

‘Mark told Luke that he was right.’

a. Marco gavea razon.

Mark had.3SG reason

‘Mark was right.’

b. Luca gavea razon.

Luke had.3SG reason

‘Luke was right.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

33. La Maria la chiama la Bruna quando che riva tardi.

the Mary she.SCL call.3SG the Bruna when that arrive.3SG late

‘Maria calls Bruna when she is late.’

a. La Maria riva tardi.

the Mary arrive.3SG late

‘Mary arrives late.’

b. La Bruna riva tardi.

the Bruna arrive.3SG late

‘Bruna arrives late.’

c. No se capisse.

not SE understand.3SG

‘I do not understand.’

34. Par via che ala Maria no ghe interessa la Bruna, ela ghe risponde mal.

since that to.the Mary not her.IOCL interest the Bruna she her.IOCL answer.3SG bad

‘Since Mary does not care about Bruna, she talks back at her.’

a. La Maria ghe risponde mal ala Bruna.

the Mary her.IOCL answer.3SG bad to.the Bruna

‘Mary talks back to Bruna.’

b. La Bruna ghe risponde mal ala Maria.

the Bruna her.IOCL answer.3SG bad to.the Mary
 ‘Bruna talks back to Mary.’

- c. No se capisse.
 not SE understand.3SG
 ‘I do not understand.’

35. Quando che riva me fradel a casa de Marco, el se mete vardar la television.
 when that arrive.3SG my brother to house of Mark he.SCL SE put.3SG watch.INF the television
 ‘When my brother arrives at Mark’s place, he starts watching television.’

- a. Me fradel se mete vardar la television.
 my brother SE put.3SG watch.INF the television
 ‘My brother starts watching television.’

- b. Marco se mete vardar la television.
 Mark SE put.3SG watch.INF the television
 ‘Mark starts watching television.’

- c. No se capisse.
 not SE understand.3SG
 ‘I do not understand.’

36. Marco ghe ga dito a Luca che lù gavea razon.
 Mark him.IOCL have.3SG said to Luke that he had.3SG reason
 ‘Mark told Luke that he was right.’

- a. Marco gavea razon.
 Mark had.3SG reason
 ‘Mark was right.’

- b. Luca gavea razon.
 Luke had.3SG reason
 ‘Luke was right.’

- c. No se capisse.

Summary

This study aims to investigate syntactic change in situations of language contact. The languages included in the study are heritage Italo-Romance languages spoken in Argentina, Belgium, Brazil, Canada and Italy.

All heritage speakers are either sequential or simultaneous bilinguals and the dominant language of the larger national society may affect their heritage language at different levels. Heritage speakers acquired their native heritage language naturalistically, but their competence differs from that of native monolinguals as a consequence of language contact.

Since heritage speakers use their native language only in limited contexts, they are unbalanced bilinguals. Their weaker language is their native language, while the stronger language is the dominant language of the society.

This study focusses on discourse features involved in the distribution of different types of subject pronouns. I investigate how different subject pronouns interact with syntax and with information structure and what happens to discourse features when languages get in contact with others.

In Chapter 1 I introduce the discussion on language contact and heritage varieties.

I discuss the role of information structure and discourse features in shaping the pronominal system of null-subject languages.

Chapter 2 presents an approach to subject pronouns in null-subject languages. The main hypothesis is that all subject pronouns have the same internal structure; the differences in their interpretation depend on a discourse feature [R(eferential)] realised in the D-layer: when subject pronouns encode this feature, they are overt and referentially specific enough to obviate or switch reference; when subject pronouns lack this feature, they refer to the most salient discourse antecedent and are normally not phonologically realised. The hypothesis is tested on heritage

Italo-Romance varieties. I discuss the internal structure of subject pronouns and propose that both ‘strong’ and ‘weak’ pronouns have the structure of a DP; the two types of pronouns are distinguished by an [uR] feature encoded by D in strong pronouns and lacked in weak pronouns. The [uR] feature makes the pronoun referentially specific enough to switch the reference to a non-salient discourse antecedent. For weak pronouns, the lack of [uR] implies that the pronoun refers to the most salient discourse antecedent. In consistent null-subject languages, weak subject pronouns can undergo PF-deletion under ϕ -feature identity with finite T, resulting in a phonologically null pronoun.

This proposal captures the distribution of null and overt subjects in heritage Italo-Romance varieties. These languages are well-behaved null subject languages, as shown by the fact that the distribution of null and overt subjects depends on a precise interaction of agreement and discourse factors. This approach combines traditional views on the interaction of discourse and syntax in the realisation of null subjects to the analysis of null subjects presented in Holmberg (2005), Sheehan (2006) and Roberts (2009).

In Chapter 3, I extend the present approach to Venetan, a language that allows for overt pronouns even when they lack [uR]. This is the case of subject clitics. Most works on subject clitics claim that they are agreement markers rather than real pronouns; however, analysing subject clitics from the perspective of information structure, it is possible to notice that their behaviour is clearly pronominal. The hypothesis is that Venetan subject clitics can be analysed as pronouns that may or may not be referentially specific enough to obviate or switch reference. Specifically, when they encode [uR], they behave as regular overt pronouns; when they lack [uR], they are phonologically realised counterparts of null pronouns. I show that subject clitics in Italian Venetan generally encode only ϕ -features, while subject clitics in heritage Venetan encode [uR] more frequently, allowing for a pronominal behaviour. Consistently with the approach on null and overt subjects I proposed in Chapter 2, [uR] is associated with the D-layer that is traditionally taken to be lacked by subject clitics. Therefore, the behaviour displayed by subject clitics in heritage Venetan cannot be explained if their analysis as markers of ϕ -agreement is maintained. I

propose that subject clitics, just as strong and null pronouns, are DPs and that the different behaviour displayed by subject clitics in heritage and Italian varieties of Venetan depends on the presence of [uR]. An analysis of subject clitics as ϕ -heads cannot capture the distribution in Italian varieties of Venetan too. The agreement-like behaviour of subject clitics in these varieties depends on an adjacency requirement that triggers cliticisation of the pronoun at PF, just as in the case of French subject clitics. This adjacency requirement represents an alternative to the PF-deletion under feature identity: in both cases, the ϕ -composition of the pronoun is a subset of the feature composition of T. In Italian Venetan, the pronoun generally remains overt if it is adjacent to the verb. In heritage Venetan, when subject clitics are not assigned [uR], they are generally deleted at PF under feature-identity, just as in other consistent null subject languages.

In Chapter 4 I test the Position of the Antecedent Hypothesis on the three types of Venetan subject pronoun: strong, clitic and null.

(1) Position of the Antecedent Hypothesis (PAH)

The null pronoun prefers an antecedent which is in the Spec-IP position, while the overt pronoun prefers an antecedent which is not in the Spec-IP position.

The study involves three Venetan varieties: Italian Venetan, Argentinian Venetan and Brazilian Venetan. The results of the study show that the antecedent preferences of subject clitics in heritage Venetan varieties match those of overt pronouns. Besides, the predictions made by the PAH are captured by the model proposed in Chapters 2 and 3, supporting the idea that at least some discourse-related features display clear syntactic effects. I show that, with respect to the predictions made by the PAH, subject clitics in Venetan are generally interpreted on par with strong pronouns and select an antecedent that is not in Spec-TP; in some cases they allow also for an interpretation akin to that of null subjects and select an antecedent in Spec-TP. Antecedent selection preferences of subject clitics depend on the presence of [uR]: subject clitics encoding [uR] prefer an antecedent that is not in the matrix Spec-TP, while subject clitics lacking [uR] prefer an antecedent in matrix Spec-TP. In Italian Venetan, subject clitics

282 The syntax of subject pronouns in heritage languages

can be of both types. This difference depends on the type of subordinate in which the subject clitic is realised; accessibility (defined in terms of c-command and salience structure of the sentence) is responsible for the constraint on the interpretation of subject clitics in complement *that*-clauses. Finally, I reflect on the fact that heritage varieties are always spoken in contact with other languages. In this respect, I proposed that the distribution of [uR] in different Venetan varieties depends on feature-reassembly.

In Chapter 5 I discuss the role of salience in determining antecedent selection by different types of pronouns. I show that the PAH is too strong and that the salience structure of a sentence is the main factor determining the correct interpretation of a pronoun at the discourse level. In particular, the PAH is challenged by the cases in which the discourse role of the antecedents of pronominal forms does not match their syntactic position. I show that subject pronouns are sensitive to syntactic and discourse factors alike: the interaction of these factors leads to precise conditions of antecedent selection. In contact situations, these conditions are not lost, but rearranged and may lead to a complexification of the system. Building on multiple-factor approaches to antecedent selection, I propose that preferences of all types of pronouns in different constructions can be reconciled by the Salience Structure Hypothesis.

(2) Salience Structure Hypothesis

In case of referential ambiguity:

- a. the preferred antecedent for null subjects is the referent with the highest salience value;
- b. the preferred antecedent for overt subjects is the referent with the lowest salience value.

Strong pronouns generally prefer an antecedent with a low salience value, while *pro* generally prefers the antecedent with the highest salience value. As far as the proposal made in the previous chapter on the encoding of a discourse-related [uR] feature on pronouns, the Salience Structure Hypothesis correctly predicts that the feature will be assigned to overt strong pronouns and not to *pro*. Subject clitics display a more

ambiguous behaviour, especially in heritage Venetan varieties. In these varieties, subject clitics generally prefer an antecedent with a low salience value, except for the case in which such antecedent is not an argument of the verb. In the same context, strong subject pronouns display a preference for the antecedent with a low salience value, even though it is not an argument of the verb. I propose that, in the case of unaccusative constructions, the difference can be attributed to the theta-structure of the verb or to the processing load in the interpretation of such antecedence relationship. Therefore, the interpretation of subject clitics depends on a fine-grained system of syntactic, discourse and semantics factors. Conversely, strong pronouns behave as logophors, in that their antecedent selection is mainly ruled by discourse-related factors, regardless of the syntactic configuration.

Chapter 6 concludes the dissertation. I summarise the theoretical proposals made in the previous chapters on the structure and distribution of subject pronouns and on the role of contact; finally, this chapter outlines some suggestions for future research that emerged from the data collected for the present dissertation.

Samenvatting in het Nederlands

Deze studie heeft als doel syntactische verandering in situaties van taalcontact te onderzoeken. De talen die in het onderzoek zijn opgenomen, zijn Italo-Romaanse erfgoedtalen die worden gesproken in Argentinië, België, Brazilië, Canada en Italië.

Alle erfgoedsprekers zijn ofwel sequentieel ofwel simultaan tweetalig. Daarnaast, kan de dominante taal van de grotere nationale samenleving de erfgoedtaal op verschillende niveaus beïnvloeden. Erfgoedsprekers hebben hun moeder-erfgoedtaal naturalistisch verworven. Echter verschilt hun competentie met die van eentaligen als gevolg van taalcontact.

Omdat erfgoedtaalsprekers hun moedertaal slechts in beperkte contexten gebruiken, zijn ze onevenwichtige tweetaligen. Hun zwakkere taal is hun moedertaal, terwijl de sterkere taal de dominante taal is die gesproken wordt in de samenleving.

Deze studie richt zich op discourskenmerken die betrokken zijn bij de verspreiding van verschillende soorten persoonlijke voornaamwoorden. Ik onderzoek hoe verschillende persoonlijke voornaamwoorden omgaan met syntaxis en met informatiestructuur en wat er gebeurt met discourskenmerken wanneer talen in contact komen met andere talen.

In hoofdstuk 1 introduceer ik een discussie over taalcontact en verschillende erfgoedtalen. Ik bespreek de rol van informatiestructuur en discourskenmerken bij het vormgeven van het pronominale systeem van null-subject-talen.

Hoofdstuk 2 gaat over de benadering van persoonlijke voornaamwoorden in null-subject-talen. De hoofdhypothese betreft dat alle persoonlijke voornaamwoorden dezelfde interne structuur hebben; de verschillen in hun interpretatie hangen af van een discourskenmerk [R(eferential)] gerealiseerd in de D-laag: wanneer persoonlijke voornaamwoorden dit kenmerk coderen, zijn ze openlijk en specifiek genoeg om referentie te ondervangen of om te schakelen; wanneer persoonlijke voornaamwoorden dit kenmerk missen, verwijzen ze naar het meest opvallende

discours-antecedent en worden ze normaal gesproken niet fonologisch gerealiseerd. De hypothese wordt getest op Italo-Romaanse erfgoedtaal variëteiten. Ik bespreek de interne structuur van persoonlijke voornaamwoorden en stel voor dat zowel 'sterke' als 'zwakke' voornaamwoorden de structuur van een DP hebben; de twee soorten voornaamwoorden worden onderscheiden door een [uR]-kenmerk dat wordt gecodeerd door D in sterke voornaamwoorden en ontbreekt in zwakke voornaamwoorden. De [uR]-functie maakt het voornaamwoord referentieel specifiek genoeg om de verwijzing om te zetten naar een onopvallende discoursantecedent. Voor zwakke voornaamwoorden houdt het ontbreken van [uR] in dat verwijst het voornaamwoord naar het meest opvallende discoursantecedent. In consistente null-subject-talen kunnen zwakke persoonlijke voornaamwoorden PF-deletie ondergaan onder ϕ -kenmerk-identiteit met eindige T, wat resulteert in een fonologisch nulvoornaamwoord. Dit voorstel legt de verspreiding van null en openlijke onderwerpen in Italiaans-Romaanse erfgoedtaal variëteiten vast. Deze talen zijn fatsoenlijke null-subject-talen, zoals blijkt uit het feit dat de verdeling van nul- en openlijke onderwerpen afhangt van een precieze interactie van overeenstemmings- en discoursfactoren. Deze benadering combineert traditionele opvattingen over de interactie van discours en syntaxis bij de realisatie van nul onderwerpen met de analyse van nul onderwerpen gepresenteerd in Holmberg (2005), Sheehan (2006) and Roberts (2009).

In hoofdstuk 3 breid ik de benadering uit tot Venetaans, een taal die openlijke voornaamwoorden toestaat, zelfs wanneer deze [uR] missen. Dit is het geval bij onderwerp clitica. De meeste werken over onderwerp clitica beweren dat ze overeenkomst markers zijn in plaats van echte voornaamwoorden; bij het analyseren van subject clitica vanuit het perspectief van de informatiestructuur, is het echter mogelijk om op te merken dat hun gedrag duidelijk pronominaal is. De hypothese is dat Venetaanse onderwerp-clitica kunnen worden geanalyseerd als voornaamwoorden die al dan niet referentieel specifiek genoeg zijn om referentie te ondervangen of te veranderen. In het bijzonder, wanneer ze [uR] coderen, gedragen ze zich als gewone openlijke voornaamwoorden; als ze [uR] missen, zijn ze fonologisch gerealiseerde tegenhangers van nulvoornaamwoorden. Ik laat zien dat onderwerp-clitica in het

Italiaanse Venetaans over het algemeen alleen φ -kenmerken coderen, terwijl onderwerp-clitica in het erfgoed Venetaans vaker [uR] coderen, wat een pronominaal gedrag mogelijk maakt. In overeenstemming met de benadering van null en openlijke onderwerpen die ik in hoofdstuk 2 heb voorgesteld, wordt [uR] geassocieerd met de D-laag die traditioneel wordt beschouwd als ontbrekend door onderwerp-clitica. Daarom kan het gedrag van subject clitics in erfgoed Venetaans niet worden verklaard als hun analyse als markers van φ -overeenkomst wordt gehandhaafd. Ik stel voor dat onderwerp-clitica, net als sterke en null-voornaamwoorden, DP's zijn en dat het verschillende gedrag dat onderwerp-clitica vertonen in erfgoedtaal- en Italiaanse varianten van Venetaans afhangt van de aanwezigheid van [uR]. Een analyse van onderwerp-clitica als φ -hoofden kan de distributie in Italiaanse variëteiten van Venetaans ook niet vastleggen. Het overeenkomstachtige gedrag van onderwerp-clitica in deze varianten hangt af van een nabijheidsvereiste die de clitisering van het voornaamwoord bij PF veroorzaakt, net als in het geval van Franse onderwerp-clitica. Deze eis van nabijheid vertegenwoordigt een alternatief voor de PF-deletie onder kenmerkidentiteit: in beide gevallen is de φ -samenstelling van het voornaamwoord een subset van de kenmerksamenstelling van T. In het Italiaanse Venetaans blijft het voornaamwoord over het algemeen openlijk als het grenst aan het werkwoord. In erfgoedtaal Venetaans, wanneer onderwerp clitica niet zijn toegewezen [uR], worden ze over het algemeen verwijderd bij PF onder kenmerk-identiteit, net als in andere consistente nul onderwerptalen.

In hoofdstuk 4 test ik de positie van de antecedent-hypothese op de drie soorten Venetaanse persoonlijke voornaamwoorden: sterk, clitisch en null.

(1) Positie van de antecedent hypothese (PAH)

Het nul-voornaamwoord geeft de voorkeur aan een antecedent dat zich in de Spec-IP-positie bevindt, terwijl het openlijke voornaamwoord de voorkeur geeft aan een antecedent dat zich niet in de Spec-IP-positie bevindt.

De studie omvat drie Venetaanse variëteiten: Italiaanse Venetaans, Argentijnse Venetaans en Braziliaanse Venetaans. De resultaten van de studie laten zien dat de

288 The syntax of subject pronouns in heritage languages

eerdere voorkeuren van onderwerp-clitica in Venetaanse erfgoedvariëteiten overeenkomen met die van openlijke voornaamwoorden. Bovendien zijn de voorspellingen van de PAH vastgelegd in het model dat is voorgesteld in Hoofdstukken 2 en 3, wat het idee ondersteunt dat tenminste enkele discoursgerelateerde kenmerken duidelijke syntactische effecten vertonen. Ik laat zien dat, met betrekking tot de voorspellingen van de PAH, onderwerp-clitica in Venetaans over het algemeen worden geïnterpreteerd als sterke voornaamwoorden en selecteer een antecedent dat niet in Spec-TP staat; in sommige gevallen laten ze ook een interpretatie toe die lijkt op die van nul subjecten en selecteren ze een antecedent in Spec-TP. Voorafgaande selectievoorkeuren van onderwerp-clitica zijn afhankelijk van de aanwezigheid van [uR]: onderwerp-clitica die coderen voor [uR] geven de voorkeur aan een antecedent dat niet in de matrix Spec-TP staat, terwijl onderwerp-clitica zonder [uR] de voorkeur geven aan een antecedent in matrix Spec-TP. In het Italiaanse Venetaans kunnen onderwerp-clitica van beide typen zijn. Dit verschil hangt af van het type ondergeschikte waarin de subjectieve clitica wordt gerealiseerd; Toegankelijkheid (gedefinieerd in termen van c-commando en opvallende zinstructuur) is verantwoordelijk voor de beperking van de interpretatie van onderwerp-clitica in complementaire *that*-clausules. Tot slot sta ik stil bij het feit dat erfgoed taalvariëteiten altijd worden gesproken in contact met andere talen. In dit opzicht heb ik voorgesteld dat de distributie van [uR] in verschillende Venetaanse variëteiten afhangt van het opnieuw samenstellen van kenmerken.

In hoofdstuk 5 bespreek ik de rol van opmerikbaarheid bij het bepalen van de antecedent selectie m.b.t. verschillende soorten voornaamwoorden. Ik laat zien dat de PAH te sterk is en dat de opmerikbaarheid-structuur van een zin de belangrijkste factor is die de juiste interpretatie van een voornaamwoord op discoursniveau bepaalt. In het bijzonder wordt de PAH uitgedaagd door de gevallen waarin de discoursrol van de antecedenten van pronominale vormen niet overeenkomt met hun syntactische positie. Ik laat zien dat persoonlijke voornaamwoorden gevoelig zijn voor zowel syntactische als discoursfactoren: de interactie van deze factoren leidt tot precieze waarden voor voorafgaande selectie. In contactsituaties gaan deze waarden niet verloren, maar worden ze herschikt en kunnen ze leiden tot een complexisering van het systeem.

Voortbouwend op multi-factor benaderingen van de antecedentselectie, stel ik voor dat voorkeuren van alle soorten voornaamwoorden in verschillende constructies kunnen worden overeengestemd met de opmerikbaarheid-structuurhypothese.

(2) Opmerikbaarheid-structuurhypothese

In geval van referentiële ambiguïteit:

- a. het geprefereerde antecedent voor nul onderwerpen is de referent met de hoogste opmerikbaarheidswaarde;
- b. het geprefereerde antecedent voor openlijke onderwerpen is de referent met de laagste opmerikbaarheidswaarde.

Sterke voornaamwoorden geven over het algemeen de voorkeur aan een antecedent met een lage opmerikbaarheidswaarde, terwijl pro over het algemeen de voorkeur geeft aan het antecedent met de hoogste opmerikbaarheidswaarde. Wat betreft het voorstel dat in het vorige hoofdstuk is gedaan over de codering van een discours gerelateerd [uR]-kenmerk op voornaamwoorden, voorspelt de opmerikbaarheid-structuurhypothese correct dat het kenmerk zal worden toegewezen aan openlijke sterke voornaamwoorden en niet aan pro. Onderwerp clitica vertonen meer een dubbelzinnig gedrag, vooral in erfgoedtaal Venetaanse variëteiten. In deze varianten geven onderwerp clitica over het algemeen de voorkeur aan een antecedent met een lage opmerikbaarheidswaarde, behalve in het geval dat een dergelijk antecedent geen argument van het werkwoord is. In dezelfde context vertonen sterke persoonlijke voornaamwoorden een voorkeur voor het antecedent met een lage opmerikbaarheidswaarde, ook al is dit geen argument van het werkwoord. Ik stel voor dat, in het geval van niet-accusatieve constructies, het verschil kan worden toegeschreven aan de theta-structuur van het werkwoord of aan de verwerkingsbelasting in de interpretatie van een dergelijke antecedentierelatie. Daarom hangt de interpretatie van onderwerp-clitica af van een fijnmazig systeem van syntactische, discours- en semantische factoren. Omgekeerd gedragen sterke voornaamwoorden zich als logoforen, in die zin dat hun voorafgaande selectie

290 The syntax of subject pronouns in heritage languages

voornamelijk wordt bepaald door discours gerelateerde factoren, ongeacht de syntactische configuratie.

Hoofdstuk 6 concludeert de dissertatie. Ik vat de theoretische voorstellen samen die in de vorige hoofdstukken zijn gedaan over de structuur en verspreiding van subjectieve voornaamwoorden en over de rol van contact; ten slotte schetst dit hoofdstuk enkele suggesties voor toekomstig onderzoek die naar voren zijn gekomen uit de gegevens die voor de dissertatie zijn verzameld.

Curriculum Vitae

Alberto Frasson was born on January 11th 1988 in Treviso, Italy. In 2007, he started his BA in Linguistics at Ca' Foscari University of Venice (Italy), where he graduated *cum laude* with a thesis on Southern Serbian dialects and the Balkan Sprachbund. His passion for languages and language change led him to continue his studies with an MA in Slavic Languages and Cultures at Ca' Foscari University of Venice (Italy), where he graduated *cum laude* with a thesis on the evolution of Russian and Slovenian compound tenses. In 2017, he began working as a PhD in the ERC 'Microcontact' project at Utrecht University, where his research focussed on subject pronouns in heritage languages. During his PhD, he was able to carry out fieldworks in Bosnia-Herzegovina and Brazil, where he collected linguistic data by heritage speakers of different languages. The data he collected during his fieldwork has been used in this dissertation, the final result of his research.