Dottorato di ricerca in Scienze del Linguaggio
Scuola di dottorato in Scienze del Linguaggio, della Cognizione e della Formazione
Ciclo 23°
(A.A. 2011 - 2012)

Balkan Romance:
Aspects on the Syntax of Istro-Romanian

Tesi di dottorato di Iulia-Georgiana Zegrean, matricola 955503

Direttrice del Dottorato	**Tutrice della dottoranda**
Prof.ssa Alessandra Giorgi	**Prof.ssa Giuliana Giusti**

Abstract:

The present study presents the initial results of the documentation and (tentative) analyses of some aspects of the grammar of Istro-Romanian (IR), an understudied seriously endangered dialect from the Eastern Romance family, spoken in Croatia. The study is based on data which I collected throughout 2009 and 2010. Chapter 1 offers a brief overview of present-day IR, with references to the community and the linguistic identity. Chapter 2 discusses the properties that IR has in common (or according to which it differs) from languages belonging to the Balkan Linguistic Area, with a systematic comparative look at the three other Eastern Romance languages/dialects. Chapters 3 and 4 present original data on the Noun Phrase and Verb Phrase in IR, in the framework of recent generative studies. Finally, Chapter 5 takes a step further into possible formal comparative analyses – with either Romanian or Slavic, by looking at the behavior of clitic elements in IR.

Estratto per riassunto:

Questo studio presenta i risultati iniziali del lavoro di documentazione e dell'analisi di alcuni aspetti grammaticali del dialetto istro-romeno. L'istro-romeno è una varietà romanza precedentemente poco studiata, parlata in Croazia, appartenente alla famiglia delle lingue romanze orientali, seriamente a rischio d'estinzione. Lo studio è basato su dati empirici che ho raccolto nel 2009 e nel 2010. Il Capitolo 1 offre una panoramica sull'istro-romeno oggi, con riferimenti alla comunità e alla identità linguistica. Nel Capitolo 2 vengono discusse proprietà che l'istro-romeno ha in comune con (oppure secondo le quali si differenzia dalle) lingue appartenenti alla Lega Linguistica Balcanica, guardando sistematicamente i dati delle tre altre varietà romanze orientali, per ragioni comparative. Nei Capitoli 3 e 4 vengono presentati dati originali riguardanti i sintagmi nominale e verbale in istro-romeno, secondo l'approccio generativista. Il Capitolo 5 offre una possibile strada per l'analisi comparativa con il romeno e/o il croato di un fenomeno sintattico specifico, ovvero la posizione degli elementi clitici.

Acknowledgements

I would like to express my deepest gratitude to Livija Fable, my main informant, who with great enthusiasm and patience answered all my questions until late at night. Her husband, Redento Fable, was of immense help in mediating the interaction between me and Livija (I am not a native speaker and knew just a few words in IR at the beginning). Special thanks to Livija's sisters Nella and Vanda, and her nephew Goran for extra-sessions of data-collection. I would also like to thank Mauro and Robert Doričić from Žejane, they were the first ones to open the doors for me in establishing contacts with the community, and Zvjezdana Vrzic for having sent me the invitations to join community events.

I am deeply endebted to Giuliana Giusti who first suggested me to focus on IR and has continuously assisted me in collecting and analyzing the data, and to both her and Guglielmo Cinque for believed in me and supported this research on Balkan syntax at the University of Venice. Giuliana has been "tutoring" me ever since I first arrived in Venice, in 2004, as an Erasmus student, all the way to the end of this life-chapter. No words are enough to express my gratitude for her support through these long years of good and bad times.

Also, I must thank professors Anna Cardinaletti and Alessandra Giorgi, whose classes and seminars I have been attending throughout the years I spent at the Department of Language Sciences at Ca' Foscari. My colleagues, PhD students and post-docs, with whom I shared the "Aula Dottorandi", projects, hopes and worries for the future. In particular, I thank Marco Coniglio, for the inspiring collaboration that started as parallel to my own research, and which resulted in three co-authored papers.

Table of contents

Abstract

Acknowledgements

Chapter 1

Istro-Romanian

1.1.   Istro-Romanians: the community and the dialect

Istro-Romanian (henceforth also IR) is a dialect of (Proto-)Romanian[1] spoken in the Croatian peninsula of Istria. The population migrated to the geographical area of Mount Učka before the 16th century (Puşcariu 1926), but it is hardly clear to historians, linguists or anthropologists when exactly they arrived in the region nor where they originate from. Throughout the centuries, IR has been subject to alloglotic influence (Croatian, Slovenian, Italian – Venetian).

Today there are less than 200 fluent L1 speakers the Croatian Peninsula Istria, plus other 1000 speakers around the world. The majority of the speakers are elderly or middle-aged, very few children have at least passive competence. They form the smallest ethno-linguistic groups in Europe, and their number has rapidly diminished throughout the past century[2]. The idiom of this population, Istro-Romanian, is classified as a *seriously endangered language* (UNESCO Atlas of World's Languages in Danger, available at www.unesco.org). In 2007, it was included on the *List of protected intangible cultural heritages* by the Ministry of Culture of the Republic of Croatia. In the past century, the number of speakers of IR diminished rapidly: "At the end of the nineteenth century the number of Vlaški- or Žejanski-speaking villagers in Istria was estimated at around 3,000. By the early 1960s, this number fell to 1,500; by the early 1990s it dropped further to 400, less than a decade later there were an estimated 200 native speakers of Vlaški or Žejanski in the villages, and we estimated their number to around 150 in 2010." (http://www.vlaski-zejanski.com, Z. Vrzić).

The criteria according to which Istro-Romanian has been classified by UNESCO as *seriously endangered* is the following:

- Most of its youngest speakers are largely in the parent or grandparent generation.
- The absolute number of fluent speakers is very small.

---

[1] Proto-Romanian is also known as Common, Primitive or Ancient Romanian. See Section 1.3.1.

- A minority of the overall population of the community speaks the language.
- The language is used in a limited and continuously dwindling number of domains.
- The language is not used in any new domains.
- No practical orthography is available to the community.
- Education is provided in the majority language(s) only and the language is not used in written form or institutional environments.
- Only some members actively support language maintenance.
- There are no language learning materials.

Currently, a project of revitalization coordinated by prof. Zvjezdana Vrzić addresses some of these criteria.

The debate as to the origin and classification of IR is, for many (linguists and non), a central topic. The two majorly acknowledged theories are: *a)* that "the Istro-Romanians would have detached themselves from the common proto-Romanian ethnic and linguistic trunk in the same period as Aromanians (*around the 10th century*). That would mean that Istro-Romanian – like Daco-Romanian, Aromanian and Megleno-Romanian – should be one of the four "historical dialects" of the Romanian language arising from *proto-Romanian*. The adjective "historical" means here that, in spite of genetic and structural similarities between these dialects, there was neither continuous nor regular contacts between the related Romanian groups after the breaking up of the original community." (Kovačec 2009), or *b)* that "the ancestors of the Istro-Romanian should have been detached from the Daco-Romanian ethnic and linguistic trunk *not earlier than around the 13th century*. Consequently, the Istro-Romanian can't be qualified as a "historical dialect" of the Romanian language on an equal footing with the Daco-Romanian, Aromanian and Megleno-Romanian: the Istro-Romanian is simply a Dacoromanian dialect which after the 12th/13th century developed in the western parts of the Balkan Peninsula. By reason of the loss of continuous and regular contacts with the Daco-Romanian ethnic and linguistic corpus, it could be qualified as a "historical dialect" of *Daco-Romanian.*" (Kovačec 2009) (italics mine). [3]

IR has two major varieties, one spoken in Žejane (Northern variety) and one spoken in Šušnjevica and several hamlets around it (Southern variety) (see Fig.2).

---

[3] See Fig.1 for a map of the (historical) dialects of Romanian.

Fig. 1. Map of regions inhabited by speakers of the four dialects of Romanian.

Fig.2. Geographical areas with Istro-Romanian villages, north and south of Mount Učka.

The name of the idiom, Istro-Romanian, does not correspond to how the speakers themselves name it. As a matter of fact, the speakers call their idiom and themselves after the name of the respective village (Žejane/Žeiăn: Žeiăntsi –žeiănski; Šušnjevica/Susńévitsę – susńévski; Nova Vas/Noselo: Novošåni – novošånski; Brdo/Bărda: Briiåni – briiånski etc.). The speakers in the south use the name Vlås/Vlåš (sg. Vlåh) for themselves and vlåski/vlåški for their idiom. Frequently the same names (Vlasi – vlaški) may be used by their monolingual croatophone neighbours. However Croatians more often use the denominations Čiribirci – čiribirski, which have a humorous connotation. In order for the misunderstandings to be avoided, linguists have introducedthe name Istro-Romanian (Rom. istro-română, Germ. Istro-Rumänisch) towards the end of the 19th century. The name reflects the geographical position and the analogy with the names Daco-Romanian, Macedo-Romanian (or Aromanian) and Megleno-Romanian. This artificial denomination

has not been assimilated by the speakers themselves, who continue to use the terms we have mentioned above.

Language and identity are very closely related. In the case of Istro-Romanian, it seems to be a rather controversial matter. We quote Frăţilă (2009):

"Being aware of their identity, the Istroromanians consider themselves Romanians and, when talking with foreigners, they call themselves, in Croatian, *rumuni*. The term can be found in their dialect as well (Petrovici-Neiescu 1964: 196). Other sources (Dahmen 1989: 452; Kovačec 1999: 132) state that the Istroromanians do not have a Romanian national awareness, saying that they are Croatians (especially those living in the north) or Istrians (more than a half of those living in the south) (Filipi 2002: 47). According to a testimony of the Venetian historian Irineo della Croce (1698: 334), they used to call themselves rumeri, which continues, with some phonetic transformations, lat. *romani* (sg. *romanus*)".

It is beyond the purpose of this linguistic study to take any position on the matter.

1.2. How to write Istro-Romanian

The Romanian (i.e. Morariu 1928, Cantemir 1959) and Croatian researchers recording and transcribing texts during the past one century and a half, have each adopted different writing systems, and have reached no agreement on its written norm not even in recent times.

I do not wish to take any position here as to 'how to write' IR. For reasons of coherence with older transcriptions of recorded data, I will mainly follow the system in Kovačec (1971, 1978) and Sârbu (1992). I have chosen not to change the original orthography when using examples from the literature, thus some differences with the transcription of my own data may occasionally be noted. Vrzić (2009) outlined a very recent proposal to unify the writing system, a proposal which has been adopted for the website dedicated to the preservation of the dialect, www.vlaski-zejanski.com:

"This spelling system is based on two main principles: *a*. it is linguistically adequate, i.e., it represents all the distinctive sounds (phonemes) of the language in a consistent

and non-ambiguous manner, and *b.* it is easy to learn by the community members, both its first and second-language speakers, who are already literate in Croatian."

The system proposed by Vrzić (2009) can be consulted in the Appendix.

1.3.   Theoretical interest

1.3.1. Some brief notes on Eastern Romance

Daco-Romanian, Istro-Romanian, Aromanian, and Megleno-Romanian dialects once formed a single language without any significant dialectal variation. This language, which is called Ancient Romanian (*străromâna*), Primitive Romanian (*româna primitivă*) or Common Romanian (*româna comună*), has not been documented in written texts. However, it was reconstructed on the basis of the present day dialects. It is the latter term that is more commonly used (and used here), as it captures the most important aspect about that of the language, namely no significant divergences between the varieties, which developed in a later stage.

An interesting viewpoint, which we will however not dwell upon, is Tagliavini's (1969:37):

"The comparison of the Daco-Rumanian dialect with Arumanian (because Istro-Rumanian is only a branch of the Daco-Rumanian dialect and Meglenitic is a branch of the Arumanian dialect...) reveals the former unity of the primitive language from which the two dialects developed."

The historical relation between the four Eastern Romance varieties is irrefutable. This relation is reflected in all modules of grammar, as we will point out throughout this study, mainly in Chapter 2 when we examine the properties common to languages belonging to the Balkan Linguistic Area that Eastern Romance displays. In particular, we will compare and contrast IR to Romanian, Aromanian and Megleno-Romanian.

1.3.2. The state of the art and the present contribution

There are no formal analyses of the syntax of IR, while some research has been carried out on the other two Eastern Romance dialects (Friedman 1994, Tomić 2004, Campos and Stavrou 2004, Campos 1995, 2005, 2006). Pioneer work needed to be done in order to

have a complete picture of the four dialects of Daco-Romanian so as to make further comparative studies possible. Considering that IR has developed in isolation from any other variety of Romanian for almost five centuries (Puşcariu 1926), its lexical and, to some extent, also its morpho-syntactic modules have long been under alloglotic influence: Slavic (Croatian and especially the Čakavian dialect, Slovenian) and Venetian (cf. Kovačec 1971).

The ultimate goal of the research on IR was to investigate those syntactic phenomena that are of interest from a comparative perspective, either with Romance or with Slavic languages. In this way, the results might add new insights to what are generally viewed as *Balkan Sprachbund* (Balkan Linguistic Area) features (Friedman 1994, Lindstedt 2000, Joseph 2001) and lay the groundwork for other comparative and contrastive analyses. Unlike the major languages of the *Balkan Sprachbund* (Albanian, Greek, Romanian, Bulgarian and Macedonian), minor languages/dialects spoken in the area have rarely been recorded and have not undergone a process of standardiazation. The present will bring its contribution to filling in the gap in the resources available for one such dialect, namely Istro-Romanian, and in its description and theoretical analysis.

The aims of the present study are presented in the bullets bellow. They have been achieved to different degrees of in depth research, given the initial state of the research on Istro-Romanian.

▫ Enrichment of the IR corpora as a result of fieldwork activity, transcribed and glossed;

▫ Providing a formal analysis of the syntax of IR following current linguistic theories (Minimalism, Cartography);

▫ Laying the groundwork for further microcomparative analyses of the four Daco-Romanian dialects and for more global comparative work on natural languages;

▫ Documenting the features which are present in IR and are common to the *Balkan Sprachbund*, as possible contribution to new insights on typological classifications;

▫ Contributing to the theoretical discussion on the contact phenomena which influence language evolution by documenting features associated to language attrition.

Needless to say, this is but a small contribution. Much is to be done in the future in terms of both language preservation, documentation, and, relevant to the purposes of the present study, formal analysis.

1.4.   Previous studies and resources

In this subsection, we shall but mention a few of the most complete studies (i.e. grammars) of Istro-Romanian, and publications of collected linguistic material throughout the past century. [4] These studies can be divided into collection of recorded texts, and descriptive grammars. Apart from monographs, articles dealing with specific aspects of IR as a community (origin, identity, idiom, grammatical aspects) have been published in volumes. We will refer to some of these studies, either for examples (when explicitly mentioned), or previous claims on certain morpho-syntactic phenomena.

The fundamental monographs for the study of the IR dialect are due to Iosif Popovic, *Dialectele române din Istria*, [5] two volumes published in 1909 and 1914 respectively, Sextil Puşcariu in collaboration with M. Bartoli, A. Belulovici şi A. Byhan, *Studii istroromâne*, [6] three volumes (1906; 1926; 1929) and August Kovačec, *Descrierea istroromânei actuale*, [7]1971.

In 1928, Leca Morariu publishes the volume of collected texts *Lu Fraţi Noştri. Libru lu Rumeri din Istrie*, [8]. Between 1929 and 1934, he publishes *De-ale Cirebirilor*, [9] four volumes comprising texts from different geographical areas inhabited by speakers of IR. Traian Cantemir conducted fieldwork in the 1930s. His collection of texts, mainly stories, including a glossary, became a volume in 1959. A few other Romanian philologists ran fieldwork in the same timespan. Their texts came to light as chapters of the *Atlasul Lingvistic Român* (1938-1942). [10]

Relatively recent studies (in the second half of the 20th century) have been conducted by August Kovačec, Radu Flora, Emil Petrovici, Petru Neiescu, Richard Sârbu, the German researchers J. Kramer şi W. Dahmen, Ion Coteanu, Elena Scărlătoiu. Apart from

---

[4] I thank professors Petru Neiescu and Richard Sârbu for having generously shared with me both material which was hard to find (P. Neiescu), and original, yet unpublished texts at the time when we met in 2009 for the "Days of the Vlaški and Žejanski Language" organized by Z. Vrzić. Also, prof. August Kovačec who sent me precious articles which he had published mostly in French and Croatian journals throughout the last four decades of the 20th century.

[5] 'Romanian dialects of Istria' (my translation).

[6] 'Istro-Romanian Studies' (*idem*).

[7] 'The description of present day Istro-Romanian' (*idem.*).

[8] 'To our brothers. The book of the Romanians from Istria' (*idem.*)

[9] 'Of the *Cirebire*' (*idem.*)

[10] 'The Romanian Linguistic Atlas' (*idem*).

documenting oral texts, some of these authors were concerned with the status of the idiom (the language versus dialect debate), the influence of the Slavic neighbouring linguistic varieties, its present state and its chances of survival.

1.5.   Data collection: informants and methodology

The present study differs from previous grammars of IR, as it attempts to give a view on the data from the perspective of recent developments in linguistic theory, in particular generative grammar (cartography and minimalism).

We checked through the data offered by our investigations, in the light of linguistic typology, on the functionality of some morpho-syntactical basic structures of IR.

The first approach to IR was in Trst, in May 2009, when I met and shortly interviewed Mauro Doricić, from Žejane, aiming at obtaining data on the nominal expression (determiners, possessors, quantifiers). Later, I integrated it with other pieces of data obtained through online communication with Mauro's son, Robert, who often consulted his grandmother.

The second and most relevant stage for this study was meeting speakers from the Southern villages and hamlets, in July 2009, at the *Days of the Vlaški and Žejanski Language* event organized in Žejane and Šušnjevica. That event hosted invited talks, but also cultural manifestations, and it gathered many of the community members both from Istria and abroad (mainly North America). On that occasion I met my main informant, Livija, born in Brdo, a speaker of the Southern variety.

The bulk of the data included in this study has been provided by Livija. Her sisters Nelia and Vanda were also occasionally consulted. The fieldwork trips were conducted in January (21st-24th) and June (25th-27th) 2010, the last one with Giuliana Giusti, in Umag (Croatia), where Livija and her family lives. Some of the data obtained in January has been rechecked in June with the same speaker and occasionally with her sisters. However, the obvious shortcoming of the collected material is that it was not checked with a greater number of speakers. Given the context, it would have been nearly impossible to do so in due time, so it must be left for future enterprises.

The fieldwork methodology applied followed the lines in Mathewson (2005) (see also Kriva 2011). The guidelines are given below:

1. Set up an explicit discourse context for the utterance to be tested;
2. Elicit either a translation into the object language of a sentence in the discourse context, or a judgment about the acceptability of an utterance in the discourse context (the Felicity Judgment Task);
3. Record and include negative data, in order to know what is not possible as well as what is.

Most of the interaction was initially done by means of an shared language, namely Italian, and occasionally through Croatian (with the support of Livija's husband Redento, as I am not a speaker of Croatian myself). As a matter of fact, my informants spoke Istrian/Venetian. At a later stage, I could interact in Istro-Romanian, which facilitated a more direct communication to the purposes of data collection.

The material has been organized in major topics, which then became chapters of this dissertation. More data on aspects not directly related to the Noun Phrase, the Verb Phrase or Clitic elements, can be found in the Appendix.

1.6.   The approaches in the present study

The four main chapters of this dissertation have different aims. Chapter 1 deals with Istro-Romanian in relation to the properties shared by languages of the Balkan Linguistic Area. In checking those properties, we follow Tomić (2004). Chapters 2 and 3 are also descriptive in nature, but they are set in the framework of recent generative studies. Chapter 4 provides a more in depth analysis of a particular phenomenon, namely clitic elements, which have been documented along the lines of a pre-set questionnaire.

All of the topic we deal with are in need for finer analyses, which we have to leave for the future.

Chapter 1

Chapter 2

The Romance Balkans:
Balkan Sprachbund properties extended to Istro-Romanian

2.1.   Introduction

Comparative research on languages not necessarily belonging to the same family, but to the same geographical area, has proven that prolonged contact between different linguistic communities eventually influence languages diachronically. We may think of this phenomenon as a much larger representation of bilingual acquisition at early stages: two (or more) (unrelated) languages influence the output of one another, but only within the possibilities offered by UG. The difference is that with language contact, the two (or more) co-habitant linguistic varieties end up influencing each other in such a way that at a later stage a certain phenomenon which was initially found in only one language will have been acquired and integrated into the second one; instead, a child acquiring two languages, will eventually correctly set the parameters of each of the language with little or no influence from one to the other. Needless to point out that in the process of language contact (leading to the sharing of properties by -unrelated- languages) whole linguistic communities are involved, and that it is a very slow and long process involving generations and generations of speakers.

The aim of this chapter is to explore to what extent those properties that are shared by the languages belonging to the Balkan Linguistic Area are to be found in Istro-Romanian. In doing so, we will first examine the following issues, based on the literature on Balkan languages. Firstly, Section 2 addresses the question of whether Balkan comparative syntax is possible, from the perspective of Joseph (2001). Section 2.3 is an overview of the literature on the Balkan Linguistic Area, starting from emphasizing the goals of linguistic research on unrelated languages spoken in a geographically specified area. At this point, definition of a Linguistic Area will be adopted. Next, we will consider the issue of the membership of the Balkan Sprachbund, from the point of view of various authors throughout the 20th century. We

11

will see how some languages have been considered to be "more Balkan" than others, based on how many core properties they share. Section 2.4 will briefly point out some theoretical considerations on comparative studies involving Romanian, so as to then turn, in Section 5, to those properties that Balkan Romance (Romanian, Aromanian and Megleno-Romanian) have been shown to exhibit. In doing so, we will consider the inventory in Tomić (2004), and check whether and to what extend Istro-Romanian shares these properties with the other, much more well-studied, Balkan Romance languages. Finally, some conclusions will be drawn in Section 2.6.

2.2.   Is Balkan Comparative Syntax possible?

Joseph (2001) discusses whether it is possible or even enlightening to do comparative Balkan syntax keeping in mind the aim of UG (establishing the extent to which all natural languages are alike and the ways in which they differ): cross-linguistic comparative studies are "at the heart of the enterprise" of UG. He addresses the following paths which comparative syntax may go along:

a) examining one syntactic phenomenon in various unrelated languages
b) looking at genetically related languages (i.e. Wakernagel 1892: Indo-European word patterns)
c) a combination of the two (specific syntactic properties in related languages) (i.e. Kayne 1981 on clitics in Romance, Weerman 1989 on Verb-Second in Germanic, etc.)
d) "areal" comparative studies: the languages under examination are not (or may not be) genetically related: e.g. Balkan languages (i.e. Rivero 1990, 1994 on verb movement, Terzi 1992 on control phenomena and clitic-climbing, Joseph 1980, 1983 on *tough*-movement constructions in finite subordinate clauses, Rudin 1988 on multiple *wh*-questions, etc.)

Thus, "areal" comparative research has brought important contribution to the syntactic studies. It goes without saying that syntax is not the only field on which research has concentrated. In fact, the most prominent module of the grammar in terms of geographical linguistic influence is the lexicon (i.e. lexical borrowings). However, we will only be concerned with syntactic phenomena in what follows.

2.3.   The Balkan Sprachbund

2.3.1. Goals of research on the Balkan Sprachbund

A *Sprachbund* (Linguistic Area) is a geographically defined convergence area where genetically unrelated languages display significant similarities which could not have developed accidentally on independent grounds (nor by inheritance). For a linguistic area to be considered as such, boundaries are also a relevant factor. Hock's (1991:494) definition takes into account both factors: "Languages which may be quite distantly related or which exhibit no discernable genetic relationship may come to converge to the extent that they form a group that is structurally quite distinct from the surrounding and/or genetically related languages."

I will use the term *Sprachbund* as a technical term. Joseph (1999:fn3) notes that the possible English translations, i.e. "linguistic union" or "language league" are "clumsy and "infelicitous". To him, "convergence area", rather than the more widespread term "linguistic area" is probably the most suitable English denomination for a *Sprachbund*.[1]

Balkanology in general (and Slavicists in particular) have tried to find sources for long noticed parallelisms between non-genetically related languages spoken in the Balkan area, on the one hand, and for particular phenomena internal to Slavic Balkan which however cannot be explained by diachrony internal to Slavic studies. Given the scarcity of historical language sources for some Balkan languages, the research would also concentrate on the language subareas which reflect the contact in the language continuum.

Andrej N. Sobolev has long argued for the hypothesis that the properties common to Balkan varieties are not due to standard, national/majority languages, but rather to the dialects and local varieties which have been in contact for centuries.[2] Thus, the aim of Balkanologists should be that of describing a sufficient number of dialects which would have two consequences: first, it would be possible to have representative corpora of diachronic development of single dialects, and second, to analyze the relations between the linguistic

---

[1] Cf. Campbell, Kaufman and Smith-Stark (1986), for example, who use the term "linguistic area" for the Meso-American languages.

[2] Sobolev's long activity in collecting, analyzing, comparing and mapping data (mainly syntactical phenomena) from major dialects of the Balkan languages materialized in Small Dialectal Atlas of the Balkan Languages" (2005: *Malyj dialektologičeskij atlas balkanskih jazykov. Seria grammatieska. Vol. I.,* Munchen).

varieties along the language continuum. Such an approach will help researchers create hypothesis on the interactions between dialects and on the birth and development of common properties.

Going in this direction, Sobolev (2004) sketches some of the results of the project "Small Atlas of the Balkan Dialects" which he coordinated. Sobolev (2004) analyses the areal distribution of 65 properties of Balkan dialects. This project involved data collection between 1996 and 2000 in eleven villages in Montenegro, Eastern Serbia, Western Macedonia, Eastern and Southern Bulgaria, Central and Southern Albania and North-Western Greece. His analysis suggested that the set of "typical Balkan properties"[3] for should be relativized. This is unsurprising: the more in depth research goes into (dialectal) microvariation, the finer the parameters turn out to be. But if the goal is to individuate "Balkanisms", that is, properties common to all/most Balkan languages, macrovariation may not be fine enough, on the one hand, whereas microvariation may be looking at too subtle differences. Thus, research should set the granularity level parameter by parameter.

The task of Balkan studies is not only to find and describe common properties, but also to account for *when* and *how* have these properties become innovations in some of the languages spoken in the relevant geographical area by contact and (gradual) assimilation from a neighbouring linguistic variety.

2.3.2. Balkan languages

Throughout the decades there has been little debate as to which major languages pertain to the Balkan Sprachbund (as we will see below). Instead, the controversial aspect that emerges from comparing different studies is related to *which and how many properties* do the languages members of the Balkan Linguistic Area share. Moreover, in recent times, dialect studies have refined the classification of Balkan linguistic varieties (whether languages or dialects), thus also refining the list. Currently, the languages / major dialects included in the Balkan *Sparchbund* (cf. Joseph 1999, Tomić 2004) are the following (in alphabetical order):

1) Arly Romani;[4]

---

[3] See sections 3.4 and 3.5 for an overview of the "typical Balkan properties" in the literature starting in the first half of the 20th century.

[4] Romany is the language of Indic Gypsies, belonging to the Indo-Arian genetic family.

2) Albanian, both major dialects: Geg (North) and Tosk (South), but especially Tosk;

3) Bosnian / Croatian / Serbian,[5] especially Torlak dialects of Southeast Serbia;

4) Bulgarian;

5) Modern Greek;

6) (Slavic) Macedonian;[6]

7) Romanian: (Daco-)Romanian, Aromanian, Megleno-Romanian, and Istro-Romanian;[7]

8) Turkish (mostly relevant for the lexical module).

One clarification note is order from the very beginning: "Balkan" in *Balkan Sprachbund* is not to be intended as a *purely geographical* notion. In fact, Schaller (1975) distinguishes between "*Balkan languages*" and "*languages of the Balkans*". The implication is that all "Balkan languages" are also "languages of the Balkans", but not all "languages of the Balkans" are "Balkan languages". The "languages of the Balkans" are considered as such on purely geographic terms, and include the "Balkan languages" *and* [8]: Armenian (spoken in Bulgaria), Circassian (Adygey variety, spoken in Kosovo, former Yugoslavia), German and Hungarian (both spoken in Romania), Italian (spoken in Istria, former Yugoslavia), Judezmo (also known as Judeo-Spanish, a dialect of Spanish spoken in Greece by Iberian Hebrews)[9]; Ruthenian (also known as Rusyn, perhaps a dialect of Ukrainain, spoken in the Vojvodina area of former Yugoslavia); Slovak (spoken in a small enclave in Vojvodina), and Slovenian.

---

[5] For the purposes of this introduction I will not distinguish between Bosnian, Croatian and Serbian as separate languages, and generally employ the label Serbo-Croatian (SC), unless the differences between the three become relevant . In that case, I will specify which language I am making reference to.

[6] (Slavic) Macedonian is to be distinguished from Ancient Macedonian, spoken in the first millennium B.C. in the Balkans, but presumably related to Ancient Greek. English generally uses two different terms for the two: Slavic Macedonian is referred to as Makedonski, while Macedonian refers to the ancient language.

[7] I will briefly discuss below the degree to which each of the four manifest core Balkan properties.

[8] The languages of the communities that are due to recent are not included in this list of languages of the Balkans, for obvious reasons.

[9] Cf, for instance, Montoliu and Auwera (2004), for phenomena which indicate that Judeo-Spanish may have been influenced by Modern Greek (such as vagueness reading between present and past irrealis, mood harmony with imperfect apodosis constructions), thus sharing some properties typical for Balkan languages (including Turkish, a *peripheral* Balkan language, according to the broad view).

Throughout the past century, many researchers have noted that not all Balkan languages are situated on the same level within the Balkan Lingustic Area. The main criterion against which the various degrees of membership have been discussed was to what extent does a language share core properties with most other languages in the *Sprachbund.* Although terminology differs,[10] the same sense can be caught in a series of classical studies on the topic, as reported by Tomić (2004):

- Weigand (1928): Albanian, Romanian and Bulgarian are *typically Balkan languages*, whereas Greek, Serbian and Greek are only *geographically Balkan.*

- Sandfeld (1930): Greek, Bulgarian, Albanian and possibly Serbian typically exhibit Balkan properties, while Turkish shares many lexical properties with each of the other languages.

- Schaller (1975): Albanian, Romanian, Bulgarian and Macedonian are Balkan languages of the *first degree,* Greek and Serbian – of the *second degree*, Turkish – of the *third degree.*

- Birnbaum (1968): Romanian and Aromanian are "most Balkan", followed by Bulgarian, Macedonian, Modern Greek and Albanian (in this order).

- Solta (1980): northern Greek dialects and southern Serbian dialects can be treated as true members of the Balkan Linguistic Area.

The membership question thus appears to be receiving different answers depending on the criteria and tools that each research has used (a matter which we will not look more deeply into here).

---

[10] Note however that the studies mentioned here do not consider all and the same languages. Also, if we take Serbian, for instance, Weigand (1928) considers it as only *geographically* Balkan, Sandfeld (1930) includes it among typical Balkan languages, but only marginally with respect to Greek, Bulgarian and Albanian, for Schaller (1975) it is a Balkan language of the *second degree*, Birnbaum (1968) does not include it in his survey, while Solta (1980) considers only southern Serbian dialects as being *true members* of the Balkan *Sprachbund.*

The second crucial observation in relation to a linguistic area is that the languages typically belong to different genetic families. Specifically, the languages that are spoken in the Balkans and which share properties in all modules of the grammar (to a different extent), belong to five distinct branches of the Indo-European languages:

- Albanian
- Modern Greek
- Romance languages (Romanian, Aromanian, Megleno-Romanian and Istro-Romanian)
- Slavic languages (Bulgarian, Macedonian, Serbo-Croatian)
- Indo-Iranian (Arli Romany/Gypsy)

As already pointed out, not all of these languages have the same (number of) features in common. When it comes to standard (i.e. national) languages, it is even more difficult to establish which ones are "the most Balkan" ones, intended – which ones share most properties, also because we would not deal with varieties that have developed (especially) through contact among speakers (it suffices to think about the role that national academies have played and still play in establishing norms). Let us say that roughly Balkan languages can be divided into three groups:

1. Albanian, Romanian and Bulgarian have the most properties in common;
2. Serbo-Croatian and Greek seem to share with the other languages a lower number of properties;
3. Turkish – (uncontroversially) shares mainly vocabulary.

Finally, we should mention that while it nowadays appears evident that Balkan languages form a linguistic union in the sense we have described, it hasn't always been granted the deserved attention from a theoretical linguistics viewpoint. We will only refer the reader to Graur (1936), who criticized openly the concept of "Balkan linguistics" in the same sense as "Romance linguistics" or "Indo-European linguistics". His claim was that what was at stake

instead was relationships of borrowings or of mere influences, which he viewed as "superficial" and not "essential" for classification of relations amongst different languages.[11]

2.3.3. Linguistic continuum: influences and changes

The process which has brought genetically unrelated languages spoken in the geographical area of the Balkans to share properties may have, in principle, one of the following two explanations: a "single donor" which had massive influence on all other languages, or a "convergence" hypothesis, by which it is contact between different languages that gradually led to sharing not only part of the vocabulary through borrowings, but also phonological, morphological and syntactic phenomena.

A possible "single donor" hypothesis for the Balkan *Sprachbund* properties has failed to convince researchers in Balkano linguistics. What is generally acknowledged is the so-called "convergence" hypothesis[12]. This approach is centered around the concept of inter-communication between different linguistic communities which *share* the same environment, thus which are in contact and need to facilitate inter-translatability. Multilingualism is thus, in this view, the main trigger for those changes in languages which reflect properties of co-habitant or neighbouring languages (among other changes which take place diachronically in natural languages). Lindstedt (2000) describes this phenomenon as a shared drift: the parallel changes are not a result of calques from one language (for instance, the most prestigious one) to all the others, nor developments from a single substrate (against the "single donor" hypothesis).

In fact, although Greek was the most prestigious language in use throughout the period in which the Balkan linguistic properties were developing, the fact that the number of Greek phenomena that percolated in the other languages is quite low is not surprising. Furthermore,

---

[11] "Remplacer dans la linguistique la notion de parenté par celle d´ "affinité", comme on veut le faire maintenant, c´est accorder a la phonétique et même au vocabulaire et la syntax le pas sur la morphologie; c´est, par conséquent,remplacer l´essentiel par le superficiel" (Graur 1936, *apaud* Du Nay 1996:127).

[12] I find Joseph's (1999:3, fn8) observation very relevant for the "convergence hypothesis":

> "[O]ne of the most fascinating aspects of the Balkan Sprachbund is that virtually all of the significant structural convergences represent also significant divergence from earlier states of each of the languages."

See also fn3.

among Balkan languages, it is not the case that "the number of Balkan *Sprachbund* properties is greatest, probably because, for the native speakers of Greek, the need of changes for the sake of communication has not been very urgent" (Tomić 2004:3,fn2). On the other hand, "Balkan Romani is spoken by a relatively small number of non-native speakers; accordingly, Balkan Romani has a relatively small number of Balkan Sprachbund properties" (*idem*).

Uncontroversially, the properties tend to be more numerous in those regions of the Balkans where the highest number of different linguistic communities co-habit. According to Tomić (2004:3), the nucleus of Balkanisms is to be located in the southern parts of the lakes Ohrid and Prespa, where Greek, Albanian, Macedonian, Aromanian and Balkan Romani intersect. Tomić further notes that even the regional and local dialects spoken in this area share a great number of syntactic phenomena, alongside with lexical items.

In what follows, I will briefly run a literature overview on the linguistic phenomena which have been (traditionally) assumed to be typical of Balkan languages.

2.3.4. Establishing membership: the shared properties criterion

The Balkan Linguistic Area is considered such on the basis of the areal distribution of properties across languages and dialects spoken in this part of Europe. Theoretically, for a property to be typologically areal it needs to be:

        a)   shared by at least three languages in the area, of which at least two must belong to different genetic families,

but at the same time the property must NOT be

        b)   present in <u>all</u> the languages of the genetic family to which the language of the area belongs.[13]

Assessing the typological properties on the basis of which membership could be granted is a quite difficult matter, and studies have had different views on if and how it could be achieved (cf. the references cited throughout this chapter). However, we could generally state

---

[13] Tomić (2004:4, fn5) refers to Birnbaum (1965), who points out that there is no ban on two or more languages belonging both to the same genetic family and to the same *sprachbund*. In fact, this is the case for instance with Romanian and the dialects of Romanian, all of which belong to the Balkan Linguistic Area. Also, Bulgarian, Macedonian and Serbo-Croatian are genetically related.

that the linguistic discussion on the Balkan *Sprachbund* membership has centered around several properties at the macro-level, listed below:[14]

- *phonology*: the presence of the schwa vowel, /ə/;
- *morphosyntax*:
  - substitution of synthetic declension markers by analytic ones;
  - grammaticalization of the category of definiteness through postpositive definite articles;
  - pronominal doubling of objects;
  - analytic expression of future tenses;
  - analytic perfect tenses with "have"-auxiliary;
  - (almost complete) loss of the infinitive and its substitution by subjunctive.

In what follows, I will summarize the Balkan properties according to some of the most influential studies in the previous century.

2.3.5   20th century studies of Balkan properties

2.3.5.1.   Sandfeld (1926/1930)

Sandfeld (1926/30) classified the over one hundred properties that he has recorded into *general concordances* (in his terms) and *concordances between different, individual Balkan languages.* He exemplified the properties with data from Albanian, from Balkan Romance – mostly (Daco)Romanian, but also Aromanian and Megleno-Romanian, from Balkan Slavic – Bulgarian and Macedonian (to which he refers as "Bulgarian speech in Macedonia"), and from Modern Greek. The general properties in Sandfeld (1926/1930) are reported in Tomić (2004):

- *morphosyntax general concordances:*
  - postpositive articles;
  - loss of the infinitive;
  - future formation with the "will" auxiliary;
  - syncretic forms for Genitive-Dative cases;
  - simultaneous use of a "self-standing" and a not "self-standing pronoun" as well as "the use of a pronoun in association with a noun" (i.e. pronominal clitic-doubling);

---

[14] Also see the concluding remarks of this section.

o constructions with a pronominal and a propositional complement, in which the former has same reference as the subject of the latter;

o verbs which take two direct objects;

o loss of distinction between question words which in Latin are represented by *ubi* and *quo*, on the one hand, and *ibi* and *eo*, on the other;

o the use of conjunctions with the meaning 'and' at the beginning of affirmative clauses which follow negative ones;

o use of a paratactic conjunction with the meaning 'and' instead of a hypotactic one.[15]

2.3.5.2. Schaller (1975)

Schaller (1975) divides Balkan properties into *primary* and *secondary* (in his terms). While secondary Balkanisms are treated as phenomena that are restricted in areal coverage and which are not prime in the structure of the language, primary properties are widely spread and cover base constructions in all the languages. Tomić (2004) reports Schaller's (1975) two primary phonological phenomena, and the eight primary morphosyntacic phenomena:

- *phonology*

    o vowel system without quantity, openness and nasality distinctions;

    o the presence of the schwa vowel, /ə/;

- *morphosyntax*:

    o merge of the Dative and the Genitive;

    o postpositive article;

    o analytic comparison;

    o common pattern for constructing the numerals from 11 to 19;

    o loss of the infinitive and its replacement by subordinate clauses;

---

[15]   Parataxis = the juxtaposition of clauses or phrases without the use of coordinating or subordinating conjunctions:

(i)   *It was cold; the snows came.*

Hypotaxis = the subordination of one clause to another by a conjunction.

- o use of the "will"-auxiliary in the construction of periphrastic future tenses;
- o doubling of objects (by "short" pronominal forms);
- o use of the "short pronominal form" for the expression of possessiveness.

2.3.5.3. Solta (1980)

Solta (1980) classifies the six properties of Balkan languages that he treats in three groups: *morphological, syntactic* and *other* (also *different* or *special*)[16] Balkanisms:

- - *morphology:*
    - o existence of the postpositive article;
    - o the merge of the Genitive and the Dative.
- - *syntax:*
    - o the loss of the infinitive;
    - o synthetic expression of futurity.
- - *other*:
    - o the Vocative "as a living category";
    - o the periphrastic comparison of adjectives.

2.3.5.4. Gołąb (1984)

Gołąb (1984) emphasizes the fact that although the different genetic origins of the Balkan languages makes it so that the lexicon is only shared to a very limited extent (i.e. Turkish words), there is a "striking" similarity (or even identity) in grammatical constructions. Alongside with what he calls *negative structural pattern*, namely the loss of the infinitive to the benefit of subjunctive clauses, he lists ten *positive structural patterns* (below). He exemplifies the patterns with data from Macedonian – a Slavic language, and Aromanian – from the Romance family.

---

[16] According to Solta (1980: 223), the phenomena of the third group are not strictly speaking Balkansims, although they are characteristic for the Balkan languages. One notices a contradiction in these two statements, though.

- o future tense formation with a particle which etymologically represents the third singular of the verb *want* + subjunctive mood of the lexical verb;
- o the so-called "future-in-the-past" ("a pattern derived from the future tense by the replacement of the present tense markers by past tense ones", Gołąb 1984:6);
- o present optative-subjunctive mood (formed by a modal particle + the present tense of the verb);
- o imperfect optative-subjunctive mood (formed in the same way as the present optative-subjunctive mood, but with the imperfect tense of the verb);
- o compound perfect (formed by the present tense of the auxiliary verb have + an indeclinable form of the perfect participle);
- o compound pluperfect (a derivative of the compound perfect, through the replacement of the present tense of the auxiliary by its imperfect tense);
- o "futurum exactum", or future perfect (a derivative of the compound perfect, through the replacement of the present tense of the auxiliary by its future tense);
- o postpositive definite article;
- o dative-possessive as a single morphosyntactic category.

2.3.5.5. Lindsted (2000)

On the basis of the number of Balkan properties which are found in a language, the Finnish linguist Jouko Lindstedt computed a "Balkanization factor" for each Balkan language (which he classifies into five language groups). According to his indices, the Balkan Slavic languages are the "most Balkan", whereas Romani is the "least Balkan".[17]

Language	Score
Balkan Slavic	11.5
Albanian	10.5
Greek, Balkan Romance	9.5
Romani (Gypsy)	7.5

---

[17] Note that Lindsted (2000) does not include Turkish. This corresponds in fact to the most embraced approaches to Balkan languages, which although share a non-trivial amount of Turkish words (due to historical reasons), are similar in such ways that seem to systematically exclude the Turkish language.

As for the shared properties, Lindsted classifies the twelve main phenomena he takes into consideration into three groups: *argument-marking*, *verbal system*, and a one-member group of *analytic comparison* of adjectives and adverbs:

- *argument-marking*:
  - enclitic articles;
  - object reduplication;
  - prepositions instead of case suffixes;
  - dative/possessive merger;
  - goal/location merger;
  - relativum generale.

- *the verbal system*:
  - Aux (+ Comp) + finite verb;
  - *volo* future;
  - future in the past as conditional;
  - *habeo* perfect;
  - evidentials.

2.3.5.6.   Ad interim concluding remarks

Pace the different classification and the different labeling of the phenomena, we cannot but note the convergence of the results which Sandfeld (1926/1930), Schaller (1975), Solta (1980), Gołąb (1984), Lindsted (2000) (among others) have reached.

On the abstract level, the two main points which need to be underlined are: a) there are *general* (Sandfeld 1926/1930) or *primary* (Schaller 1975) properties which are shared by most (or all) Balkan languages, and *secondary* properties which are to be found in a restricted number of languages (most likely due to geographical proximity which translates into direct contact); b) accordingly, some languages are "more Balkan" than others, to use Lindsted's (2000) conclusion.

As we have seen, the most salient Balkan features in the literature have been long observed and fairly widely discussed. They cover all modules of the grammar: vocabulary, phonology,

morphology and syntax. Let us summarize with Joseph (1999) which are the *general/primary* properties taken into account in recent literature.[18]

- *phonology*
  - o absence of "overlay" articulatory features (i.e. length, nasalization, etc.);
  - o the presence of the mid-to-high central vowel schwa, /ə/;
- *morphology*
  - o the enclitic definite article;
  - o invariant future tense marking derived from "want";
  - o syncretic forms for Dative and Genitive case;
  - o analytic comparative for adjectives and adverbs;
- *syntax*
  - o doubling (cross-indexing) of syntactic objects (both direct and indirect) by a clitic pronoun;
  - o the loss of the infinitival complement clauses (replaced with subjunctive forms of the verb);
  - o the use of a special verbal form to indicate confirmativity/evidentiality.

It goes without saying that these features are by no means to be intended as typical *only* of Balkan languages. They are all part of UG thus in principle possible to detect in any language (and in fact we do find them in other unrelated languages; for instance, the enclitic definite article is also typical of Scandinavian varieties and Northern Russian dialects; similarly, future formation with a volitional verb is to be found in English and other Germanic languages). Also, these properties are not necessarily identically realized in *all* languages belonging to the Sprachbund. However, the presence of clusters of these features in the relevant languages is crucial for defining (the extent of) the linguistic area.

---

[18] In Section 2.5 we will examine the syntactic properties which Istro-Romanian may share with the other Balkan Romance varieties.

2.4. The role of Balkan Romance

The role of Balkan Romance in emergence of common Balkan features has been considered by various researchers (see, for instance, Reichenkron 1962, Solta 1980). It is, of course, a difficult task to establish how and to what extent may Balkan Romance have had an influence on other Balkan languages.

Another issue when considering Balkan Romance is the extent to which the four varieties belong to the *primary* or *secondary* Balkan languages. This point can be briefly commented upon with Joseph (1999):

"Actually, in counting Romanian among the Balkan languages, the greatest attention belongs to Aromanian, spoken in pockets in northern and central Greece, in Albania, and in Macedonia, and to Megleno-Romanian, spoken in a few areas in northern Greece; Daco-Romanian is a Balkan language to some extent, but not fully so, and Istro-Romanian is largely irrelevant as far as the Balkan Sprachbund is concerned." [19] (Joseph 1999:3)

"Still, of the varieties of Romanian other than Istro-Romanian, Daco-Romanian is the least like the core Balkan languages — and the most like other Romance languages — while Aromanian and Megleno-Romanian are far more Balkan in their character, at least insofar as the perfect and the preterite tenses are concerned." [20] (Joseph 1999:13)

One piece of argumentation for the prime role of Aromanian within the Balkan Sprachbund comes from comparative research between Balkan Romance and South Slavic.

Rosetti (1986:324) divides the Romanian groups of population is Northern and Southern (of Danube). He maintains that "[the ancestors of the Arumanians] must be distinguished from the Rumanians in Serbia, recorded during the entire course of the Middle Ages in the Serbian kingdom. The language of the Rumanians in Serbia, as well as Istrorumanian, presents characteristic features of Daco-Rumanian and belongs to the northern group of the Rumanian

---

[19] Underlined characters are mine.

[20] The fact that this quotation refers to a specific property is indicative of the more general trend in considering Romanian as a *primary* Balkan language. (Underlined characters are mine).

language while Arumanian constitutes its southern group." This view is consistent to that of Balkanologists, who assign a primary role to Aromanian in the development of the Balkan Linguistic Area.

According to Gołąb (1984), "there is no doubt that the Balkan languages, in our particular case Aromanian and Macedonian, show a set of common structural features in their grammars, which leads any Slavist knowing the earlier or the oldest stage of Macedonian, i. e. the Old Church Slavonic, to the following question: what is the source of these structural features in the South Slavic languages?" Gołąb's belief was "that this was continental Balkan Romance, or more specifically in the case of Macedonian, the primary Arumanian" [21] (Gołąb 1984: 9).

Along similar lines, Sobolev (2008) argues that along with mutual lexical borrowings, some structural innovations in Eastern Serbian and Macedonian dialects which are not to be found in other Eastern Balkan varieties may be "grammatical aromanianisms (recent loan-translations and calques)" (Sobolev 2008:115)[22].

Tomić (2004), however, questions the primary of the role of Aromanian (as suggested by previous studies) in determining for the rise of Balkan common features: "But why should a language spoken by a limited number of inhabitants of the peninsula be the source for these calques? Moreover, since the Balkan *Sprachbund* features are not typical for all Romance languages, their provenance in Aromanian would still be in need of explanation" (Tomić 2004: 2)

Furthermore, Sobolev (2008:113) notes that "the linguistic appearance of the Western Balkans does not seem to be identical to that of the Eastern Balkans. Some structural innovations in Eastern Serbian and Macedonian dialects which are not present in the Eastern Balkan area may have Aromanian as their source language."[23] However, Sobolev considers it

---

[21] Underlined characters are mine.

[22] *Idem.*

Drawing the data mainly from the "Small Dialectal Atlas of the Balkan Languages" (2005: *Malyj dialektologičeskij atlas balkanskih jazykov. Seria grammatieska. Vol. I.,* Munchen), Sobolev (2008) mainly takes into account the following structures: prepositional direct object, the reformation of the periphery of instrumental case meanings, the direct object taken by the verb "to be".

[23] For Sobolev (2008), the Balkan language continuum can be divided into two different language subareas: the Western Balkan Linguistic Area and the Eastern Balkan Linguistic Area. The languages classified as Western

premature at the present state of research to make "these kinds of strong generalizing statements concerning the role of Balkan Romance languages and especially of Aromanian in the formation of Balkan Slavonic or common Balkan language structures" (Sobolev 2008:115).

2.5. The shared properties of Romance Balkan languages: where does Istro-Romanian stand

We have seen that Istro-Romanian has generally not been considered to be part of the core Balkan group. Nevertheless, given the genetical relation with the other Eastern(/Balkan) Romance varieties and also the co-habitation and permanent contact with Croatian (in particular, with the local dialect, Čakavian), it is reasonable to assume that Istro-Romanian may displays Balkan properties. In what follows, we will attempt at establishing "how Balkan" this linguistic variety is. In doing so, I will base the investigation on a recent inventory of properties (Tomić 2004). To this purpose, I will concentrate on Balkan Romance (Romanian, Aromanian and Megleno-Romanian). I will only mention (Serbo-)Croatian when the properties of Istro-Romanian are analogous to what is observed in (Serbo-)Croatian.

2.5.1. Nominal Cases [24]

Across Balkan languages (with the exception of Serbo-Croatian), case inflections on nominal elements are very rarely marked; instead, case relations are represented through prepositions.

a) Balkan Romance (in particular, Romanian and Aromanian) have distinct forms for Dative (and Genitive) for all nouns, and distinct Vocative forms for some nouns. Megleno-Romanian only shows Vocative marking.

Balkan are: Eastern Serbian, Macedonian, Albanian and Aromanian dialects, while dialects of Bulgarian and Greek are classified as Eastern Balkan.

[24] The Aromanian and Megleno-Romanian examples in this section are taken from Tomić (2004). Some Romanian examples are also hers. The Istro-Romanian ones are either taken from the literature (when specified), or my original data, in which case no specification is mentioned.

b) Serbo-Croatian displays full paradigms for all cases, while the Locative case is expressed prepositionally plus Dative on the nouns when referring to locations, and Accusative when referring to direction, respectively.

c) Istro-Romanian: the northern variety, Zejanski, has preserved synthetic marking for oblique (Dative = Genitive), while the southern variety, Vlaski, resorts exclusively to prepositional marking for oblique. None of the two varieties marks Accusative case (neither prepositionally, nor inflectionally).

2.5.1.1. Vocative case forms

a) In Romanian, we note that only personal nouns and common nouns denoting humans (in the singular form) have distinct Vocative markers. With plural nouns, the inflection corresponds to the paradigm for the Dative case.

(1) a   *Popescule!*   (Nom. *Popescu*)
       Popescu-the.M.Sg.Voc

  b   *Bunico!*   (Nom. *bunică*)
       grandma.F.Sg.Voc

  c   *Fetelor!*
       girls.F.Pl.Voc/Dat

Aromanian and Megleno-Romanian also display distinct inflections for Vocative singular human nouns (Tomić 2004:11 notes that the use of the Vocative for the feminine is pejorative). The forms are identical in the two dialects:

(2) a   *Profesore!*   (Nom. *profesor*)
       professor.M.Sg.Voc

  b   *Soro!*   (Nom. *soră/ã*)
       sister.F.Sg.Voc.Pej

b) Kovačec (1971:107) notes that in the northern variety of Istro-Romanian, masculine nouns with Slavic origins are marked for the Vocative with –*e*, while those originating from Latin are marked with –*(u)le*, but he adds that all his informants accept the

alternative marking, too. In the southern variety, some nouns are marked by *–e*, others by *–u* (in the latter case, we may say that it is not Vocative marking, since the form is identical to the definite Nominative form of the noun). As for feminine nouns denoting humans, the Vocative is formed with the morpheme *–o,* just like in the other Romanian dialects.

(3) a.   *Gospodine! / Gospodinule!*   (Nom. *gospodin*)
         man.M.Sg.Voc

    b.   *Soro!*   (Nom. *sora/ę*)
         sister.F.Sg.Voc

2.5.1.2.   Accusative case forms and markers

a) No Balkan Romance language marks Accusative case through inflections. Romanian, however, employs the preposition *pe* for proper names and [+human, +specific] bare common nouns, as well as for [+human, +specific] nouns preceded by the indefinite article.

(4) a   *L-am căutat pe Ion / prietenul meu / un prieten.*
         him.Cl have.1Sg looked.for ACC Ion /friend-the mine / a friend
         'I looked for Ion / my friend / a (specific) friend'.
    b   *Caut consilier / un consilier.*
         look.for.1Sg advisor / an advisor
         'I'm looking for a/any advisor'.

b) IR does not mark Accusative in any way: there is no (equivalent) preposition like *pe*. Moreover, clitic doubling for proper names and [+human, +specific] nouns is not displayed (compare with Romanian).

(5)   *Poč vedę Lara.*
      can.1Sg see Lara.
      'I can see Lara'.

2.5.1.3. Dative/Genitive case forms and markers

a) In Romanian and Aromanian only definite nouns have oblique case forms, which are identical for Dative and Genitive. In Aromanian, the obligatory prepositional case marker *a* precedes the inflected noun.

(6) a *Maria i-a adus copilului un cadou.* (Rom.)

Maria 3sg.Dat.CL has brought child-the.M.Sg.Dat a present

'Maria brought the child a present'.

b *cadoul copilului*

present-the child-the.M.Sg.Gen

'the child's present'

(7) a *Petre lji are dată lilice a featiljei.* (Ar.)

Petre 3sg.Dat.CL has given flower to girl-the.F.Sg.Dat

'Petre has fiven a flower to the girl'.

b *sor-sa a profesorlui*

sister his.M.Sg.CL to professor-the.M.Sg.Gen

'the professor's sister'

Like Slavic Balkan, Megleno-Romanian resorts exclusively to prepositional marking for Dative and Genitive. Indirect objects are introduced by the preposition *la* 'to/at' when animate, and by *di* 'from' if inanimate.

(8) a *Petre ăi deadi la feata flor.* (MR)

Petre 3sg.Dat.CL give.3Sg.Aor to girl-the.F.Sg flower

'Petre gave a flower to the girl'

b *capu la profesoru / maţa*

head-the.M.Sg at professor-the.M.Sg / cat-the.F.Sg

'the head of the professor / cat'

c *picioru di masa*

leg-the from table.the.F.Sg

'the leg of the table'

a) In IR, we note that feminine common and proper nouns also have a specific inflection, different from the definite article. Masculine common nouns are inflected with (what looks like) the definite article, while masculine proper nouns are uninflected. The prepositional marker *lu* is obligatory.

(9) a   *Dat-a lu sore / fråtele / Nele regålu.*
         given has to sister.F.Dat / brother.M.Dat / Nela.F.Dat present-the
         '(S)he has given the present to the/her sister /brother / Nela'.

    b   *fiju lu sore / lu fråtele / lu Dejan*
         son of sister.F.Sg.Gen / of brother.M.Sg.Gen / of Dejan
         'the sister's son / brother's son / Dejan's son'

In the northern variety, synthetic marking (similar to Romanian) both on feminine and masculine definite nouns is still preserved. An optional extra prepositional marker, *a*, can sometimes be found.

(10)   *omului / (a) lu omu*
         man.M.Sg.Dat/Gen / to of man-the.M.Sg

2.5.2. Definite articles

Tomić (2004:18) briefly synthesizes the position and use of the definite articles in Balkan languages:

"All the Balkan languages other than (standard) Serbo-Croatian have articles. In Modern Greek and Arli Balkan Romani the articles are pronominal words that inflect for person, number and case, in the other Balkan languages they are clitics that encliticize to the noun, of the noun is the only DP constituent, and to its pronominal modifier(s) otherwise".

a) When two (coordinated) adjectives co-occur, Romanian definite articles appear on both modifiers, while Megleno-Romanian and Aromanian articles only appear on the first element.

(11) a   *băiatul*   (Rom.)

        boy-the.M.Sg

        'the boy'

   b   *micul şi neastâmpăratul băiat*

        little-the.M.Sg and restless-the.M.Sg boy

        'the little restless boy'

(12) a   *porcu*   (MR)

        pig-the.M.Sg

        'the pig'

   b   *micu alb porc*

        little-the.M.Sg white pig

        'the little white pig'

(13) a   *feata*   (Ar)

        girl-the.F.Sg

        'the girl'

   b   *mushata mintimenã featã*

        beautiful-the.F.Sg clever.F.Sg girl

        'the beautiful clever girl'

b) Istro-Romanian also has enclitic definite articles.[25] Unless the adjective is contrastively marked and postnominal (14 c), the definite article is completely dropped (14 b).

---

[25] Chapter 3 will discuss in more detail the distribution and interpretation of the definite article. One note is that the presence of the definite morpheme does not always indicate definiteness/specificity. Also, the occurrence of the definite article on preverbal adjectives is fairly restricted, and the language resorts to different strategies. See Chapter 3.

(14) a   *porcu*   (IR)

             pig-the.M.Sg

             'the pig'

     b   *mic åb porc*

             small.M.Sg white.M.Sg pig.M.Sg

             '(the) small white pig'

     c   *porcu mic ši åb*

             pig-the.M.Sg small.M.Sg and white.M.Sg

             'the small white pig'

More will be said about definiteness (and the interaction with demonstratives, when present) in a dedicated part of this work, namely Chapter 3.

2.5.3.   Pronominal clitics

All Balkan languages (except for Arli Romani) have full and clitic pronouns for Dative/Genitive and Accusative. Across the Balkan Sprachbund, pronominal clitics form clusters with other clitic(-like) elements: auxiliaries, negation, the subjunctive marker.[26]

In Balkan Romance, pronominal clitics (and the clitic cluster in which they appear) are syntactically preverbal and phonologically proclitic.

We shall not dwell on this topic here. Chapter 5 is entirely dedicated to clitic elements (both pronominal and verbal), dealing with first hand data and a possible thread of analysis. We thus refer the reader to Chapter 5.

2.5.4.   Subjunctives

As discussed by many authors, one of the most widespread property of the Balkan *Sprachbund* which opposes this group to many other Indo-European languages (including languages belonging to the same genetic family) is the loss of the infinitive, which has been replaced by subjunctive constructions.

---

[26] In Romanian, there is a very small class of so-called 'semiadverbs' 'or clitic adverbs' which are confined to a specific position in the string of clitic elements are obligatorily adjacent to the verb in Romanian.

a) Romanian, Aromanian, Megleno-Romanian have almost completely lost the infinitive constructions.

b) Infinitival constructions are less limited in Istro-Romanian.

c) Serbo-Crotian infinitives are found in constructions in which subjunctives can also appear.

The following subsections take into account infinitival subordinate clauses as complements to different types of matrix predicates (modal and intentional verbs) and infinitivals as sentential modifiers (relatives and adverbial clauses). Finally, we will briefly consider bare subjunctives as root contexts.

2.5.4.1. Complements of modal verbs

a) Modal verbs in Romanian are either impersonal (*a trebui* 'must'), or they must agree in person and number features with the main verb – since they share the same subject. In unmarked contexts (answers of out-of-the-blue questions with wide scope), the subject surfaces to the left of the main verb. *A putea* 'can' and the root modal *a şti* 'can/to be able to' are optionally able to take a bare infinitve as a complement in standard Romanian.[27] To our knowledge, Megleno-Romanian (and Aromanian) all modal verbs agree with the main verb and never take the infinitve (cf., for instance, Ammann & van der Auwera 2004, Hill & Tomić 2008 for recent generative research on the topic).

(15)   *Trebuie să ştii citi.*   (Rom.)

must.Impers Subj.Mark know read.Ing

'You must be able to read'.

(16) a   *Poţ s-ti duţ.*   (MR)

can/may.2Sg Subj.Mark 2Sg.Refl.CL go.2Sg

'You can/may go'.

   b   **Poţ s-il duca.*

    can/may.2Sg Subj.Mark 3Sg.Refl.CL go.3Sg

---

[27] *A putea* + infinitive is preferred in Transylvania.

b) Istro-Romanian modal verbs *moręi* 'must', *putę* 'can/may', *vrę* 'want/will' always agree in person and number with the subject and select an infinitival complement. There is also an impersonal root (deontic) modal *treba* 'must', clearly related to Romanian *a trebui*.[28]

(17) a   *Lęmnele mores fi lunž.*   (IR)   (Kovačec 1971:175)

       wood-the.M.Pl must.3Pl be.Inf long.M.Pl

       'The wood must be long'.

    b   *Atunča treba lęmne pripravi de iărna.*   (Sârbu & Frăţilă 1998, Glosar)

       then must.Impers wood prepare.Inf of winter

       'Then we(generic)/one must prepare the wood for the winter'.

2.5.4.2.   Complements of intentional verbs

Crosslinguistically, three classes of intentional verbs can be distinguished: the 'intend' class, the subject of which must be co-referent with the subject of the embedded verb; the 'order' class, the referent of the subject must be disjoint from the referent of the embedded subject; the 'want' class, the complement of which may either have a co-referent or a disjoint subject.[29] We will illustrate the three classes for Romanian (embedded subjunctive) and the 'want' class for Istro-Romanian (embedded infinitive). [30]

(18) a   *Intenţionez să plec.*   (Rom.)

       intend.1Sg Subj.Mark. go.1Sg

---

[28] Etimologically, *treba / a trebui* have Slavic origin (cf. also Serbo-Croatian and Slovenian *treba*).

[29] Needless to say, in other Indo-European languages these verbs of intentionality select for infinitival complements (with null subjects). Cf. English:

   i. I_i intend PRO_i to go.

   ii. I_i ordered PRO_{i+} / him_j to go.(the index 'i+' indicates that the plural referent must include the antecedent; it is an instance of obligatory partial control)

   iii. I_i want PRO_i / him_j to go.

[30] We do not have data for the *order* class in IR:

'I intend to leave'.

b   *Am ordonat ca el să plece.*

have.1Sg ordered Subj.Mark he Subj.Mark go.3Sg

'I have ordered that he should leave'.

c   *Vreau să plec / să plece.*

wanr.1Sg Subj.Mark go.1Sg / Subj.Mark go.3Sg

'I want to leave / him/her to leave'.

(19) a   *Vreţ čevå popí?*   (IR)   (S&F 1998:116)

want.2Pl something eat.Inf

'Do you want something to eat?'

b   *Io-åş vrę-nca se rient'e cuvínta.*   (Sârbu 2009:60)

I aux.1Sg.Opt want still Refl forward speak.Inf

'I'd like people to speak (the language) in the future.'

Tomić (2004:40) notes that Croatian intentional verbs take either subjunctive or infinitival complements, depending on the (dis)joint reference of the matrix and embedded subjects. Thus, the 'order' class will take a subjunctive, and so will verbs in the 'want' class when the two subjects do not co-refer. The 'intend' class will always select for an infinitival complement.[31] From the data available to me, it does not appear that such a distinction can be consistently made for IR.

2.5.4.3. Subjunctive constructions as sentential modifiers

a.   Relative clauses

Relative clauses with a subjunctive verb are (optionally, depending on the language) introduced by a *wh-* relative item.

a) Romanian may or may not display and overt relativizer; Megleno-Romanian and Aromanian cannot omit the *wh* item.

---

[31] She also notes that (standard) Serbian differs from (standard) Croatian, in that for the 'intend' class both subjunctive and infinitve are available, with a preference for the subjunctive.

(19) a   *Caut o fată (care) să ştie englezeşte.*   (Rom.)

look.for1Sg a girl which Subj:Mark know.3Sg.Subj English

'I'm looking for a girl that knows English.'

b   *Ubides capelă cari s-mi veagljă di soare.*   (MR)

look.for.1Sg hat which Subj.Mark -1Sg.Acc.CL protect.3Sg.Subj from sun

'I'm looking for a hat that would protect me from the sun.'

c   *Aflai tu sone una fustane cai/tsi s-mi ariseascã.*   (Ar.)

find.1Sg.Aor   at   end   a   dress   which/what   Subj.Mark   -1Sg.Acc.CL

please.3Sg.Subj

b) This set of data cannot be replicated in Istro-Romanian, which does not resort to a subjunctive form of the embedded verb.

(20) *Ver âŋca făče calunu cu če ver strilei la Iardåsi.*   (IR)   (K 1971:189)

want.2Sg still do.Inf cannon-the with what want.2Sg fire.Inf at Iardasi

'You moreover want to build a cannon with which to fire Iardasi.'

b.   Adverbial clauses

When introduced by equivalents of the preposition *for*, the subjunctive clauses function as adverbial purpose modifiers.

a) Romanian:

(21) *Ion e prea periculos pentru ca sa-l angajăm.*   (Rom.)

Ion is too dangerous for that.Mod Subj.Mark -3Sg.M.Acc.CL hire.1Pl

'Ion is too dangerous for us to hire him.'

b) Istro-Romanian also employs subjunctive mood (marked by *neca* plus indicative, or *se* plus indicative or aorist)[32] for purpose adverbial clauses. Alternatively, the equivalent of

---

[32] In IR the subjunctive mood is formed by *neca* (or, rarely, *se*) and the indicative form of the verb.

the complementizer *that* may introduce a purpose adverbial with an indicative mood (a construction that is not available in Romanian).

(22) a   *Atunče l-a zdrenit din mašuna fåra ...*   (IR)

        then 3.Pl.M.Acc.CL have.3Pl taken.out from stable out…

        neca påscu pâr la sera.   (K 1971:192)

        Subj.Mark feed.3Pl until at evening

        'Then they took the out of the table so that they would feed until the evening.'

  b   *Mâre sęra viro tu se na rem pogovarui.*   (K 1971: 193)

        Tomorrow evening come.2Sg.Imp you Subj.Mark 1Pl.Refl.CL aux.1Pl.Aor talk

        'Come tomorrow evening so that we talk.'

  c   *Ši atunče a mes omu nuntru ke-l va zvadi.*   (K 1971: 192)

        and then have.3Sg gone man-the inside that 3Sg.M.Acc.CL will.3Sg take.out

        'And then the man went inside to take him out.'

c.   Bare subjunctive constructions

Across the Balkan languages (but not only, see Aaman and Auwera 2004), a bare subjunctive construction express wishes, intentions, suggestions or mild commands.

  a) Romanian, Megleno-Romanian and Aromanian display analogous constructions with the same meaning and function:

(23) a   *Să te duci!*   (Rom.)

  b   *S-ti duţ!*   (MR)

  c   *S-ti duts!*   (Ar)

        Subj.Mark 2Sg.Acc.CL go.2Sg

  b) Istro-Romanian expresses wishes, suggestions, etc. mainly through what is called "restrictive mood" (present or perfect), which may (but need not) be preceded by

subjunctive markers *neca* or *se* (the latter is also a marker for conditional, cf. Italian *se* 'if') or by *când* 'when'. It is thus not the case that IR also emplys a bare subjunctive construction for these pragmatic meanings.

(24) *(Se / Neca / Când) io reš fost fi coló!*   (IR) (K 1971:148)

   if / Subj.Mark / when I aux.1Sg.Restr been be.Inf there

   'If (only) I had been there!'

From the comparative data taken into account in this section 5.5., we conclude that the use of the subjunctive is less widespread in IR than in the other Balkan Romance varieties. It has not come to replace the infinitive in all and the same contexts. Thus, IR is "less Balkan" in this respect than Romanian, Megleno-Romanian and Aromanian.

2.5.5. Future tenses

Future tenses are expressed analytically in (most) Balkan languages.[33] Three constructions are available[34]: *a)* "will" inflected for number and person plus subjunctive; *b)* non-inflected "will" plus subjunctive; *c)* non-inflected "will" plus finite verbal forms analogous to the forms of the subjunctive in the respective languages.

a) Romanian, Aromanian and Megleno-Romanian differ in which of the three structures they reflect. In particular, Romanian the future is constructed with inflected "will" and a bare infinitival verb, or (colloquially) with the invariant clitic element *o* (arguably a morpheme originating from "will") plus a subjunctive-like structure with the inflected main verb. Thirdly, an impersonal (archaic) construction formed with invariable *va* "will" and a subjunctive construction. Finally, Romanian expresses intentional future tense with a subjunctive construction preceded by the inflected auxiliary *avea* "have". Aromanian displays a subjunctive construction preceded by an invariable form of "will", while in Megleno-Romanian the future tense is expressed by the subjunctive construction alone.

---

[33] These analytic structures have supposedly evolved "from configurations in which subjunctive constructions appear in complement positions of lexical <will>-modals" (Tomić 2004:43).

[34] Though, of course, not all of them are present in every language.

(25) a *Eu voi pregăti cina.*   (Rom.)

I will.1Sg.CL prepare.Inf dinner-the

b *Eu o să pregătesc cina.*

I will.CL Subj.Mark prepare.1Sg dinner-the

'I will prepare the dinner.'

c *Pedeapsa Domnului va să vină.*

punishment-the God.Dat will.CL Subj.Mark come.3Sg

'God's punishment will come.'

d *Eu am să pregătesc cina.*

I have.1Sg Subj.Mark prepare.1Sg dinner-the

'I will (I intend to) prepare the dinner.'

(26) *Va s-yin s-ti ved mãne.*   (Ar.)

will Subj.Mark –come.1Sg Subj.Mark -2Sg.Acc.CL see.1Sg tomorrow

'I will come to see you tomorrow.'

(27) *Si vină.*   (MR)

Subj.Mark come.2Sg/Pl.Subj

'(S)he/they will come.'

b) Istro-Romanian future tense is formed by the inflected form of *vrę* "will" and the infinitive form of the main verb.[35]

(28) *Io voi cântå.*

I will.1Sg sing.Inf

2.5.6. Evidentials

The grammaticalization of structures expressing evidentiality is reflected differently in the various languages.

---

[35] The status of the auxiliary *vrę* will be discussed in some detail in Chapter 4. The unmarked linear order is Verb > *vrę*, if the subject is not overt. When the (pronominal) subject is present, the auxiliary will precede the verb: Subj > *vrę* > Verb.

a) Romanian expresses evidentiality by constructions with the auxiliary "will" plus the invariable element expressing anteriority *fi* 'be' and present or past participles. Aromanian arguably has not grammaticalized expressions for evidentiality. Megleno-Romanian resort to two different constructions, both of which involve inversion: the inverted perfect (with the participle preceding the inflected auxiliary "have"), and the inverted past perfect (perfect of the auxiliary: participle of "have" preceding the inflected auxiliary "have", all being followed by the past participle of the main verb). The first construction expresses either reported or epistemically deduced events, or surprise, admiration, amazement or irony, while the second one expresses "non-witnessed events or actions that have begun at a point a point in the past and last until the moment of speaking" (Tomić 2004:54).[36]

(29) a   *Eu voi fi pregătind cina.*   (Rom.)

        I will.1Sg be.Perf prepare.PresPrt dinner-the

        'I will (presumably) prepare the dinner.'

     b   *Eu voi fi pregătit cina.*

        I will.1Sg be.Perf prepare.PastPrt dinner-the

        'I will have (presumably) prepared the dinner.'[37]

(30) a   *Tu fostai ăncrilată!*   (MR)

        you be.3Sg.Inv.Perf clever

            1.   'You are clever!' (admiration)

            2.   '(I am surprised to see that) you are so clever'

            3.   'You clever!' (irony)

     b   *Zisi ca ţ-li vutau dat.*

        say.3Sg that 2Sg.Dat.CL -3Pl.F.Acc.CL have.3Sf.Inv.Perf given.Part

        '(S)he says that (s)he had given them to you.'[38]

---

[36] She also notes that although structurally parallel to Albanian evidential forms, Megleno-Romanian evidential constructions are used similarly to Macedonian evidentials.

[37] This construction is identical to the future in the past. The interpretation can thus vary between past presumptive and anteriority in the future.

b) To our knowledge, Istro-Romanian does not have grammaticalized structures which express evidentiality.

2.6. Ad interim conclusions

In this chapter, we have presented data relevant for those properties which seem to be common to the members of the Balkan Linguistic Area. Before doing that, however, we offered an overview of Balkan languages as it emerged from the first studies which took into account the similarities between them, and that, on the basis of those similarities and differences, classified the languages spoken in the relevant geographical area as "more" or "less" Balkan.

In our attempt to establish "how Balkan" Istro-Romanian is, we also aimed at pointing out whether the properties under examination were shared with the other Eastern Romance varieties, namely (Daco-)Romanian, Aromanian and Megleno-Romanian. We therefore presented our data in parallel to examples from these three languages/dialects.

Finally, we concluded that Istro-Romanian is the "least" Balkan of the four, and also that, as expected, that it shares an smaller number of properties with those languages that belong to the "core" of the Balkan Linguistic Area.

---

[38] In these Megleno-Romanian examples I have faithfully reproduced the orthography and glosses in Tomić (2004:54, 77b and 78a), although I believe that a more appropriate orthography and gloss for the inverted perfect auxiliary is the following:

    i.   *fost-ai*

            been.PastPrt –have.2Sg

    ii.   *vut-au dat*

            had.Past.Prt have.3Sg/Pl given

These constructions involve Long Head Movement and are also present in other Balkan languages, including Romance Balkan (cf. Rivero 1994). In Romanian, a hyphen separates the participle from the enclitic auxiliary.

Chapter 3

The Noun Phrase

3.1. The noun

This chapter is dedicated to the noun phrase, starting from the nominal features such as gender and number, to the functional categories related to the nominal expression, and finally to modifiers of the noun.

3.1.1. Gender and number

In this section we will briefly mention the morphological patterns for gender and number in IR. We refer the reader who is interested in a more thorough description and classification of IR nouns according to gender morphology and number marking to Pușcariu (1929) and Kovačec (1971). [1]

IR displays three genders: *masculine, feminine* and *neuter.* The Latin neuter disappeared morpho-syntactically in Danubian Latin due to the evolution of nominal endings (ILR 1969:217). Thus, what is called *neuter* in Eastern Romance varieties is *de facto ambigender*: Latin singular neuter was reanalysed as masculine, while plural neuter patterned with feminine.[2] In the case of IR, borrowings from the surrounding Slavic varieties may also display *neuter* morphology (see below).

A noun is specified for gender from the lexicon. This feature is interpretable on the noun, and it is shared with all its modifiers (when present), which must all agree with it.

Masculine nouns may not display any specific morphology (Ø), or can be identified by the ending in *–u, -o, -e,* or by a truncated final syllable.

---

[1] More precisely, see Kovačec (1971:81ff).

[2] Cf. Section 3.2. for the reanalysis of the Latin indefinite article under two genders.

(1)   a. *fečor* 'boy', *lup* 'wolf', *om* 'man, husband', *fil* 'son'

              b. *žeruncl'u* 'knee', *ocl'u* 'eye', *socru* 'father-in-law'

              c. *nono* 'grandfather', *mulo* 'donkey' ³

              d. *câre* 'dog'

              e. *cå* 'horse', *viţé* 'calf'   (cf. Rom. *cal, viţel*).

    <u>Feminine nouns</u> end in *–a* (and the allomorph *–e* following palatals or fricatives) *–e,* and Ø or *ę* in the South.

(2)   a. *fęta* 'girl', *ženska* 'woman', *våca* 'cow', *nona* 'grandmother'

              b. *mul'are* 'wife', *pâre* ,bread'

(3)   a. *mâra* (Ž.) – *mâr* (S.) 'hand'

              b. *fęţę* (S.) 'girl'

    <u>Neuter nouns</u> are borrowings from Croatian. They are morphologically marked by a final *–o*:

(4)   *nebo* 'sky', *zlåto* 'gold'

    As for number morphology, let us sketch the main patterns in what follows. For <u>masculine nouns</u> (and adjectives), number morphology is neutralized in many cases: all masculine nouns ending in the labials /p, b, f, v, m/, some palatals and the rhotic /r/ have the same overt realization for singular and plural. (5.a-b):

(5)   a. SG *lup*   - PL *lup*ₒ   'wolf'

              b.   *fečor* -   *fečor*ₒ   'young man/son'

Those nouns that for the singular end in dentals /t, d, n/, velars /k, γ, h/, alveolar fricatives /s, z/ and the lateral /l/ display the number opposition through allomorphs:

---

³ These nouns are borrowings from Italian/Venetian. In Italian, the *–o* is the masculine singular morpheme for nominal categories (nouns, adjectives, quantifier adjectives, demonstratives, etc.).

(6)   a. SG *sused*   PL *susez*   'neighbour'

        b.   *dråc*   *dråč*   'devil'

There are many other masculine nouns that are very easy to recognize as either singular or plural, given the (almost) perfect parallelism to their Romanian counterparts. Both in Romanian and in Istro-Romanian we can encounter the plural marker /i/ (7.a). The morphological alternation in (7.b) mirrors the Romanian *or/oare* typical for nouns that are masculine when singular but feminine when plural, whereas the alternation *ø/ure* in (7.c) is another illustration of the same masculine singular – feminine plural phenomenon exemplified for Romanian in (7.d):

(7)   a. SG *codru*   PL *codri*   'wood'

        b.   *pičor*   *pičore*   'leg'

        c.   *vârh*   *vârhure*   'tip'

        d.   *vârf înalt*   *vârfuri înalte*   (Rom.)

             tip$_{masc}$ high$_{masc}$ tip$_{PLfem}$ high$_{PLfem}$

Masculine adjectives ending in /p, b, f, v, m/, the fricatives /š, ž/, the rhotic /r/, the affricates /ţ, č/ may or may not display number opposition (8a-b). [4] The obligatory marking of the plural is observed especially due to the influence of Serbo-Croatian and Italian, both of which employ –*i* as masculine plural morphology. [5] In fact it is mainly borrowed adjectives that always mark the plural with –*i* (8c, d). Another case of obligatory plural marking is the one exemplified in (8.e), typically a Serbo-Croatian phonological phenomenon:

(8)   a. SG *bur*   PL *bur(i)*   'good'   (Rom.:*bun,-i*)

---

[4] Kovaček (1971:91) mentions that this list has been built up on theoretical rather than empirical grounds. In fact I have not found in the texts available to me from Sârbu (1992) any clear evidence to support the optionality of adjectival number marking. I will have to postpone this issue until I will be able to have first hand data on the matter.

[5] 'Borrowed' is used here to group together all lexical elements that entered the language through a borrowing process from the other languages (thus except for Proto-Romanian) with which IR has been in contact throughout the centuries. Whether these words do or do not have Latin origins is an independent matter.

b.	*åb*	*åb(i)*	'white'	(Rom.: *alb,-i*)
c.	*fin*	*fini*	'delicate, beautiful'	(Ital.: *fino,-i*)
d.	*gotov*	*gotovi*	'finished, done, ready'	(SC: *gotov,-i*)
e.	*bol<u>ân</u>*	*bol<u>ni</u>*	'ill'	

There are three other cases of separate forms for singular and plural given in (9), clear traces of the (Proto-) Romanian origin of these adjectives:

(9) a. SG *måre*   PL *mår_ø*   'big'   (Rom.: *mare-mari*)

    b.   *viɉu* [6]   *viɉ_ø* 'alive'   (Rom.: *viu-vii*)

    b.   *verde*   *verz_ø*   'green'   (Rom.: *verde-verzi*)

A rather interesting phenomenon involves the masculine singular adjectival forms ending in –*(i)le*. Historically –*le* is the enclitic definite article which in Romanian attaches to certain masculine singular nouns,[7] but in IR it has come to merely indicate adjectival masculine singular morphology.[8]

(10) a. *do taliɉanskile rat*

      until Italian war

    b. *Neca naidę totile ånu.*

      that_{Subj} eat.3Sg all year-the

      'So he could eat all year long.'

                        (ex. from Sârbu & Frățilă, *Glosar*)

As already mentioned, most feminine nouns end in –*a* or –*e* (in certain phonological contexts, i.e. following some palatals or fricatives). Generally, they form the plural in –*e*, but in cases like (11.c) there is no distinction between the singular and the plural form:

---

[6] *ɉ* is a palatal sonorant fricative.

[7] and to feminine plural nouns or adjectives, but we believe this is unrelated to the present remark.

[8] We will come back to this phenomenon in Section 3.2.1 dedicated to definiteness.

(11)   a. SG *cåpra*   PL *cåpre*   'goat'   (Rom.: *capră - capre*)

   b.   *fęta*   *fęte*   'girl'   (Rom.: *fată - fete*)

   c.   *fil'e*   *fil'e* [9]   'daughter'

Other feminine nouns display the –*a* – Ø or –*e* – Ø opposition plus some other phonological alternation:

(12)   a. SG *mâra*   PL *mâr*_ø   'hand'

   b.   *uše*   *uš*_ø   'door'

As a note for the purposes of this section, in IR nouns can show overtly the definite/non-definite opposition. The variation depends on the phonological ending, i.e. masculine nouns ending in –*u* and feminine nouns ending in –*a* are realized identically whether they are definite or non-definite. This issue will be discussed more thoroughly in Section 3.2. *The articles.*

3.1.2. Noun classes

> *Gianni mi-a dat regalu, ne boca de vir.*
> Gianni to.me has given present-the, not bottle-the of wine.
> 'Gianni gave me the present, not the bottle of wine.'

Nouns can be concrete or abstract. Further, both concrete and abstract nouns can be subclassified as singular, plural and collective; proper and common; countable and uncountable (mass) nouns. In what follows, we will illustrate the aforementioned (sub)classes in IR. Where available, we will provide examples for all genders. Also, the plural form is provided, when applicable.

   Roughly speaking, <u>concrete</u> nouns pick out people or objects. The referent can be animate or inanimate. Animate nouns can designate human or non-human entities. Here are some examples: [10]

---

[9] *l'* is a palatal lateral.

[10] For a more complete description of the nouns given in this section, we have also marked the vowel in the accented syllable, and specified the plural form when irregular.

(13) *Concrete objects*

▪animate   [+ human]

F   **ženska / mul'ǻre** 'woman', **fęta** 'girl', **fętiţa /fečorina** 'little girl', **camerista** 'maid';

M   **muški / muščarac** 'man', **fečor** 'boy', **fečorin** 'little boy', **radnic** 'worker', **advocate** 'lawyer' (pl. advocaţ), **doctor** 'doctor/physician', **prevt (**var. **preft)** 'priest' (pl. prevţ), **soldǻt** 'soldier' (pl. soldǻţ).

         [- human]

F   **vǻca,** 'cow' (pl. vǻč), **gal'ira** 'hen' (gal'ir), **cǻpra** 'goat', **mačka** 'cat', *o***ie** 'sheep' *(*pl. oi);

M   **cǻ** 'horse' (pl. cǻl'), **bo** 'ox' (pl. boi), *ǻ***sir** 'donkey' (=pl), **câre**, (pl. câr) /**brec**, (pl. breč) (S) 'dog'.

▪inanimate

F   **galeta** 'bucket', **lingura** 'spoon' (pl. lingur), **cńiga** 'book', **coşul'a** 'shirts', **maşina** 'car' (pl. maşiń), *o***cna** 'window';

M   **carbur** 'coal' (=pl), **cumpir** 'potato' (=pl), **foc** 'fire' (pl. focure);

N   **camion** 'truk' (pl. camione), **oltǻr** 'altar', **cuţit** 'knife'.

    <u>Abstract</u> nouns denote abstract, immaterial notions or concepts. The may be root or derived with suffixes that impart such meaning (i.e. English *happiness*). Examples are provided in (14).

(14) *Abstract nouns*

F   **vera** 'faith/religion', **vesel'e** 'happiness', **gripa** 'flew', **jivot** '(way of) life', **briga** 'worry', **cazioniţa** 'occasion', **criza**, *e* 'crisis', **democraţiia** 'democracy', **fome** 'hunger', **sriča** 'luck';

M   (we could not find examples)

N   **red** 'order'.

    Both concrete and abstract nouns combine with the (in)definite article.

Examples of singular and plural nouns can be seen throughout (13) and (14) above. Some collective nouns are given in (15), while (16) illustrates uncountable (mass) nouns.

(15) *Collective nouns*

F   ***o*ste** 'army';[11]

M   (we could not find examples)

N   **buk*e*t** 'bouquet' (pl. bukete), **n*a*rod** 'people, nation' (pl. narodure).

Collective nouns trigger singular agreement on the verb.

(16) *Uncountable nouns*

F   **c*å*rne** 'meat', ***å*pa** 'water', **s*å*re** 'salt', **far*i*na** 'flour';

M   ***o*l'ie** 'oil', **l*å*pte** 'milk', **put*ă*r** 'butter', **m*e*so '**meat', **café** 'coffee', **o*ţ*et** 'vinegar'.

As for the determiners which mass nouns take, the definite article and the demonstratives are allowed, in which case a specific meaning is obtained.

Finally, the class of <u>functional</u> nouns is exemplified in (17.a-c): quantifier nouns, nouns denoting containers (of), and group nouns.

(17) *Functional nouns*

        a. quantifier nouns

F   **ve*č*inom** 'majority'; [12,13]

M   (we could not find examples)

        b. container

F   **st*â*cla** 'bottle', **ş*å*liţa** 'cup';

M   (we could not find examples)

N   **pak*e*t** 'packet'.

---

[11] Also 'war'.

[12] This noun is only used in Žejane.

[13] For Quantifier expressions, cf. Section 3.3.

c. group

F **cumpan*i*a** 'company/group (of people)';

M **grozd** 'grappolo (d'uva)';

N **buk*et***, *e* 'bouquet' (pl. bukete), **cup** 'pile' (pl. cupure).

Container and group nouns combine with the nouns they modify by means of the functional preposition *de*, and trigger singular agreement on the verb:

(17') *Česta buket de rož je/*scu mušåt.*

    this bunch of roses is/*are beautiful

    'This bunch of roses is beautiful.'

3.1.4. Nouns without argument structure. The R relation

Examples of nouns with a semantic structure are given below. They have been subclassified in: kinship, extended relationships, body parts (nouns that mark inalienable or obligatory possession) and personal objects (alienabile possession).

(18) *Relational nouns*

▪kinship

F **m*å*ia** 'mother', **f*i*l'e**'daughter' (=pl), **s*o*ra** 'sister' (pl. suror/surăr) ,**n*o*na** 'grandmother', **str*i*na** 'aunt (father's side)'.

M **č*å*ie(J)/č*å*če(S)** 'father', **fil'** 'son' (=pl), **fr*å*te** 'brother' (pl. fraț), **n*o*no** 'grandfather' (pl. noń), **str*i*ţ** 'uncle (father's side)' (=pl.), **br*å*tit"** 'little brother', **cumn*å*t** 'brother-in-law' (pl. cumnåţ).

N **rod** 'relative' (pl. rordure).

The examples below illustrate how these nouns combine with possessives. More data is presented in Sections 3.2.3 and 3.2.3.

(18') a. *Lara se va maritå*   *dupa filju lu cralju.* */*dupa lu cralju flilju*

    Lara refl will3Sg wed   after son-the of king-the /*after of king-the son-the

'Lara will marry the king's son.'

   b. *sora lu Lare*   */ sora l'ei*   */ l'ei sora*

      sister of LaraGen   / sister her   / her sister

      'Lara's / her sister'

▪extended relationships

F   **prii*å*tel'i*ţ*a** 'friend', **sos*e*da / sus*e*d*ę*** 'neighbour'.

M   **prii*å*tel'** 'friend' (=pl), **sos*e*d / sus*e*d)** 'neighbour' (pl. sosez / susez), **šef** 'boss' .

▪body parts

F   **g*u*ra** 'mouth' (pl. gure) [14], **i*rima** 'heart' (pl. irime), **gl*å*va** 'head' (pl. glåv).

M   **fr*u*nte** 'forehead' (pl. frunţ), **o*cl'u** 'eye' (pl. ocl'i), **d*i*nte** 'tooth/teeth' (pl. dinţ).

N   **cuv*e*t** 'elbow' (pl. cuveture), **gut** 'throat' (pl. guture).

▪personal objects

F   **c*å*sa** 'house', **borş*i*ţa** 'purse' (pl. borşiţ), **c*a*mbra** 'room'.

M   **brag*e*ş** 'trousers', **o*č*ale** 'glasses'.

N   **ople*č*** 'blouse' (pl.opleče).

(18')   c. *Česta {mę} mižul {*mę} zelen âi pre scånd.*

        this {my} glass {my} green is on table

        'This glass of mine is on the table.'

As for relationship nouns, we refer to Section 3.2.2. and 3.2.3. for more data.

---

[14] The plural marker for these nouns (*-e*) is worth mentioning since it is different from the Romanian equivalent counterpartes: *guri* 'mouths', *inimi* 'hearts'.

3.2. The determiner

Structurally speaking, those elements that are in complementary distribution with the article (be it definite or indefinite)[15] are generally labeled as determiners. They form closed classes and can be described on distributional criteria. Also, in (19.d) we see that they always precede adjectival modifiers of the noun.

(19)   a. *ur (*česta/*čegod) fil'(*u)*

      a   this any   son-the

    b. *nušte (*češti/*urii) cumpir(i)*

      sm' these/some potatoes-the

    c. *(*ur/*česta/*čegod) fil'u*

      a   this any   son-the

    d. *ur/česta mušat (*ur/*česta) fečor*

      a/this handsome a/this boy

In what follows, we will describe the various types of determiners in IR in terms of paradigms and distribution. We will also tackle the question as to how they may have developed (from Latin), and briefly mention the current trend of analysis as to their role and respective positions in the syntactic architecture.

3.2.1. The articles

3.2.1.1. The indefinite article

In all Romance languages, the indefinite article is a free morpheme preceding all the elements in the nominal expression. Allegedly, it evolved in the same way from Latin into the different Romance varieties.

    As for Eastern Romance, the first evolutionary phase, the 'Danubian Latin', did not have an article system (cf. Nicolae 2009, quoting Iordan & Manoliu 1965). At that stage, Latin *unus* had a weak quantitative meaning, representing a preliminary stage in the process of grammaticalization. The article system of Eastern Romance (both the definite and the

---

[15] This is however not crosslinguistically true for all structures. For instance, Romanian allows the co-occurrence of the definite article and the demonstrative, i.e. when the noun+article precedes the demonstrative.

indefinite article) seems to have developed later, during the stage in the language usually referred to as 'the Common Romanian', a stage which preceded development and split of the four major dialects (Daco-Romanian, Aromanian, Megleno-Romanian and Istro-Romanian), i.e. before the 9th century.

The grammaticalization process of the numeral *unus* into the indefinite may have begun when *quidam* (which had the functions of an indefinite article in Latin) disappeared, according to some authors (cf. Nicolae 2009 and references therein). It seems more natural to assume that it has been a gradual process, tracing back to 4th century Latin, when "*quidam* slowly disappeard and *unus* weakened its use from an indefinite [+specific] meaning to an indefinite [-specific] meaning, replacing in the process of communication the word *quidam*" (Nicolae 2009:4).

Heine & Kuteva (2006:104-105) identify five stages of grammaticalization of the numeral into the indefinite article, from the 4th century (stage I, when the *unus* functioned as a numeral with pronominal and adjectival uses), to "the generalized article stage (V), going through the stages of "presentative marker" (actually a deictic usage) (stage II), of "specific" (stage III) and, subsequently, "non-specific indefinite marker" (stage IV). By the end of the 'Common Romanian' period, the indefinite marker had reached the fifth stage. Thus, its functions were already well-established when Eastern Romance split into the four major idioms.

In his study of the grammaticalization patterns of the Romanian articles, Nicolae (2009) considers Romanian to be a non-fully fledged stage V indefinite article language. Stage V corresponds to the generalized capacity of *un-* to occur with plural nouns. However, Romanian displays an incomplete plural paradigm, with *unor* as a Gen/Dat indefinite article (morphologically marked by *–or*), but with *nişte* (roughly 'some') for Nom/Acc.[16]

The IR paradigm actually displays both *ur-* and *nušcar-* for plural. (As a note, the suffixal oblique case morpheme has been well-preserved only in the northern variety, Žejanski.) The presence of the plural indefinite article *uri/-e* indicates that, contrary to present-day Romanian, IR is a fully-fledged stage V language, perhaps on a par with Spanish (Heine & Kuteva 2006: 105):

---

[16] According to Nicolae (2009), the indefinite article *unor* has been employed for Gen/Dat because it was the only way to indicate case relations when the noun phrase has indefinite reference (differently from the other Romance languages, oblique case is marked morphologically in Romanian).

(20)   **un** *dia ven-iam* **un- o- s** *hombres [...]*

one day come-.3Pl.Pret.Imperf one. M. Pl men

'One day there came some men'

Case	Singular		Plural	
	Masculine	*Feminine*	*Masculine*	*Feminine*
Nom/Acc	(ur/ân) om	(o) fẹt*a/-ẹ*[17]	(uri/nušcarl'i) omir	(ure/nušcarl'i) fẹte
Gen/Dat	(urvẹ) om	(url'ẹ) fẹt*e*[18]	(urorẹ/nušcarorẹ) omir / fẹte	

Table 1: The indefinite article

3.2.1.2. The definite article: evolution, morphology and referential properties

Like Romanian (and other dialects of Romanian) but unlike other Romance languages, IR has enclitic definite articles with dedicated forms for number (singular and plural) and gender (masculine, feminine):

(21)   *Blagoslovit fečor**u** ân utrob**a** a tå, Isusu.*

blessed son-the.M in womb-the.F of yours.F Isusu

'Blessed is the fruit of thy womb, Jesus."

(Sârbu 2009, text collected in Žejane in 2000)

The table below gives the respective forms of the enclitic article. As in Romanian, the plural Gen/Dat form is identical for masculine and feminine.[19]

---

[17] -*a* is used in Zejane, whereas *ẹ* is the Nom/Acc feminine singular ending (nouns, adjectives) in the southern varieties of IR.

[18] This alternation is similar to what we find in Romanian. It is due to the Latin morpheme -*ae* for feminine genitive singular.

[19] As already mentioned, oblique case marking has been well-preserved only in the northern variety, Žejanski. It has been almost completely lost, or is rarely encountered in the southern varieties. This empirical fact has been pointed out by Kovačec (1971), and my data confirmed it.

Case	Singular		Plural	
	Masculine	*Feminine*	*Masculine*	*Feminine*
Nom/Acc	(*-u, -le*)	(*-a*)	(*-i, -l'i*)	(*-e*)-*le*
Gen/Dat	(*-u-*)-*lui*	(*-e*)-*l'ei*	(*-i-*)-*lor*	(*-e*)-*lor*

Table 2: The definite article

These forms are all derived from Latin and are similar to the corresponding forms in all other Romanian dialects.

Let us focus on the hypothesis of the emergence from Latin of the definite article in Eastern Romance. This is a particularly interesting phenomenon, given that in these Romance varieties the nominal declension with the definite article is different from the other Romance languages.

It is well-established that historically, definite articles have developed from the demonstrative attributes (cf., Greenberg 1978, Diessel 1999, among many others). Diessel (1999:113, 128-9) proposed the following three stages in the grammaticalization evolution of the definite article:

(22)   exophoric demonstrative > anaphoric demonstrative > definite article

As for Romance languages, Vincent (1997) argues for the existence of one more stage between the anaphoric demonstrative and the definite article. This stage corresponds to Late Latin, when *ille* and *ipse* mark definiteness in the nominal phrase, alongside with their cataphoric function (*ille* used as a definite first mention followed by a restrictive relative clause), and anaphoric function respectively (*ipse*) (Vincent 1997:154-6).

Heine & Kuteva (2006: 103) (following Hawkins 2004, Ch. 4), present an even more fine grained view of grammaticalization process from the demonstrative to the definite article from a semantic/pragmatic perspective, articulated in five stages:

(23)   Stage I: the deictic restriction (near hearer/speaker vs. far from hearer/speaker) is
            abandoned; the item can identify the referent relative to some whole pragmatic

set; this stage is restricted to anaphoric reference to objects existing in the immediate situation of utterance.

Stage II : broader range of contexts, expanding from the visible (i.e. *mind the step*) to non visible and larger situations (i.e. *the king has abdicated*);

Stage III: the article can be used universally or generically (i.e. *the lion is a mammal*);

Stage IV: the definite article can be used to also encode specific indefinite reference in addition to definiteness – it can thus simply assert existence;

Stage V: the definite article is used for purely syntactic purposes, such as agreement, noun or class marker (in those languages which have noun classes, i.e. Bantu).

According to Heine & Kuteva (2006:104), no European language has reached the fourth stage, except for possibly a dialect of Swedish spoken in north-eastern Scandinavia. The Romance languages have reached the third stage. Nicolae (2009) suggests that Romanian grammaticalization process of the demonstrative to the definite article may have reached the Stage V. He claims that the combination of the definite article with nominal constituents is part of morphology, rather than syntax.

Roberts & Roussou (2003: 131-136), building on Giusti (1998) and Lyons (1999), present the grammaticalization of *ille* in Romance, which involves three different processes: firstly, a morphophonological reduction (*ille > le*); secondly, semantic bleaching (loss of the demonstrative property); and thirdly categorial change (demonstrative > article).[20]

Slavonic-Romanian documents (dated at the beginning of the 14th century) display occurrences of the enclitic definite article fully functional and with a complete paradigm (cf. Cornilescu 2009).[21] Given that none of the south-Danubian dialects has proclitic definite articles, we cannot but assume that the grammaticalization process had been concluded before Aromanian, Megleno-Romanian and IR detached themselves from the common linguistic trunck (irrespectively of the debate as to when exactly that happened).

---

[20] Cornilescu (2009) shows that Old Romanian had low articles, i.e. appearing on the lower nominal constituent in the DP, which is taken as evidence that it has developed out of a postposed demonstrative (cf. Renzi 1993, among others).

[21] Except for proclitic *lui*, the oblique definite marker for masculine singular, preserved until today in Modern Romanian. It is however a controversial matter as to how it grammaticalized.

Coming to the IR definite article paradigm, let us consider in some depth the two possible forms for Nom/Acc masculine plural. Kovačec (1962:78.ff10) notes that –*i* is derived from Latin demonstrative *illi* through the intermediary form –*l'i*. In fact, in the Southern variety of IR instances of –*l'i* have been recorded: *tare* (Sg.) -*tarl'i* (Pl.) "strong, hard", *bur* (Sg.) - *burl'i* (Pl.) "good", and in both varieties the definite plural form of *cåre* (also *cårele*, see below) is *cårl'i* (which.Pl, Ital. "i quali"). Actually, –*l'i* has been almost completely lost and the reason is due to the existence of plural indefinite forms of certain (truncated) nouns belonging to the basic vocabulary which already end in –*l'i* (*cå* - *cål'i* "horse, horses", *fil'* - *fil'i* "son, sons") and this should have produced forms like **cål'il'i*, **fil'il'i*. We therefore frequently find in the texts nominal elements which are ambiguous between a definite or a bare reading (*cål'i* "(the) horses", *fil'i* "(the) sons") for plurals, too.

We must note, however, that the above arguments from Kovačec (1962:78.ff10) are not very strong. Theoretically, nothing prevents such forms as **cål'il'i, *fil'il'i* to exist. As an illustration, take Romanian lexical entry *lalea* "tulip" – *laleaua* "the tulip" – *lalelele* "the tulips". We have no other comment on the matter at this stage.

One further phenomenon worth noticing is the occurrence of the masculine singular adjectival forms ending in *-(i)le*, that is the enclitic definite article which in Romanian and IR attaches to masculine singular nouns (ending in –*e*) (cf. 24). In order to avoid forms overlapping with adverbs, IR has come to adopt the bound morpheme –*(i)le* for masculine singular adjectives (25) without it assigning definite reference.

(24) a. *câine**le*** (Rom.)
          dog-the.M.Sg

      b. *Câre**le** mi*   *-l åm*   *la ńive.*   (IR)
          dog-the.M.SG to.me   -it have.1SG at fields
          'I keep the dog at the fields.'

                                        (ex. from Sârbu 2009)

(25) Adj. *rumunski**le** brod*   vs   Adv. *rumunski cuvintå*
           a/the Romanian ship   Romanian (to) speak

Kovačec (1962) claims that this use of the morpheme –*le* can be explained by the need for IR to strengthen the difference between adjectival forms (Nom, masc, sg.) and (manner)

adverbial forms, which in most cases would be identical (i.e. Rom.: Un copil *frumos* care cântă *frumos*, IR: Ur *mušåt* fečor cåre *mušåt* cântå "A beautiful child who sings beautifully").[22] When present in the language,[23] the neuter singular adjectival forms are also employed as manner adverbs. At the same time, for those adjectives which lack a neuter form in *–o*, the masculine singular form is thus also employed for neuter (26.b). Such elements would thus become morphologically indistinguishable between, on the one hand, Masc. Sg. Adjective, and on the other hand, N. Sg. Adjective which would in any case be identical to the Adverb. Therefore the need to attach a nominal (in the sense of [+N]) morpheme to adjectives, namely *–(i)le.*

(26) a.   Adj.   Masc.Sg.   oštr*u*
            Adj.   F.Sg.   oštr*a*
            Adj.   N.Sg.   / Adv.   oštr*o*   'hard (on somebody)'
     b.   Adj.   M./N. Sg.   gåbir
            F. Sg.   gåbira   'yellow'

Moreover, the *–(i)le* suffixation has been generalized to other noun modifers and determiners (i.e. relative and quantifier adjectives), indicating that this process was extended to such items which can function *a)* either as adjectives or adverbs (cf. 26.a), *b)* either as adjectives or pronouns (cf. 27, 28). In (27), the first form is pronominal or adverbial, while the second is adjectival.

---

[22] There are a few exceptions: the Masc. Sg. adjectival form may differ from the manner adverb, i.e. adj. *bur*, adv. *bire* (cf. Rom. adj. *bun*, adv. *bine*).

[23] Adjectives borrowed under contact from the dialect of Croatian spoken in the area (Čakavian), or new borrowings from Croatian always display all three genders. As for lexemes originating from Latin, some have derived neuter forms along the Croatian model (as in (i.b)). Sometimes Masc/Fem and Masc/Fem/Neuter forms coexist in the language (probably subject to speaker variation).

   i.   a. *nervoz, -a*
           b. *nervozan, -zna,- zno*   'nervous, irritated, angry'   (Sârbu&Fraţilă 1998)
           (cf. Rom. *nervos, -oasă*, Ital. *nervoso, -*a, Cr. *nervozan, zna, zno*)

(27)　　cât / cât*ile*　　　　'how much, how many'

　　　　câtvå / câtv*ile*　　　'(a) little, (a) few'

　　　　câtgod / cât*ile*god　'any' (quantity)

　　　　nušcât / nušcât*ile*　'some' (quantity)

　　　　acâta / acât*ile*　　　'this much'

　　　　cum / cum*ile*　　　　'how' (manner) / 'such'

　　　　cåre / cåre*le*　　　　'who, which'

　　　　tot / tot*ile*　　　　　'all'

　　　　saki / saki*le*　　　　'each'

(28)　a.　*Tot åm　　poidit.*

　　　　　all have.1Sg eaten

　　　　　'I have eaten everything'.

　　　b.　*Preste　totile ånu　lucru　　-n oštarie.*

　　　　　during　all year-the　work.1SG　in restaurant

　　　　　　　　　　　　　　　　　　　(ex. from Sârbu & Frăţilă 1998)

As for the oblique cases, they are also inherited from Latin demonstratives as given below:

(29)　a.　M.Sg.　　　Gen/Dat.　　Lat.　illuius/illui < *-lui*　　(cf. Rom. *-lui*)

　　　b.　F.Sg.　　　Gen/Dat.　　Lat.　illaeius/illeius < *-l'ei*　(cf. Rom. *-ei*)

　　　c.　M./F.Pl.　Gen/Dat.　　Lat.　illorum [24] < *-lor*　　(cf. Rom. *–lor*)

The first part of this section was a survey of the morphological properties of the definite article in IR. In what follows we will start considering the syntax and semantics of the definite morphemes, i.e. their referential properties.

In his introduction to the 1992 collection of Istro-Romanian texts, Sârbu very briefly notes that "certain subsystems or structures of the language have developed under the influence of a foreign pattern", resulting in phenomena such as "the dropping of the articulated forms of preposed adjectives" (Sârbu 1992:40). If confirmed by the data, this fact would be an

---

[24] The Latin feminine plural form "illarum" was lost.

61

important piece of evidence in favour of a contact approach [25] to IR complex nominal expressions and definite reference. We will see, however, that although the general observation is correct, there are pieces of data (both in the texts and collected by myself) that do not confirm it. Moreover, we will see that IR has developed a different structure for definite nominal expressions when modified by preposed adjectives.

In order to check whether this observation is correct, we first have to: *a)* check whether adjectives can occur both pre- and postnominally and under what conditions, and *b)* to clarify how referentiality is mapped morpho-syntactically in IR in the absence of modifiers. It is indeed licit to suppose that Croatian (which lacks definite articles) may have influenced IR to a certain extent (as Sârbu 1992 notes). We will start with the second observation, and afterwards (Section 3.2.1.2) we dwell in detail on the interaction between definite markers and adjectival modification.

Let us first note that in IR the presence of the definite article does not (necessarily) indicate definite reference. Firstly, we have seen in the previous section that the presence of *-(i)le* can be misleading because it functions as a masculine singular adjectival marker. Secondly, under certain conditions, [26] the article is visible even when the interpretation of the nominal expression is neither definite nor specific. Informants give an indefinite interpretation to the noun 'regal*u*' in (30.a), which is intended to be a habitual event (people always bring gifts on people's birthdays); while in sentence (30b) which has a specific temporal reference, the indefinite is realized by *ur*:

(30) a. *Vaika-mi ducu regalu.*

      always to.me bring3Pl gift-the

      'They always bring me a gift' (context: every year for my birthday).

    b. *Livjia mnji-a dat ur regal.*

    Livija to.me has given a gift

    'Livjia gave me a gift'.

---

[25] Cf. also Aljović (2002, 2005) for an analysis in terms of 'specificity' for short and long adjectival forms in Serbo-Croatian.

[26] Which exactly are these conditions is not easy to identify. More research is needed to clarify this point, which I must leave aside here.

Generally, however, the definite reference is indeed obtained as expected:

(31)   *Verít-av   nevęstele ši   noi ganęim că iǎle.*   Subject
       come have wives-the and we speak   like them
       'The wives came and we speak like them'.   (ex. from Sârbu 2009)

(32)   *Ier   Lara n-å   poidit paninu.* [27]   Object
       yesterday   Lara not has eaten sandwich-the
       'Yesterday Lara didn't eat the sandwich'

According to Longobardi (1994), the syntax-semantics mapping of definite article (in D°) is realized as in (33). In (34) we give examples of how Longobardi's proposals may account for IR Nominal Expressions:

(33)   a. [D [NP x]   definite descriptions
       b. [[D N] [N̶]]   proper names
       c. [[D expletive] [N]] generic Noun Phrases

(34)   a. *Fečori   ier   čuda   čiteit-av.*
          boys-the yesterday much   read have.3.Pl
          'The boys read a lot yesterday.'

       b. *Redento ier   čuda   čiteit-a.*
          Redento yesterday   much   read has
          'Redento read a lot yesterday'.

       c. *Căprele   mârâncu iǎrba.*
          Goats-the eat.3Pl grass
          'Goats eat grass'.

---

[27] The definite reference of the object noun in this example (collected by myself) has been established through providing the appropriate context to the informant.

In languages that have enclitic definite articles, the structure in (33.a) predicts that common nouns to raise to D, where the article is inserted. This is due to the enclitic nature of the article, that morphologically and phonologically cliticizes on the noun. Noun raising to functional projections above NP has been assumed by many authors (cf. Cinque 1994, Longobardi 1994; for Romanian, cf. Dobrovie-Sorin 1987, Grosu 1987, Giusti 1992, among others). In fact, ever since Abney (1987), it is standardly assumed that nouns not only project Noun Phrases, but also Determiner Phrases headed by a functional head D°. More recent literature (Szabolczi 1994, Giusti 1993, 1997, etc.) showed that only articles are heads of DP. Other determiners (including demonstratives) are maximal projections occupying the specifier of DP position. The empirical observation that definite articles and Demonstratives can co-occur in some languages (i.e. Romanian, Scandinavian languages, etc.) is explained by the different syntactic status of these two categories.

3.2.1.3. The 'pseudo-article'

As for the morpho-syntactic mapping of definite reference on modified nominal expressions, let us first note the contrast in (34):

(34)   a. *Čǎste scu lęmnele   uscǎte.*

      these are wood-the dry

      'This is (the) dry wood.'

    b. *Čǎste scu uscǎte(*le) lęmne(*le).*

      these are   dry   wood

      'This is (the) dry wood.'

The examples above form a minimal pair. The parenthesis in the English translation stand to indicate that the interpretation can be either definite or indefinite (cf. "regal*u*" in (30) above). Importantly, informants reject the placement of the definite article –*le* on the prenominal adjective, resulting in a bare nominal expression. The piece of data in (34.a) mirrors

Romanian. As for (34.b), differently from Romanian, IR participial adjectives are also allowed in prenominal position. [28]

Apart from participials, adjectives are not able to function as hosts for the affixal article when pronominal (35.a), while in Romanian they must bear the suffixal article (35.b) when they occur as the first element in the nominal expression. Thus (although relatively rare and dispreferred by speakers) the article in IR is found on the noun:

(35)   a. IR   [?(?)] *mic fečoru*   /*micu   *fečor*   **Adj > N+art**

      b. Rom.   *mic feciorul   / *micul*   *fecior*   **Adj+art > N**

                little boy-the   little-the boy

What we find instead is that IR resorts to an unbound morpheme which resembles the Romanian "pseudo-article" *cel* (traditionally "adjectival demonstrative article" (henceforth also dem.art.)). The use of this "pseudo-article" is much more widely spread than in Romanian, precisely because it is often the only means for rendering a complex nominal expression definite (apart from extralinguistic context), as will be shown below. In what follows, I will call the IR *čela* "demonstrative article" (dem.art.) on a par with Romanian *cel*, and gloss it by "the" (differently from the enclitic definite article, glossed as "-the").

	Singular		**Plural**	
	Masculine/N	*Feminine*	*Masculine/N*	*Feminine*
Romanian	*cel*	*cea*	*cei*	*cele*
Istro-Romanian	*čela*	*čå*	*čel'i*	*čåle*

The parallelism with Romanian can be noted in the examples below:

---

[28] This distribution does not lead us to draw any conclusion as to the merging position of reduced relative clauses in IR (cf. Cinque 2010, cf. also Alexiadou and Wilder 1998 for the two different merging positions of adjectives). Cinque (2010) notes that the postnominal position of reduced relative clauses in Romanian is common to the other Romance languages, while it is not found neighbouring Balkan languages such as Greek or Bulgarian. It will be interesting to see if IR behaves like a Romance or Balkan language in this respect, issue which I must leave unexplored for the moment due to scarcity of data at my hand.

(36)   a. fråte**le čela** mic   **N+art.>dem.art.+Adj**   IR

brother-*the čela* young

b. frate**le cel** mic   **N+art.>dem.art.+Adj**   Rom.

brother-*the cel* young

(37)   a. **čela** mic**u**   **dem.art + Adj+art.**   IR

čela young-*the*

b. **cel** mic   **dem.art + Adj**   Rom.

cel young

'the young one'

Although in standard Romanian there is no definite article on the adjective when the noun is elided, we do find cases in colloquial, family, or regional registers in which the IR example in (37.a) finds its perfect match:

(38)   *Ăla micu a dormit toată noaptea.*   (Rom.)

that small.the slept all night long

As suggested by the gloss above, it is likely that it is the demonstrative that appears in the Romanian construction. As for IR (37.a) above, it is not clear whether it is the demonstrative or the 'demonstrative article', although intuitively informants exclude the 'distal' interpretation. We leave this question open for the time being.

As already mentioned, reasearch has shown that crosslinguistically articles seem to originate as grammaticalized demonstrative (specifically, cf. Giusti 1995 for Germanic, Giusti 2001 for Romance, Dimitrova-Vulchanova & Vulchanov 2008 for Bulgarian). According to Coene (2004), *cel* also derives from the Latin demonstrative of distance *ille* preceded by the demonstrative adverb *ecce* (*ecce-illum, ecce-illam*). Thus, *cel/cea* can be morphologically decomposed into *ce*-D.

In the literature on Romanian, *cel* was initially analysed as the equivalent of the French *celui* (Dobrovie-Sorin 1987, cf. also Cornilescu 1992 for *cel* as an 'expletive determiner', and Grosu 1994). In more recent literature *cel* was demonstrated not to be a head, but an XP (Campos 2005, Giusti 2002, cf. also Ungureanu 2009 for a proposal of a *cel*P). Just like the

Demonstrative *acel*, *cel* may have a deictic/anaphoric interpretation, differently from the enclitic article, which may lack independent reference (Giusti 1993).

Cinque (2010) notes that the syntactic and semantic properties of adjectives introduced by *cel* in Romanian (previously described in the literature, cf. Lombard 1974 Ch.6, Dobrovie-Sorin 1987, Giusti 1993, Grosu 1994, Coene 1999, Cornilescu 1992, Drăgan 2002), among which the fact that they can only follow the noun (39), indicate that these are adjectives of indirect modification which have a reduced relative clause source.

(39)   *băiatul   cel curajos   /*cel curajos băiat*
       boy-the   the courageous   / the courageous boy
       'the courageous boy'

Should IR *čela* also behave like a 'doubling' article in certain structures as distribution facts (see below) incline us to think, then there is reason to believe that although it is completely homophonous with the Demonstrative adjective, it has also come to play another function in the language.

Let us now consider the distribution of what we shall call 'demonstrative article'. In (40.a) the definite article appears on the subject. In (40.b) we see that an adjective is inserted and the definite article still attaches to the noun which stays in its base position (cf. 40.a). However, this is highly marked construction. Speakers would marginally accept it and they would suggest (and productively use) two other constructions given in (41.a) and (41.b) below.

(40)   *Io-m (ur) mic fečor.*
       I have (a) little boy
               a. *Fečoru   -i ân cåsa.*
                  boy-the   is in house
               b. **Micu fečor   /*$^{?(?)}$*Mic fečoru /   -i ân cåsa.*[29]
                  little-the boy /   little boy-the   is in house

---

[29] Postnominal adjectives are very marked and context-dependent, see Section 3.5. Note that for instance *porcu mic* 'pig-the little' is a newborn pig, or *brecu mic* refers to a puppy, so they are similar to compounds.

(41) a. *čela micu   fečor /*^{??}*čela mic   fečor*   **čela   Adj+art   N**

    the little-the boy   /   the little boy

b. *fečoru čela mic*   / **fečor čela mic*   **N+art *čela*   Adj**

c.**micu čela fečor*   / **mic čela fečor*

The data shows that when the noun is modified by an adjective, *čela*-insertion becomes obligatory in order to obtain a definite interpretation. Interestingly, we note that the adjective bears a morpheme which appears to be the definite article itself in (41.a), if it precedes the noun. The following question arises: why are *micu fečor* 'little-**the** boy' and *mic fečoru* 'little boy-**the**' ungrammatical/degraded, and thus it is obligatory to insert the demonstrative article in co-occurrence with the definite article? We will try to give an answer to this question in the following sections.

    Let us further note that when the context suggests that two modified nouns are contrasted, we (marginally) find a N+art. preceding the adjective, which may or may not be introduced by *čela*, (42). If the noun is elided, the definite article is 'transferred' onto the adjective, and *čela* is inserted.

(42)   *Io-m   doi fečori, ur   mic   ši   ur måre…*

    I   have two boys, one little and one big

                Fečoru (čela)   mic / čela micu   ei ân câsa,   **N+art (*čela*) Adj**

        boy-the (the) little   / the little-the is in house,

                fečoru (čela) mårle / čela mårle   mes -a   bę   åpa.

        boy-the (the)big   / the big-the   gone -has drink water

'I have two boys, a little one and a big one. The boy that is little / the little one is in the house, the boy that is big / the big one went to drink some water'.

    Consider other types of noun modification and the interaction with *čela*. In (43.a) below, *čela* is excluded. [30] The possessive adjective is a maximal projection in SpecDP. However, when an adjective is inserted, the presence of *čela* becomes obligatory, as in (43.b).

---

[30] It is, of course, grammatical if *čela* is a demonstrative.

(43) a. *(*Čela) mę porc ei mic.*

      the   my pig is little

b. ^{??}*(Čela) mę åb porc ei mic.*

      the   my white pig is little

Up to now we assumed that *čela /čå / čel'i /čåle* are 'demonstrative articles'. The parallelism with Romanian 'demonstrative articles' can be noted in the examples in (44),[31] while the main difference is pointed out contrastively in (45).

(44)   a. *fečoru čela mic*   **N+art.** *čela* **Adj**   IR

       b. *băiatul cel mic*   **N+art.** **cel** **Adj**   Rom.

         boy-the the young

---

[31] Both Romanian *cel* and IR *čela* are also used for the formation of the superlative together with the morpheme *måi*, and in IR also for the comparative (although exceptionally, since intonation/stress suffices to distinguish between the comparative and the superlative). Notice that *čela måi tireru fråte* 'the younger son' displays the doubling phenomenon, similarly to (36.a)/(40).

    Furthermore, although in standard Romanian there is no definite article on the adjective when the noun is elided (i.b), we do find cases in colloquial, family, or regional registers in which the IR example in (i.a) finds its correspondent (ii). As suggested by the gloss, it is the Dem that appears in the Romanian construction. However, *ăsta/ăla* is not syntactically identical to the Dem *acest/acel* (iii), and it can, at least for some speakers including myself, behave like *cel* (iv).

     i.   a. *čela micu*   (IR)   b. *cel mic*   (Rom.)

           the young-the   the young

           'the young one'

     ii.   *Ăla/Ăsta micu a dormit toată noaptea.*

          that/this little-the has slept all   night

     iii.   a. *acel / * ăla băiat*   **≡Dem.**

           b. *băiatul acela / ăla*   'that child'

     iv.   a. *copilul ăl mic*   **≡Dem.art.**

           'the little child'

           b. *cel / ăl / *acel mai frumos (băiat)*

           'the most handsome boy'

(45)   a. *čela micu   fečor*   **čela Adj+art   N**   IR
          b. **cel mic(u) băiat*   (cf. (44.b))   Rom.

As for IR (44.a) above, we have not yet said anything with respect to what motivates the claim that it is a 'demonstrative article' and not a Demonstrative. This will be done in the following section.

3.2.1.2.1. The status of *čela /čå / čel'i /čåle*
The important question whether *čela /čå / čel'i /čåle* might actually be the Demonstratives 'that/those' with which they are homophonous, c n finally be answered by pinpointing the environments in which *česta/ čåsta/ češti/ čåste* 'this/these' are banned.

    As in Croatian, in IR the Demonstrative is always prenominal. [32] It does not co-occur with the definite article (46.a), unless the noun denotes a kinship relation (46.b). Of all noun modifiers, it surfaces as the leftmost one (47).

(46)   a. *česta fečor(*u)*
             this   boy
          b. *česta fråtele   / susédu*
             this   brother-the   / neighbour-the
(47)   *Češti mel'   trei   musaţ   fraţi   čuda lucru.*
          these my.Pl three handsome brothers much work.3.Pl
          'These three handsome brothers of mine work a lot'.

          Dem > Poss > Num > Adj > N

---

[32] In my quantitative study, 89% (64/72) Demonstratives in Sârbu (2009) surfaced prenominally. Of the postnominal ones (8/72), only one was produced by a speaker from the Southern villages. In this paper we are dealing with the Southern variety of IR. Apparently in the Northern variety (*Žejanski*), the Demonstrative can follow the noun and the definite article is obligatorily inserted. This fact mirrors Romanian data (ii).
          i. *fečor*(u) česta/čela*
          ii. *băiatul acesta/acela*   'boy-the this/that'   (M. Doričić, Trieste, 2009)

What we have already noticed instead is that there are occurrences of postnominal *čela* 'that' (when the nominal expression is modified), but there are no such occurrences of *česta* 'this'.

(48) a. *fečoru   čela /*česta mic*

        boy-the **čela** / this little

It logically follows that any occurence of *čela* in postnominal position cannot be a Demonstrative, therefore it must be the 'demonstrative article'. As for the prenominal postion of *čela*, there are two possibilities: a) it is always a Dem, or b) it is either a Dem. or a 'dem. art.', function of the configuration in which it appears. The relevant data are given in (49):

(49) a. *česta mic(*u)   fečor(*u)*

        this little(*-the) boy(*-the)

      b. *čela mic fečor*

        that little boy

      c. *čela micu fečor*

        the little-the boy

(49.a) shows that the Demonstrative does not co-occur with the definite article, be it on the noun or on the adjective. In a parallel fashion, in the absence of the article (49.b), *čela* is the Demonstrative. Conversely, the presence of the article (49.c) indicates that *čela* must be doubling 'dem. art.'. Note that in this configuration the definite article *must* appear on the pronominal adjective.

Finally, (50) strenghtens the hypothesis that the presence of the article, in this case on the noun, a marked construction with a postnominal adjective, is indeed a test for identifying *čela* as a 'dem. art.'.

(50) $^?$*čela   /*česta fečoru mic* [33]

---

[33]As in modern Greek, the data at my hand seem to indicate that only adjectives that can be used predicatively allow for the dem.art. and def.art. to co-occur (cf. Alexiadou & Wilder 1998, Lekakou & Szendrői 2007 for restrictivity as a criterion).

In any case, the IR construction does not appear to be parallel to Greek demonstrative doubling (also known as

the/*that   / this   boy-the little

Summing up**,** the definite article cannot attach to a prenominal adjective. The main strategy is to insert *čela*, an element that is homophonous with the distal Demonstrative. IR Demonstratives are always prenominal and their co-occurrence of the definite article is banned. Thus, when *čela* is postnominal it functions as a 'demonstrative article' (doubling the definite article showing up on the noun), and when *čela* is prenominal: a) it is the Demonstrative when no definite article is present, and b) it is the 'dem.art.' when it co-occurs with the article. The same conditions are true when the noun is elided. The prediction is that all postnominal occurrences of *čela* should also display the definite article.

3.2.1.2.2. **Possible analysis and open questions**

On the basis of Spanish evidence, Brugè (1996, 2002) takes Demonstratives to originate lowest in the functional projections of the noun which host its modifiers. The low position of the locative reinforcer *oanča* signals the original position of the Demonstrative:

(51)   *česta mušat   fečor oanča*
       this   handsome boy   here

   Like adjectives, Demonstratives agree in gender and number with the noun. Phi-features checking is achieved through Concord, which is a relation between the Specifier and the head of the functional projection in which the Specifier is merged (cf. Giusti 2008). This operation does not trigger movement by itself, only AGREE does, covertly or overtly, in order to obtain a Spec-Head configuration. The Demonstrative surfaces on the left of all other adjectives because it eventually lands in SpecDP for checking the relevant referentiality features.

   Data indicate that the structures in which the Demonstrative is present ban the definite

---

Greek 'polydefinite constructions' or 'determiner spreading) (cf. Horrocks&Stavrou 1987, Kolliakou 1995 and following work, Grohmann & Panagiotidis 2004, Panagiotidis 2005, Campos & Stavrou 2005, Lekakou & Szendrői 2007, among many others). For instance, the doubling dem.art. cannot be reiterated. Compare Greek and IR in (i) and (ii). respectively. These issues, however, would need a separate study.

   (i) to megalo to petrino to spiti   (ii) *čela mic čela brižen fråte
       the big the stone the house   *the young the wretched brother*

article. This could be explained by assuming a Doubly Filled DP filter (Dimitrova-Vulchanova and Giusti 1998) The Demonstrative triggers reference and deictic/anaphoric interpretation, so once it moves to SpecDP it satisfies the requirements of the DP therefore the definite article need not (must not, in the minimalist account) be filled. However, the same should hold for the "doubling" construction with prenominal *čela* (49) (given that *čela* may also have deictic/anaphoric interpretation and independent reference).

Remember that apparently the only condition for the "doubling" to be realized is for an adjectival modifier to be present as was shown in (40), repeated below for convenience:

(40)/(51)   a. *čela mic**u** fečor*
            the little-the boy
         b. *fečor**u**   čela mic*
            boy-the the little

It has been argued that the nominal structure (DP) is parallel to the clausal structure (CP) (Carstens 2000, Adger 2003, Giusti 2006 a.o.). Semantic and thematic relations are established in the VP/NP-shell (cf. Larson 1988 for VP-shell), agreement relations and other morpho-syntactic requirements are realized in the intermediate layer, while the highest CP/DP layer interacts with the structure/element that selects it, at the same time establishing propositional value or referential interpretation, respectively.

(52)   a.   [$_{\text{Clause}}$ CP-layer [IP-layer [VP-shell]]]

         b.   [$_{\text{NE}}$ DP-layer [FP-layer [NP-shell]]]

Grohmann & Panagiotidis (2004) express the same concept in a framework involving Prolific Domains (corresponding to the three layers in (52) above), Copy Spell Out and Anti-Locality effects, building on an Anti-Locality Hypothesis:

(53)   *The Anti-Locality Hypothesis* (Grohmann 2003:26)
       Movement must not be too local.

Let us first consider the ban on the co-occurrence of the Demonstrative with the definite article. We assume with Brugè (1996, 2002) that the Dem originates lowest in the hierarchy of the modifiers. This assumption crucially derives that the Anti-Locality Hypothesis is non violated, namely that movement of the Demonstrative is not too local (within the same layer/domain), so the spelling out of a possibly violating copy will as a minimal copy specified for number, gender and Case (namely the definite article) is not necessary (54.a). [34]

(54) a. [$_{DP}$ *česta*   [$_{FP1}$ mic [$_{FP2}$ *~~česta~~* [$_{NP}$ fečor]]]

  b. [$_{DP}$ *čela* [$_{FP1max}$~~*čela*~~ mic-*u* [$_{FP1}$ ~~mic~~ [$_{NP}$ fečor]]]

  c. [$_{DP}$ fečor-*u* *čela* [$_{FP1max}$ *čela* ~~-u~~ [$_{FP1}$ mic   [$_{NP}$ ~~fečor~~]]]

A further assumption is that *čela* (on a par with Romanian *cel*) introduces indirect modification adjectives (cf. Cinque 2010), which are merged in the structure of DPs as predicates of reduced relative clauses (the FP$_{max}$s in (54.b-c)), above the FPs hosting adjectives of direct modification. Movement from the highest specifier of the intermediate layer to the highest layer is anti-local,[35] resulting in the spelling out of a mimal copy, the definite article. Given the affixal nature of the definite article in IR, it needs a host. Thus either the noun or the adjective could, in principle, raise. The most minimal operation is for the closest possible element to raise, namely the adjective (44.b). The alternative is also found, with the extra-move of the Noun+article to the DP.

However, while this solution neatly captures why the Demonstrative and the definite article cannot co-occur in IR, it does not seem to explain the "particular morpho-phonological effect" mentioned by Grohmann & Panagiotidis (2004) themselves for Romanian (namely that the article skips a projection to land on the noun in the equivalent of (54.c)). This questions

---

[34] This analysis would also derive the Romanian Dem>N as in (54.a), though the obligatory insertion of the article when the Dem is postnominal (N+*art.*>Dem) would be problem for the assumption that the movement of the Dem does not violate (Anti-)Locality. One way out is to say that the Dem in Romanian is merged much higher (Romanian is the only Romance language that does not have a Demonstrative re-enforcer, like Spanish *este-aqui*, Italian *questo qui*, etc. to indicate its low base position, cf. Brugé 2002), but it would be a language-specific stipulation. I must leave all these aside for the moment.

[35] This approach is similar in spirit to the Phase theory (Chomsky 2001, 2005), for which only the highest Spec of a phase is "visible" for further derivations (cf., a.o., Abels 2003).

remains open for future research and, possibly, for a better understanding of the data.

3.2.2. The demonstratives (with and without N)

Demonstratives express the near/distal opposition (cf. English 'this/that'). The case morphemes in *italics* below are identical for other determiners too, i.e. *ur* 'a', *nušcarl'e* 'somebody' , *såki* 'any', *ničur* 'no(one)', *cårl'e* 'which', *amindoi* 'both', etc. The Gen/Dat feminine singular morpheme is actually–*l'ę* with the variant *ę* for *čest"ę* 'to this'.

Case	Singular[36]		Plural	
	Masculine	*Feminine*	*Masculine*	*Feminine*
Nom/Acc	česta	čåsta	česti	čåste
Gen/Dat	čest*vę*	čest"*ę*	čest*orę*	

Table 3: The proximal demonstrative

Case	Singular		Plural	
	Masculine	*Feminine*	*Masculine*	*Feminine*
Nom/Acc	čela	čå	čel'i	čåle
Gen/Dat	čel*vę*	čel'*ę*	čel*orę*	

Table 4: The distal demonstrative

Demonstratives are obligatorily prenominal. Apparently they may also occur postnominally in certain structures[37], such as one attested postnominal demonstrative which is further modified by a restrictive relative clause.

(55)   […] *şi čeşti d-atunče poşni cuvintå čåsta nostra limba.*
        … and these since start speak this our language

---

[36] The neuter form has only been attested for the Nom/Acc singular: *česta* ('this') and *čå* ('that') identical to the masculine and feminine singular, respectively.

We must mention again that the oblique synthetic marking is only encountered in the northern variety, Žejanski.

[37] Only 8 out of 72 occurences of demonstratives in Sarbu (2009) manuscript texts were postnominal, some of which however might be the demonstrative article. Out of 8, 7 were attested in Žejane variety.

'And since then, they (would) start speaking this language of ours.'

(56) *Ręş   ziče de diţa   čåsta če jives ân Jeiǎn cum cuvintu mǎi vişe-acmo*

would.1SG say of youth this which live in Jeiǎn how speak.3PL more now

'I would say about these children which live in Zejane that they speak more now'

(examples from Sârbu 2009, collected in Žejane in 1999)

As already mentioned in the above discussion on the "pseudo-article", the Demonstrative does not co-occur with the definite article in the southern variety of IR.

(57) *Česta fečor je me fråte.*

this boy   is my brother

(58) **fečor(u) česta / *česta fečoru*

boy-the this / this boy-the

Apart from the examples found in the texts, my informant from Žejane also allowed the demonstrative to occur postnominally. In this case, the noun obligatorily bears the definite article (like in Rom.):

(59) *Fečor**u** česta je marle.*   **N+def.art>Dem**

boy-the this is big

'This boy is (a) big (boy)'

Like Croatian (and other languages, among which Italian and Spanish) but unlike Romanian, we find in IR the construction Demonstrative Adverbial Reinforcer. The two elements share the 'proximity'/'distal' feature:

(60)   a. *česta oanča*

this here   'this one here'

b. *čela colea*

that there   'that one there'

(61)   a. **česta colea, *čela oanča*

this there, that here

The position of the locative reinforcer signals the base position of the Demonstrative, which originates low in the sequence of projections associated to the nominal expression (cf. Brugè 1996, 2002):

(62)   česta mušat   fečor oanča
        this handsome boy   here

When the noun is also modifed by an adjective, the two elements can appear discontinuously, but the Dem must be prenominal and the adverbial reinforcer must occur last:

(63)   a. {*mladic}   česta (musat)   {mladic}   oanča
           young man this   (handsome) young man here
           'this handsome young man here'
        b. {*mladic}   čela (musat)   {mladic}   colea
           young man that   (handsome) young man there

3.2.3. The possessives (with and without N)

IR possessives pronouns and adjectives inflect with respect to the person of the possessor (1st, 2nd and 3rd), the number both of the possessor and of the possessee, respectively (singular and plural, there is no paucal). Also, they inflect and with respect to the gender of both the possessor and of the possessee, respectively (masculine, feminine and neuter – only for the singular), orientation towards the subject for the 3rd person (reflexive and non reflexive). As for case, only the northern variety inflects (Nom/Acc and Gen/Dat of the possessee). The functional preposition *a* introduces the possessors.

   All these oppositions are also expressed in Romanian.

Person	Case	Singular (possessee)			Plural (possessee)		
		Masc.	*Fem.*	*Neuter*	*Masc.*	*Fem.*	*N.*
1st	Nom/Acc	a mev (me[38])	a må	a mevo	a mel'	a męle	-

---

[38] The forms in the parethesis in this table are the ones used in the southern varieties.

	Case	Masc.	Fem.	N.	Masc.	Fem.	N.
			(mę)				
	Gen/Dat	a melvę	a mel'ę	-	a melorę		-
2ⁿᵈ	Nom/Acc	a tev (te)	a tå (tę)	a tevo	a tel'	a tåle	-
	Gen/Dat	a telvę	a tel'ę	-	a telorę		-
3ʳᵈ (subj. oriented)	Nom/Acc	a sev (se)	a så (sę)	a sevo	a sel'	a såle	-
	Gen/Dat	a selvę	a sel'ę	-	a selorę		-

Table 5: Possessives forms – one possessor

Gender and number (possessor)	Case	Singular (possessee)			Plural (possessee)		
		Masc.	*Fem.*	*N.*	*Masc.*	*Fem.*	*N.*
Masc. Sg.	Nom/Acc	a lui			a lui		-
	Gen/Dat	lu a lui	le a lui	-	(-lor) a lui		-
Fem. Sg.	Nom/Acc	a l'ei			a l'ei		-
	Gen/Dat	lu a l'ei	le a l'ei	-	(-lor) a l'ei		-

Table 6: Possessives forms – one possessor, non subject oriented (non reflexive)

Gender and number (possessor)	Case	Singular/Plural (possessee)		
		Masc.	*Fem.*	*Neuter*
Masc. Pl.	Nom/Acc	a lor		
	Gen/Dat	lu a lor	le a lor	-
Fem. Pl.	Nom/Acc	a lor		-
	Gen/Dat	(-lor) a lor		-

Table 7: Possessives forms – more possessors, non subject oriented (non reflexive)

As for possession relations expressed by nominal elements, we note that the traditionally so-called 'genitival possessive article' which in Romanian varies for gender and number of the possessee (*al, a, ai, ale*) is only present in Žejane IR and is invariable (*a*). The southern varieties do not have synthetic forms for Gen/Dat which are rendered through the functional morpheme *lu* (masc.)*, le* (fem.). The feminine common and proper nouns also have a specific

inflection, different from the definite article. Masculine common nouns are inflected with (what looks like) the definite article, while masculine proper nouns are uninflected.

(64) a. *fiľu lu sore*   */ lu frǻtele*

   son-the of sister   / of brother.the

   'the sister's/brother's son'

   b. *sora lu Lare*   */ lu Dejan*

   sister-the of Lara   / of Dejan

   'Lara's /Dejan's sister'

Possessive noun phrases are generally postnominal (see example above). Possessive adjectives can occur both pre- and postnominally:

(65) *{l'ei / lui} sora {l'ei / lui}*

   'her / his sister'

Hierarchically, possessives precede all nominal modifiers (and follow demonstratives). The order Poss > Adj > N is rigid in IR:

(66) a. *Lui drǻga sor verit-a.*

   his dear sister come-has

   b. **sor(a) lui drǻga*   cf. (Rom.)   *sora lui dragă*

   c. **drǻga lui sor(a)*   cf. (Rom.)   *draga lui soră*

When Dem, Poss and Adj are all present, the only possible orders are:

(67)   a. *Česta lui zelen mižol je pre scǻnd.*   **Dem > Poss > Adj > N**

   this his green glass is on table

   b. *Česta lui mižol zelen ii pre scǻnd.*   **Dem > Poss > N > Adj**

   this his glass green is on table

   c.   *Česta zelen mižol lui je pre scǻnd.*   **Dem > Adj > N > Poss**

   this green glass his is on table

'This green glass of his is on the table.'

The orders in (67.b-c) are derived from (67.a) by movement. In (67.b), the noun moves past the adjective, while in (67.c) the adjective and the noun move together past the possessive modifier.

   We should mention at this point that my informant from Žejane, who allowed a postnominal Dem, also allowed the order **Poss>Dem>N** (along with **Dem>Poss>N** and **Dem>N>Poss):**

(68)   *A nostru česta filj je mårle.*
       of our   this   son is big
       'This son of ours is (a) big (boy)'.

The examples below show that while the definite article is obligatory on the noun when only the possessive is present, it is excluded when there is an adjective:

(69)   a. *{Lui} fråtele {lui} verit-a.*
          {his} brother.the {his} come-has
       b. *Lui dragi fråte verit-a.*
          his dear brother come-has

However, when a Dem is present the article becomes ungrammatical:

(70)   *Česta fråte(*le) lui je mic.*
       this brother.the his is little
       'This brother of his is little'

3.3. The quantifier (with and without N)

This section deals with the quantifier system in IR. Original data will be presented and confronted with recent formal theories, in particular Giusti & Leko 2004 and Cardinaletti & Giusti 2006.

From a lexical viewpoint, IR employs invariable quantifiers borrowed from the Istrian dialects (presumably Čakavian):

(71) *Čuda, više* – invariabile   'much/many/a lot of'   (*čuda/više omir* 'many people')

Omârva – invariabile   '(a) little, (a) few'   (*omârva omir* '(a) few people')

Zălica – invariabile   '(a) little, (a) few'   (*zălica låpte* '(a) little milk')

Saki(le), saka, sako   'each (MASC, FEM, N)'   (*Saca căsa* 'each house')

The Romance lexical entries have been preserved only for comparative and superlative:

(72)   *munt/ (var.) mund*   (Lat. *multus/mult/munt/mund*) 'many'

       puţin   (Lat.** putinus = pittinus*) 'few'

Kovačec (1967) claims that in older stages IR could have used the Romanian-type quantifier system. Present-day IR, displays system which is structurally identical to the Croatian one, although many speakers also accept (though disprefer) the Romanian one. The forms of the quantifiers are, however, in most cases inherited from Latin and parallel to the Romanian ones.

As mentioned before, IR has developed new forms for masculine singular, by suffixation with *–i-le* (e.g. *totile*), in order to avoid form overlapping with adverbial forms. Allegedly, this enclitic morpheme originates from one of the masculine singular definite article morphemes, but it is now a marker for masculine singular.

Recent literature has shown that categorically there are three types of Quantity expressions (Giusti & Leko 2004, Cardinaletti & Giusti 2006)**:**

1) <u>Quantity Nouns</u>: X°s heading their own projection QP; they select the noun phrase as their complement, to which they assign (genitive) Case;

2) <u>Quantity Adjectives</u>: XPs in the Spec of functional projections (hierarchically ordered) modifying the noun, (cf. Cinque 1994, 2010);

3) <u>Quantifiers</u> (proper): X°s heading their own projection QP; select the noun phrase as their complement, to which they can assign (partitive) Case.

Furthermore, Cardinaletti & Giusti (2006) argued for a tripartition of Quantifiers (proper). The structures of the three constructions in English is provided:

a) <u>Existential Qs</u> are 'transitive': they select for a DP (variable) and a partitive PP; they assign 'Partitive case' to the DP.

(73)   [$_{FP}$ [$_{F'}$ Q [$_{AgrQP}$ DP ~~Q~~ [$_{QP}$ PP [$_{Q'}$ ~~Q DP~~]]]]]
       many   books   of those you gave to me

b) <u>Universal Qs</u> only select for a DP (variable=subset); they don't assign case to the DP (because of their semantics: do not have the function to relate a subset to a set).

(74)   [$_{FP}$ [$_{F'}$ Q [$_{AgrQP}$ DP ~~Q~~ [$_{QP}$ ~~DP~~ [$_{Q'}$ ~~Q DP~~]]]]]
       all   (the) books

c) <u>Distributive Qs</u> are 'transitive': they select for a DP (variable) and a partitive PP; assign 'Partitive case' to the DP.

(75)   [$_{FP}$ Op [$_{F'}$ Q [$_{AgrQP}$ DP ~~Q~~ [$_{QP}$ PP [$_{Q'}$ ~~Q DP~~]]]]]
       each   book   of those you gave to me

The property of assigning case was shown by Cardinaletti & Giusti (2006) for Italian (cf. 76 below). The existential Q assigns partitive case, thus the partitive clitic *ne* surfaces when the complement of the quantifier is extraposed. In contrast, it is the accusative pronominal clitic *li* (masculine plural) that is spelled out when extraposed from the complement of an universal quantifier.

(76) a. Ne$_i$ ho visti [$_{QP}$ molti [t$_i$]].   (existential, assigns partitive Case)

        CL I-have seen many

     b. Li$_i$ ho visti [$_{QP}$ tutti [t$_i$]].   (universal, does not assign partitive Case)

        CL I-have seen all

In what follows, we will look in turn at each of the three quantity expressions in IR. Also, we will exemplify the three types of quantifiers proper (existential, universal and distributive) when there is available data.

3.3.1. Quantifier Nouns

In Croatian (cf. Giusti & Leko 2004) and Romanian, QNs assign partitive (i.e. Genitive) case. The verb agrees with the QN (but may also marginally agree with the N in Romanian).

(77) *Većina   mojih prijatelja je došla.*   (Cr.)

     majority-NOM my-GEN friends-GEN arrived.SG

     'The majority of my friends arrived.'   (Giusti&Leko 2004)

(78) a. *Majoritatea dintre noi a/??am sosit.*   (Rom.)

        majority-NOM of (lit. among) us-NOM has.3SG/have.1PL arrived

     b. *Majoritatea prietenilor mei a/??au sosit.*

        majority-NOM friends-the.GEN my.GEN has.3SG/have.1PL arrived

In IR, *večinom* 'majority' is only found in Žejan; while no Quantifier Noun has been identified in the southern variety. This already casts doubts on the existence of such category at all. Furthermore, it does not assign partitive Case, since it occurs with a nominative subject with which the verb agrees. It is therefore reasonable to assume that *večinom* 'majority' is not a QN but a universal quantifier in the sense of Giusti & Leko (2004):

(79) a. *Samo jeiånţi, večinom jeiånţi*

        only jeian.people-NOM, majority-NOM jeian.people-NOM

        '(We are) only from Zejane, most of us are from Zejan'.

b. *Şi večinom   noi   ăm ...   curiera a trås.*

and majority-NOM   we-NOM have-1PL … bus has-3SG drove

'And the majority of us drove buses'.   (ex. from Sârbu 2009)

3.3.2. Quantifiers and Quantity adjectives

Cardinaletti & Giusti (2006) have identified a set of properties that distinguishes the two categories. These properties are listed in the table below:

Property	**Quantifiers (Q)**	**Quantity Adjectives (QAdj)**
1	can precede definite/indefinite articles	are preceded by a determiner (in languages that have definite articles)
2	can co-occur with a pronoun	cannot co-occur with a pronoun
3	allow movement of the DP in Spec,QP	cannot be stranded if the DP moves
4	can occur in discontinuous position	cannot appear in discontinuous position (in languages in which adjectives cannot do so)
5	can select a partitive PP	cannot select a partitive PP

In what follows, we will provide IR data for those properties in the table above that data allowed to infer for universal, existential and distributive quantification, in order to distinguish between quantifiers and quantity adjectives, namely properties 1-4.

Property 1:

-       universal quantifiers: IR quantity expressions precede determiners

(80)   *Tote locurle scu mušate.*   / *locurle tote*

all places-the are beautiful.

'All the places are beautiful.'

(81)　a.　　all (the) places / *the all places

　　　b.　　*toate locuri*(le) / *toatele locuri / locurile toate*　　　　　(Rom.)

-　　existential quantification: in IR, no existential Q is compatible with a definite article.

(82)　a.　　*Čuda rumuni prenč trecave.*

　　　　　many_{invar.} Romanians-Ø here pass

　　　　　'Many Romanians pass by'.　　　　　　　　　　　　(ex. from Sârbu 2009)

　　　b.　　*trei locur / *trei locurle*

　　　　　three places / three places-the

　　　c.　　*vrur loc / *vrur locu*

　　　　　some (any) place / *some (any) place-the

(83)　a.　　(the) many children / *many the children

　　　a'.　　*?mulţi(i) băieţi / *mulţi băieţii / *cei mulţi băieţi*　　　　(Rom)

　　　　　many(-the) boys / *many boys-the / the many boys

　　　b.　　(the) three children / three the children

　　　b'.　　*(cei) trei băieţi*

　　　　　the three boys

　　　　　 (*the) some (*the) children

　　　c'.　　*vreunii (*băieţi)*

　　　　　some(any)-the (*boys)

　　　　　　　　　　　　　　　　　　　　　　　(ex from Giusti 1992)

　　-　　distributive quantification: as in English and Romanian, IR distributive Qs are incompatible with the definite article: [39]

---

[39] Recall that *sakile* is only apparently definite. See again the discussion around (27).

(84) *E noi ăm sakile miseţ acaţăt.*
      and we have each month got
      'And we were paid each month'.

                                                              (ex. from Sârbu 2009)

(85) a. (*the) each/every (*the)child
      b. *fiecare copil / *fiecarele copil / *fiecare copilul*   (Rom.)
         each child / each-the child / each child-the

(86) *N-ă dat zălic núşte mai munt pénzie.*
      CL.1PL have given little some more pension
      'They gave us little higher pensions'.

                                                      (ex. from Sârbu&Frăţilă 1998)

(87) *Podfrigitam cu cumpir şi omârva de ăl' şi ul'ie şi ur lic de popăr.*
      fried have.1SG with potatoes and a little *of* garlic and oil and a little of pepper
      'We fried (it) with potatoes and some garlic and pepper.'

                                                      (ex. from Sârbu&Frăţilă 1998)

(88) *Vidio sam nekoliko mojih priatelja.*   (Cr.)
      (I) saw several my-GEN friends-GEN
       'I saw several of my friends'.   (ex. from Giusti&Leko 2004)

(89) *Au venit amândoi dintre prietenii mei cei vechi.*   (Rom.)
      have arrived both *of* friends-NOM my-GEN the old
      'Both of my old friends arrived.'

From the data above we conclude that the universal quantifier is a Q proper, while this is an irrelevant test for existential and distributive quantifiers, since they are incompatible with a definite article, as discussed by Cardinaltti & Giusti (2006).

Property 2:
  -   universal quantifiers can co-occur with a pronoun

(90) *noi toţ / *toţ noi*

we all / all we

(91) a. they all / all *(of) them

b. *noi toţi / *toţi (dintre) noi*   (Rom.)

we all / all (of) we

- existential quantification: cannot co-occur with a pronoun

(92) *čuda noi / *noi čuda*

many us /us many

(93) a. *many we / *we many -> many *(of) us

b. *mulţi noi / *noi mulţi -> mulţi *(dintre) noi*

many us / us many   many (of) us

- distributive quantifiers cannot co-occur with a pronoun

(94) *sakile noi / *noi sakile*

each us / us each

(95) a. *each we / *we each -> each *(of) us

b. *fiecare noi / *noi fiecare -> fiecare *(dintre) noi*   (Rom.)

This property leads to the same conclusion as for *propery 1*, namely that the universal quantifier is a Q proper, while the existential and the distributive quantity expressions are adjectival.

However, the universal quantifier fails to pass the test of the third property according to Cardinaletti & Giusti (2006) as a quantifier proper.

(96) a. *{tote} locurle {*tote}*   (IR)

b. *{toate} locurile {toate}*   (Rom.)

> {all} places-the {all}
>
> 'all the places'

(97)　　*Tote locurle scu mušate.*　　　　　/ **locurle tote*

　　　　all places-the are beautiful

In IR, the noun raises to SpecDP, where the affixal definite article attaches to the it. In Romanian, the DP (N+D°) further moves to SpecQP, but in IR this movement does not take place.

　　One possible cause for the unavailablity of the order **locurle tote* is due to the Slavic-like necessarily prenominal position of modifiers and quantifiers (see Giusti & Leko 2004, who show that Serbo-Croatian *mnogo* 'many' can surface postnominally with pronouns).

　　If we consider the fourth propery, namely the discontinuous position of the quantifier from the noun, the universal *toţ* patterns again with quantifiers proper:

(98)　a.　　*Maria fečori vezut-a toţ.*　　　　(IR)

　　　　　　Maria boys-the saw-has all

　　b.　　*Maria copiii i-a văzut pe toţi.*　　　(Rom.)

　　　　　　Maria boys-the CL has seen Prep all

　　　　　　'Maria saw all the boys.'

　　As for existential and distributive quantifiers, the data at our hand is incomplete for all the properties in Cardinaletti & Giusti (2006).

　　If universal Q *toţ* is a quantifier proper, while existential *čuda* is a quantifier adjective, it can be predicted that the two can co-occur. This prediction is indeed borne out for IR, as for Romanian and Croatian:

(99)　[QP [Q° *Q* [DP　Dem [D° (art.) [XP *QAdj* [X° [NP]]

(100)　a. *toţ čel'i čuda omir*　　　　　　　(IR)

　　　　b. *toţi acei mulţi oameni*　　　　　　(Rom.)

　　　　c. *svi oni mnogi dečki*　　　　　　　(Cr.)

　　　　　　all those many people

The data presented in this section on the one hand brought one example of "mixed system" in the sense of Kovačec (1967), where a Slavic-like requirement for modifiers to precede the noun takes over the (cross-linguistic) property for a universal quantifier to be able to follow the noun.

In the last section of the present chapter, we will consider the order of all types of nominal modifiers.

3.4. Noun modifiers

Noun modifiers surface in orders that are either the base-generated one (see the hierarchy below), or derived orders.

(101)   [DP D [XP possessive X [YP cardinal Y [WP ordinal W [ZP quality Z
        [HP size H [LP shape L [MP color M [RP nationality R [NP N]]]]]]]]]]

                                               (Brugé 2002, building on Cinque 1994)

Let us take the relative orders of the the all modifiers in turn.

3.4.1. The order of Demonstrative and Numerals and Noun

The data indicates that the Dem must always precede the Num modifier, be it cardinal or ordinal. As such, no movement is allowed to take place.

(102)   *a.*   *Čåste trei selişt âs prope.*   **Dem>card.Num>N**
             these three villages are nearby
      b.   **trei čeaste selişt*   ***card.Num>Dem>N**

(103)   a.   *Čåsta prva cåsa je prope.*
             this   first house is nearby   **Dem>ord.Num>N**
      b.   *prva čeasta casa   ***ord.Num>Dem>N**

(104) a**.** *Trei prve cåse ås prope.* **card.Num>ord.Num**

three first houses are nearby

b. **prve trei cåse* ***ord.Num>card.Num**

first three houses

3.4.2. The order of Demonstrative, Possessive and Noun

We have already seen some of the orders in the sections dedicated to demonstratives (3.2.2) and possessives (3.2.3). Here we provide more examples from Žejanski, which apparently allows for more flexible orders compared to the southern variety. [40]

(106) a. *česta a nostru teren* **Dem>Poss>N – unmarked**

this of ours field

b. *česta teren a nostru* **Dem>N>Poss**

this field of ours

c. *a nostru česta teren* **Poss>Dem>N**

of ours this field

(107) *A nostru česta filj je mårle.*

of our   this   son is big

'This son of ours is big.'

3.4.3. The order of Possessive, Numeral and Noun

Only one order is allowed in both varieties. All other combinations are ruled out.

(108) a. *A nostre trei fęte ås mår.* **Poss>Num>N**

of our three girls are big

b. **a nostre fęte trei* ***Poss>N>Num**

of our girls three

The position of the adjectival modifiers will be dealt with in the following section.

---

[40] Data provided by Mauro Doričić in Triest, 2009.

3.5. Adjectival modifiers

3.5.1. The position of the adjective

Let us thus now look at the positions that adjectives can occupy in IR.

Under the influence of (Čakavian/) Croatian, IR descriptive adjectives (generally) occupy a prenominal position, whereas in Romanian they generally occur postnominally when unmarked (apart from the different semantics of adjectives such as *biet* or *sărac*, namely "wretched" when preposed, "poor" when postposed, which is well-known in Romance). Apparently in certain contexts adjectival postposition in IR is optional and/or marks a stylistic change. Based on texts published in 1929 and 1959 by Leca Morariu and Traian Cantemir, respectively, and also on his own unpublished texts, Kovačec (1971) observes however that the Slavic system has not been completely adopted, given that IR adjectives are postposed much more frequently than in Čakavian and that some occurrences would be completely ungrammatical if translated literally in the latter (i.e. (120)-(122) below).[41]

IR adjectives occur pronominally (109), except for nationality adjectives derived by suffixation with –*an* which are obligatorily postposed (cf. 124 below).[42]

---

[41] Kovačec (1971:76) notes that Čakavian allows postnominal adjectives under specific conditions, such as compounds or genitives:

   i.   *mama stara*
       mother old
       'grandmother'

   ii. *Bulin Gražanov*
       Bulin Gražan.Gen.Sg
       'Bulin, son of Gražan'

[42] In a nominal raising analysis, we observe that the noun must obligatorily raise above the "nationality – *an*" adjectival projection in IR. In unmarked orders, it does not surface any higher than this position.

   i.   [DP D [XP possessive X [YP cardinal Y [WP ordinal W [ZP quality Z [HP size H [LP shape L [MP color M [RP nationality -*ski* R [PP nationality -*an* P [NP N]]]]]]]]]]]

             (hierarchy adapted from Brugé 1996, 2002, based on Cinque 1994)

(109) a. *a nostra måra tradiţiia*

of our   big   tradition

b. *glavna uliţ*

main   streets

c. *grumbo vrẹme*

ugly   weather   (ex. from Sârbu 2009)

Adjectives can surface postnominally under the following conditions:

A. the adjective expresses a general (or intrinsic) characteristic ("quality") of the noun:

(110) a. *nẹgrẹ pâre / pâre nẹgrẹ*

black bread / bread black

b. ?*diţa mica / mica diţa*

childern small / small children

B. the adjective is contrasted:

(111) *Bevu   viru   åb,   ne   viru   negru.*

drink.1SG wine=the white, not wine=the black

C.if two or more adjectives modify the noun, they can either both occur prenominally or 'sandwich' the noun. This occurrences are difficult to assimilate to Croatian, in which all adjectives are preposed.

(112)   a. *tirer   mušat   fečor*

young   handsome   boy

b. *tirer fečor mušat*

c. **fečor tirer mušat*

D. two or more adjectives are coordinated:

(113) *Mę {mušata ši måre}   cåsa   {mušata ši måre}   ei prope de måre.*
my {beautiful and big} house {beautiful and bog} is close to sea

E. those nationality adjectives derived by suffixation with –*an* (as opposed to –*ski*, borrowed from Croatian and obligatorily prenominal) are always postnominal:

(114) *Io cunosc ur taljanski fečor / ur fečor taljan / ur fečor din Italje.*
I know   an Italian boy   / a boy Italian   / a boy from Italy

F. restrictive adjectives – which may originate from a reduced relative clause, cf. Cinque (2010), can also (marginally, for some speakers) occur postnominally.

(115) *Io-m   vezut doi {otrovni}   šarpel'i {(cåre scu) otrovni}.*
I =have1SG seen two {poisonous} snakes   {(which are) poisonous}.

G. Participial adjectives occur both pre- and postonimally. Their prenominal position may be related to 'affectivity' (thus it would not be optional), but speakers' judgements are not very sharp. Some speakers notice no difference between (22.a) and (22.b), for example.

(116)   *{uscåte} lęmne {uscåte}*
{dry} wood {dry}

H. the adjective is further modified by a PP. This construction is also found in Čakavian.

(117) *oriži pomešani s krvun*   (Čakavian)
rice mixed   with blood

(ex. from Kovačec 1971:77)

(118) a. *orižile   zmišeite cu sânže*   (IR)

rice=the mixed with blood

b. *hlębe de pâre åbe ca ši låptele*

loaves of bread white as milk=the

c. *cåsa zâgârnita cu slåmę*

house covered with thatch

(adapted from Kovačec 1971:83)[43]

It must also be noted that some instances of A > N (both of which have Romance origins) are frozen expressions which in Romanian have the same order:

(119)   *de burę vol'e / irime*

of good will / heart

'with great pleasure'

(120)   *bura damareţa / bura zi / burę sęrę*

good morning / good day / good evening

The cases in which speakers differentiate between the pre- and postnominal use of the adjective are rare (cf. 121.a). Apparently the influence of the position of the Croatian adjective is stronger that the need to express semantic differences (which can be rendered via specific intonation). In fact, my informants do not assign different meanings to the two positions of the adjective in (121.b) (cf. Romanian in 28).

(121)   a. *porcu divl'u* / *divl'u porcu*

pig savage savage pig

(savage = non domestic) (savage, referred to behaviour)

b. *Redento-i ur {måre} om {måre}.*

Redento is a {tall/great} man {tall/great}

---

[43] Clarification note: "adapted from" refers to those examples which I have found in older literature and which I have checked with my informants. They judged them either as marginal or ungrammatical – in which case the sentences have been modified/labeled according to the speakers' judgements.

(122)   *Grigorescu a fost   un {mare} om {mare}*   (Rom.)

   Grigorescu has been   a {great} man {tall}

As for classificatory adjectives, some precede the noun (123.a), others follow the noun (123.b). Note that *osnova* is borrowed from Croatian, while *catolica* is derived from Latin.

(123)   a. *{osnova} škola {*osnova}*

   primary school

  b. *{*catolica} besęreca {catolica}*

    church catholic

The speaker's attitude (contrastive focus, diminutives and augmentatives, affectivity) can be expressed through a particular intonation, or by postponing the adjective thus obtaining a stylistic effect. Kovacec (1971) shows the following cases, in which speaker's attitude is strongly felt:

(124)   a. *La a tå   måie   prva rem ĭi.*

   to of your mother first would.like.1PL go

   'We'd like to go see your real mother (not your adoptive one)'

  b. *Tu   sti   haiduc   måre!*

   you   are.2SG brigand   big

   'You are a great brigand!'

  c. *Pac fost -av   åntr -o   cåsiţa siromåšna.*

   and been=have.3PL in a house(little) poor

   'And they lived in a poor little house.'

  d. *Osåndit   -l   -a   la devet an'   ån tamniţe šcure.*

   condemned =himCL =has   to nine years in prison   dark

   'They have condemned him to nine years of (hard) prison.'

         (adapted from Kovačec 1971:82)

The data above show that the possibility of pre-/postnominal alternation can be correlated with the degree to which a nominal expression is 'fixed' in the language. This observation

strenghtens the view in Kovačec (1971) that the Croatian/Čakavian system (strict A > N) has not been completely adopted even after four more decades (with respect to Kovačec's conclusions) of continuous and intense contact with the macro-language.

3.5.2. The comparative/superlative

The constructions that are used for comparative and superlative are given below:

(125)   (*čela* / *čå*)   > *måi* > A > N
      the.Masc/the.Fem   more

The parenthesis indicate that for the comparative *čela* / *čå* are employed only exceptionally, whereas the presence *čela* / *čå* are obligatory for the formation of the superlative.

   The optionality of the "demonstrative article" in superlative constructions in IR contrasts with its obligativity in the Romanian parallel construction. Croatian forms the superlative with the (bound) morpheme *naj-* preposed with respect to the adjective, thus the IR morpheme *mai* (which has identical distribution) is identified with this morpheme. Croatian does not have "demonstrative articles", so there is the tendency to drop these elements in IR under the Croatian influence. When identical constructions (i.e. when *čela* / *čå* is dropped for the superlative), the comparative and the superlative in IR are distinguished by the accent on *mai*: absent with the comparative, present with the superlative

3.5.3. The modifiers of the adjective

The modifiers of the adjectives are invariable adverbials (i.e. do not agree in number/gender). Two different constructions are given below:

(126)   *Me t''aro mare casa-i prope de mare.*
      my very big house is close to sea

(127)   *amår de munt vreme*
      bitter of much time
      'incredibly much time'

3.6. Ad interim conclusion

Chapter 3 examined the noun phrase in Istro-Romanian. After having illustrated noun classes, we took a more in depth investigation of definiteness, i.e. its morphosyntactic realization through definite articles, demonstrative adjectives, and demonstrative (or "pseudo-") articles. We offered a tentative analysis along the lines of a recent proposal for Romanian and other languages, leaving however some open questions.

Secondly, we exemplified and analyzed the properties of quantity expressions in Istro-Romanian, following two recent studies that showed how quantifiers proper are different from quantifier nouns and adjectives.

Thirdly, we presented the order(s) in which modifiers (demonstratives, numerals and possessives) are allowed to appear in Istro-Romanian

Finally, we offered data for Istro-Romanian adjectives, regarding their position with respect to the noun which corresponds to the hierarchical order as it evinces from cartographic studies.
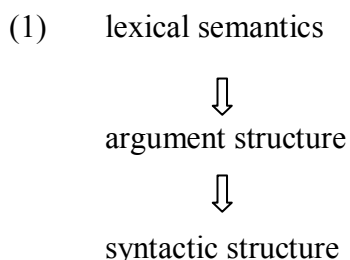
Chapter 4

The Verb Phrase

In this chapter we will be concerned with the verbal elements: their semantics and argument structure, and the expression of functional categories. We do not take any theoretical stand most of the times, but will however briefly refer to some theoretical aspects.

4.1. Classification of lexical verbs. Argument structure (transitive, intransitive, unaccusative, psychological, reflexive, meteorological verbs)

4.1.1. Transitive, intransitive, unaccusative verbs

"[A]rgument structure is an interface between the semantics and syntax of predicators (which we may take to be verbs in the general case)... Argument structure encodes lexical information about the number of arguments, their syntactic type, and their hierarchical organization necessary for the mapping to syntactic structure." (Bresnan 2001:304)

Argument structure (or theta-grid) stands at the interface between semantic roles (or theta-roles) and syntactic functions. Semantically, it is the representation of event structure. Syntactically, it functions as a subcategorization specification of the relations between argument-taking heads and their syntactic dependents (complement, specifier). Thus, it is a link between semantics and syntactic structure.

(1)   lexical semantics

⇓

argument structure

⇓

syntactic structure

In particular, it contains syntactically relevant information, whereas semantic roles are less prominently relevant for syntax. At the same time, though, argument structure is not affected by syntactic operations such as passivisation; only the mapping of the arguments into syntactic positions is affected. [1]

Lexical heads (verbs, nouns) introduce the participants in the events; the number of arguments of one and the same lexical entry can vary. For instance, the head *open* obligatorily introduces one argument (the Theme), and may optionally introduce others (an Agent, an Instrument). Similarly, the IR verb for *eat* may leave out the Theme/Patient argument. Differently from English but in line with the Romance pattern, IR is a pro-drop language, namely it needs not project overtly the subject. In syntax however, the subject is present and it features (person and number) are retrievable from the verbal inflection with which the null subject agrees.

(2)   a.   The door opened.

        b.   John opened the door.

        c.   John opened the door with the key.

(3)   (*Căprele) mârâncu iårba.*

        goats-the eat.3Pl grass-the

        'The goats eat the grass.'

A very influential approach is that of Grimshaw (1990), who defines argument structure on (purely) semantic grounds. She assumes that thematic roles are ordered

---

[1]Baker (1988) is widely referred to as the earliest hypothesis (within the Principles and Parameters framework) for a constrained mapping of theta-roles into syntax:

    i.   UTAH = Uniformity of Theta Assignment Hypothesis: Identical thematic relationships between items are represented by identical structural relationships between those items at the level of D-structure. (Baker 1988:46).

Thus, syntactic operations (which in the old P&P framework took place between the D-structure and the S-structure) do not affect the base-generation of the arguments.

Section 2.3 deals with changes to argument structure after syntactic operations have applied.

hierarchically, with the Agent as the highest and the Theme as the lowest in the hierarchy.

(4) Agent < Experiencer < Goal/Source/Location < Theme

For Grimshaw, argument prominence is an interaction of the thematic hierarchy with the aspectual one, which has Cause as the highest element. [2]

The aspectual hierarchy is based on the event structure of a predicate, where "the event structure represents the aspectual analysis of the clause, and determines such things as which adjuncts are admissible... "(Grimshaw 1990: 26). For example, the event structure of an accomplishment verb, such as *build*, can be divided in two sub-events, namely an activity and a (resulting) state.[3] Grimshaw proposes that the most prominent argument on the aspectual hierarchy is always realized syntactically as a D-structure subject, irrespective of its thematic role. Moreover, in order for an element to be an external argument, it must be the most prominent in both hierarchies. Thus, Agents are always external arguments when present. If there is a conflict (i.e. an argument is less prominent on the thematic tier, but more so on the aspectual one), only internal arguments will be projected.

We will not dwell on this theoretical aspect in any further detail. We will refer to argument structure and theta roles in what follows, as we will be dealing with verb classes.

Function of the number of (obligatory) arguments they take, and on whether the arguments are internal or external, verbs have been categorized in several classes. Transitive verbs obligatorily project one internal argument and one external argument. The Unaccusativity Hypothesis (Perlmutter 1978, Burzio 1986) distinguishes two subclasses of intransitives: unergative and the unaccusative, the syntactic structure of which is different in a crucial way. Unergative verbs (e.g., *run*) have their unique

---

[2] The other elements (beside Cause) in the aspectual hierarchy are not specified in Grimshaw (1990).

[3] In Grimshaw's view, an argument that participates in the first sub-event is more prominent than an argument participating in the second sub-event, because it will be causative or agentive, whereas (resulting) states are undergoers (patients or themes).

argument projected as an external argument, namely in [Spec, vP],[4] but project no internal argument, thus the VP does not select a complement. Unaccusatives (e.g., arrive), on the other hand, generate their sole argument in [Compl, VP] position, while the [Spec, vP] position is empty. In thematic terms, the unergative verbs assign Agent role to their unique argument, while the unaccusative verbs take a Patient or Themes

The internal argument of both subclasses of intransitive verbs surface as subjects. The XP in [Spec, vP] for unergatives, and the XP in [Compl, VP] for unaccusatives moves (i.e. by 'move α') to the surface subject position, [Spec, IP] to receive / check the nominative case.

(5) a. unergative verb $[_{vP}$ DP v $[_{VP}$ V $]$

  b. unaccusative verb $[_{vP}$ v $[_{VP}$ V DP$]$

Burzio (1986) related the ability of a verb to have an external argument with the structural case assignment. According to Burzio's Generalization, unaccusative verbs do not take an Accusative direct object. In older Government and Binding terms, unaccusatives do not assign structural case to their object.

(6) a. A verb which lacks an external argument fails to assign accusative case.

  b. A verb which fails to assign accusative case fails to θ-mark an external argument.

(Burzio 1986:178–179, 184)

The three possible argument structures of verbs are given below in brief:

---

[4] [Spec, vP] is the canonical position in which the external argument is generated. As opposed to previous views according to which external arguments were projected higher in the structure (e.g., in [Spec, IP], Chomsky 1986), the VP Internal Subject Hypothesis (CITA) enabled the assignment of all theta-roles internally to the verb phrase, with subsequent movement of the external argument in subject position. Cf. also Larson (1988), Hale and Keyser (1993, 2003), Chomsky (1995), among others for the VP-shell/vP hypothesis.

1. Transitive verb: two arguments, two theta-roles; able to Accusative case-mark its syntactic complement.

2. Unergative verb: only one, external argument.

3. Unaccusative verb: only one, internal argument; unable to Accusative case-mark its syntactic complement, which receives Nominative case in subject position.

The members of these verb classes do not have a homogeneous behavior. Some verbs can enter transitive-unaccusative alternations (e.g., *sink*), others cannot (e.g., *fall*). Unergative verbs do not display alternating patterns.

(7) a.   The enemy sank the ship.

  b.   The ship sank.

  c.   *John fell the tree.

  d.   The tree fell.

(8) a.   *The car came John.

  b.   John came.

Examples (9)-(11) illustrate members of each verb class in Istro-Romanian.

(9) *Cåprele mârâncu iårba.*   (transitive)
  goats-the eat grass-the
  'Goats eat grass.'

(10) *Våčile dormu ân štåla.*   (unergative)
  cows-the sleep in stable
  'Cows sleep in the stable.'

(11) *Vire ur cârstiån.*   (unaccusative)
  comes a man
  'A man is coming.'

It is well-known that compound past tenses in some Romance languages can display auxiliary selection requirements.[5] Verbs requiring *be* in French and Italian are ergative (in Burzio's terms), namely they do not assign a theta role to their syntactic subject, but they assign one to the object position, which however does not receive structural Accusative case in situ. Like Romanian, Spanish and Portuguese, but unlike French and Italian, IR does not display the auxiliary *be-have* alternation. Thus, the perfect auxiliary (in the active voice) for all verb classes above is *avę* 'have'. Examples are provided below:

(12)   *Ier Lara nu-a poidit paninu.*[6]   (transitive)
       yesterday Lara not have.3Sg eaten sandwich
       'Yesterday Lara didn't eat the sandwich.'

(13)   *Durmit-am o nopte la Nela.*   (unergative)
       slept have.1Sg a night at Nela
       'I slept at Nela's one night.'

(14)   *Verit-a ur cârstiăn.*   (unaccusative)
       come have.3Sg a man.
       'A man came.'

4.1.1.1. More on unaccusativity vs. unergativity

Many authors argue that unaccusativity is syntactically a unified phenomenon, but maintain that the distinction between unaccusatives and unergatives is semantically encoded. Perlmutter's original formulation of the Unaccusative Hypothesis states that the difference between unergatives and unaccusatives is syntactically represented. Two approaches to unaccusativity developed in response to the Unaccusative Hypothesis: the syntactic approach and the semantic approach.

(15)   *The Unaccusative Hypothesis*
       Certain intransitive clauses have an initial 2 but no initial 1.

---

[5] Other languages belonging to different families (i.e. Dutch, Basque) also display auxiliary alternation.

[6] *Poidi* is a Slavic borrowing, while *mâncă/ muncă* (*mârânc, mancât*) has Latin origins.

In other words, unaccusative predicates select a single internal argument, while unergative predicates select a single external argument.

Defenders of the syntactic approach, amongst which Burzio (1986), argue that all unaccusative predicates, regardless of their semantic class, share certain syntactic properties (i.e., the inability to assign accusative case, selection of a single internal argument, and lack of an external argument). Supporters of the semantic approach, on the other hand (cf. Van Valin 1990, among many others) argue that unaccusativity is not a unified syntactic phenomenon based on the fact that some verbs can test as unaccusative for one diagnostic and unergative for another.

Levin & Rappaport Hovav (1995) bring both approaches together, presenting detailed arguments from a variety of languages in defense of Perlmutter's original approach to unaccusativity, i.e., they argue that unaccusativity is syntactically a unified phenomenon, but maintain that the distinction between unaccusatives and unergatives is semantically encoded.

Burzio (1986), one of the strongest defenders of the syntactic approach to unaccusativity, transposes Perlmutter's Unaccusative Hypothesis into the Government-Binding (GB) framework of Chomsky (1981). He provides an extensive study of unaccusativity in Italian, positing several syntactic diagnostics for distinguishing between unaccusatives and unergatives. He shows that unaccusatives differ syntactically from unergatives in Italian in (at least) three distinct ways. Unaccusatives:

(a) select *essere* 'to be' as their auxiliary in the past tense (vs. *avere* 'to have') (in languages which display auxiliary selection, as does Italian),

(b) show past participle agreement with their subject, and

(c) allow for *ne*-cliticization (i.e., extraction from the direct object position).

Several of Burzio's examples are shown below for Italian.

(16)   *Auxiliary selection*

    a   *Maria è/*ha arrivata.*   (unaccusative)

           Maria is/*has arrived

           'Maria arrived.'

    b   *Maria *è/ha telefonato.*   (unergative)

           Maria *is/has called

           'Maria called.'

(17)   *Past Participle agreement*

    a   *Maria è arrivat-a/*-o.*   (unaccusative)

           Maria is arrived.Fem./*Masc.

           'Maria arrived.'

    b   *Maria ha telefonat-o/*-a.*   (unergative)

           Maria has called-Masc./*Fem.

           'Maria called.'

(18)   Ne-*cliticization*

    a   *Giovanni ne inviterà molti* t.   (transitive)

           Giovanni of-them will-invite many

           'Giovanni will invite many of them.'

    b   *Ne arrivano molti* t.   (unaccusative)

           of-them arrive many

           'Many of them will arrive.'

    c   **Ne telefonano molti* t .   (unergative)

           of-them telephone many

           'Many of them will call.'

In (18.a), *ne* 'of them' raises from the direct object *molti* (*ne*) 'many (of them)' and cliticizes to the verb *inviterà* 'will invite.' Similarly, in (18.b), *ne* raises from the subject

molti and cliticizes to the verb *arrivano* 'arrive'. However, with the unergative verb *telefonano* 'telephone' in (18.c), this movement is not licit.

Based on these data, Burzio (1986) concludes that subjects of unaccusatives pattern syntactically with direct objects of transitive predicates. Given this conclusion, Burzio predicts that unaccusative subjects should also pattern with subjects of passive predicates.

These data strengthen the claim that *unaccusative subjects are really underlying direct objects*. However, unlike direct objects, these subjects canonically receive nominative case as opposed to accusative case. Based on this observation, Burzio (1986) notes a correlation between the ability of a verb to take an external argument and its ability to assign accusative case. This correlation has come to be known as Burzio's Generalization (see above).

Since Burzio's (1986) original formulation of his generalization, much work has suggested the need to reevaluate it, particularly in light of data from a variety of languages indicating that some verbs which do not select an external argument are in fact capable of assigning accusative Case (see for instance Bowers 2002).[7]

4.1.2. Psychological verbs

Psychological verbs have long been a case study for the relation between syntax and argument structure.

From the thematic point of view, psych verbs display a common property: they all involve a human participant emotionally related to some other participant , i.e. psych verbs always make available an Experiencer role which is mapped into different syntactic positions. The interpretations usually attributed to the Experiencer role refer to mental/psychological change, or to a change of state of the [+human] participant.

> "A psych-verb is any verb that carries psychological entailments w.r.t. one of its arguments (the experiencer). A psychological entailment involves an individual being in a certain mental state. Thus *frighten* is a psych verb since *Mary frightened*

---

[7] The cliticization of *ne* is an irrelevant test in IR. Possible tests are related to word order, i.e. postverbal subjects in presentative clauses. At the time we had the sessions with the informants, we had not taken into account the relevant data.

Bill entails that Bill is in a certain mental state (i.e., fright); whereas *invite* is not a psych verb, since *Mary invited Bill* carries no entailments as to Mary's or Bill's state of mind (although it does entail that both are human)."

(Landau 2005:22)

One of the most influential studies of psychological verbs is a Belletti & Rizzi (1988). It offers a tripartite classification of psychological verbs, based on syntactic properties. Moreover, they are concerned with the mapping of the θ-grids onto D-structures (in the GB framework), since the Experiencer and Theme are generated in different syntactic positions.

(19)   *B&R (1988) tripartite classification of psychological verbs*

      a   Class 1: Nominative Experiencer (*temere/* 'fear')

            Gianni teme i terremoti.

            Gianni fears the earthquakes

      b   Class 1: Accusative Experiencer (*preoccupare/* 'worry')

            Questi problemi preoccupano Gianni.

            these problems worry Gianni

      c   Class 3: Dative Experiencer (*piacere/* 'appeal to')

            i.   *A Gianni piace la musica.*

                  to Gianni appeals the music

            ii.   *La musica piace a Gianni.*

                  the music appeals to Gianni

Their proposal is that there are two types of D-structure for psych verbs:

(20)　a　　Nominative Experiencer
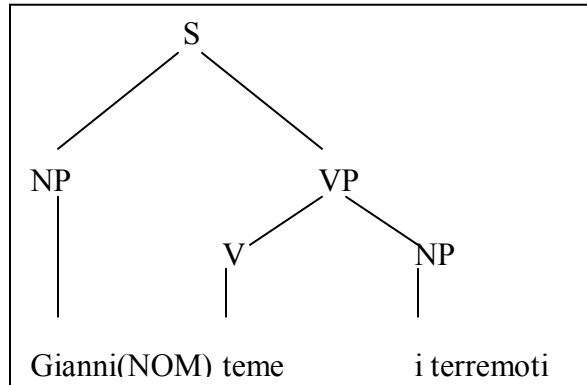
b　　Accusative / Dative Experiencer
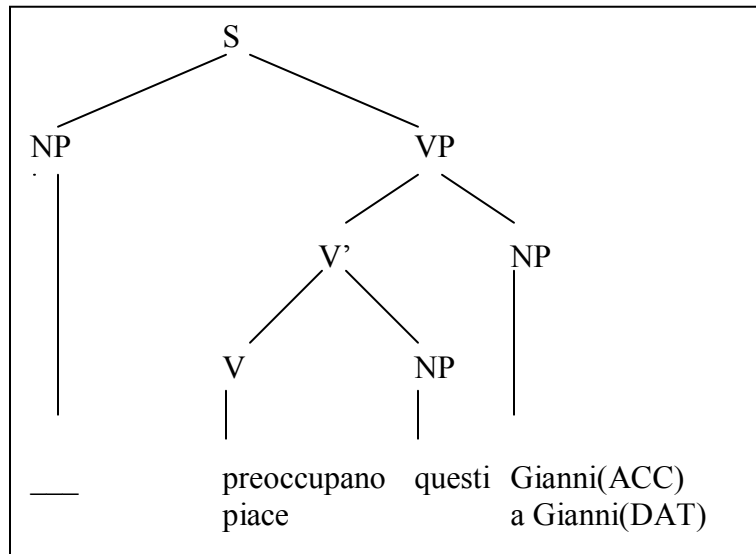
Although recent approaches to generative grammar have abandoned the two levels of representation (D- and S-structures), the tripartite classification of psych verbs is still empirically valid. In what follows, we will illustrate the main syntactic properties of members of the three classes of psychological verbs in Istro-Romanian. We will not make any theoretical claims, but merely present the facts from the data collected. We have not been able to identify psych verbs belonging to the second class (

4.1.2.1. Class 1, Nominative Experiencer: *piažę* (1) [8]

The members of this subclass are transitive verbs with a Nominative Experiencer as a subject. There is a variety of opinions as to the best characterization of the "semantic role" of their direct object; the labels used include Theme, Target of emotion, Stimulus, and Subject matter (see the discussion in Pesetsky 1995, for instance).

The order of the two arguments is free. However, when both are proper names (21), thus bearing no morphological marking nor any functional marker for the Accusative as in Romanian (22) (the preposition *pe* is used for animate, specific objects), ambiguity may arise. The unmarked order seems to be with the Experiencer linearly preceding the Theme argument. In (21), the interpretation that Gianni is the holder of the psychological state very likely involves an operation related to information structure (*Maria* as a Topic or Focus) which makes it so that Maria (Theme) precedes Gianni (Experiencer). In spoken language, intonation (precisely related to information structure) disambiguates.

(21) *Maria piažę Gianni.*

Maria like3Sg Gianni

'Mary likes Gianni' or 'Gianni likes Mary'.

(22) *Maria îl place pe Ion.*   (Romanian)

Maria himCL like.3Sg Ion

'Maria likes Ion'.

Pronominal arguments leave no room for ambiguities. The examples below illustrate the Accusative form of a pronominal Theme, and the Nominative pronominal Experiencer. Both possible orders are allowed.

(23) *Maria piažę tire.*   / *Tire piažę Maria.*

Maria like.3Sg you.Acc   / you.Acc like.3Sg Maria

'Maria likes you'

Exp NOM > Theme ACC   Theme ACC > Exp NOM

---

[8] In IR there is a pair of verbs *piažę* (1) and *piažę se* (3) corresponding roughly to 'like, fancy (something or someone)' and 'like, appeal to', respectively. The latter is a Class 3 psych verb, to be discussed below.

(24)   *Io piažes Gianni.*   */ Gianni piažes io.*

         I like.1Sg Gianni   / Gianni like.1Sg I

         'I like Gianni'

         Exp NOM > Theme ACC   Theme ACC > Exp NOM

If there is number mismatch between the Experiencer and the Theme argument, subject agreement with the verb will disambiguate (25).

(25)   *Maria piažę Gianni ši Goran.*   */ Gianni ši Goran piažę Maria.*

         Maria like3Sg Gianni and Goran   / Gianni and Goran like3Sg Maria

         'Mary likes Gianni and Goran'.

4.1.2.2. Class 2, Accusative Experiencer: *piažę* (2)

The members of this subclass of psych-verbs describe the bringing about of a change in the psychological or emotional state. They are Object Experiencer verbs, whose subject is the cause of the change in the psychological state.

A second lexical entry for *piažę* 'like' selects for an Accusative Experiencer and a Nominative Theme. Pronominal forms are illustrative. Compare with the Romanian *a încânta* (lit. 'to enchant').

(26)   *Šcola nu me piažę.*

         school not me.Cl.Acc like.3Sg

         'School doesn't appeal to me.'

(27)   *Şcoala nu mă încântă.*   (Rom.)

         school not me.Cl.Acc enchanted.3Sg

         'I'm not mad about school.'

Differently from Italian, which is more restrictive for the orders available for Class 2 psych verbs depending on whether the Experiencer is pronominal clitic or a strong

pronoun / DP, Romanian can display the reversed order of the two arguments irrespectively.[9]

(28) a. *Questi problemi preoccupano Gianni.* (Ita.)

these problems worry.3Pl Gianni

b. **Gianni preoccupano questi problemi.*

Theme NOM > Exp ACC   (a)
* Exp ACC > Theme NOM   (b)

(29) a. *Questi problemi lo preoccupano.* (Ita.)

these problems worry.3Pl he.Acc

b. *Lo preoccupano questi problemi.*

him.Cl.Acc worry.3Pl these problems

'He is worried about these problems.'

Theme NOM > Exp ACC   (a)
Exp ACC > Theme NOM   (b)

(30) a. *Problemele acestea îl îngrijorează pe Ion / pe el.* (Rom.)

problems-the these him.Cl.Acc worry.3Pl *pe* Ion / *pe* he

b. *Pe Ion / pe el îl îngrijorează problemele acestea.*

pe Ion / *pe* he him.Cl.Acc worry.3Pl problems-the these

'He is worried about these problems.'

As for IR, both orders are available when the Accusative Experiencer is a clitic pronoun:

---

[9] As can be seen in the examples, Romanian displays clitic doubling of the internal argument when it has definite reference. Also, the functional preposition *pe* introduces Accusative specific arguments.

(31)   *Šcola nu me piažę.*   / *Nu me piažę šcola.*

       school not me.Cl.Acc like.3Sg   / not me.Acc like.3Sg school

       'School doesn't appeal to me.'

       Theme NOM > Exp ACC   Exp ACC > Theme NOM

In the case of nominal DPs, recall (from Class 1 psych verbs) that IR does not mark Accusative arguments neither morphologically nor through a functional preposition (as with Romanian *pe*, or with Dative *lu* in IR). Thus, a full DP Experiencer could be either Nominative hence belonging to Class 1 (with an Accusative Theme, also unmarked), or Accusative belonging as for Class 2 (with a Nominative Theme). There would be no instruments to distinguish between the two. [10]

4.1.2.3. Class 3, Dative Experiencer: *piažę se*

The pattern of argument expression of Class 3[11] maps the Theme as Nominative as a subject, and the Experiencer as a Dative indirect object.

An intriguing form which my informants have produced regards an impersonal-like form of the verb *piažę* involving an invariable morpheme *se*. The orders available are illustrated below, both with a pronominal Experiencer and with full DPs.

(32)   *Ţie se piažę Maria.*   / *Maria se piažę ţie.*

       you.Dat *se* like.3Sg Maria   / Maria *se* like.3Sg you.Dat

       'Maria appeals to you.'

---

[10] One of my informants produced the following context:

    i.   *Noi nu piažę šcola, ma ne piažę žucă.*

       we.Nom/Acc not like school, but we.Cl.Acc like dance

Considering that the 1st person plural strong pronoun *noi* (as well as the 2nd person pronoun) has the same from for Nominative and Accusative, the first instance of *piažę* could be either Class 1 or Class 2.

[11] Members of this smallest sub-class of psychological verbs in Italian are *piacere* 'like', *dispiacere* 'dislike', *mancare* 'lack', *bastare* 'suffice', *seccare* 'vex'.

(33)   *Lu Gianni se piažę Maria.   / Maria se piažę lu Gianni.*

to Gianni *se* like.3Sg Maria   / Maria *se* like.3Sg to Gianni

'Maria appeals to Gianni.'

Exp DAT > Theme NOM   Theme NOM > Exp DAT

4.1.3. Reflexive verbs

Reflexive verbs are mono-argumental. The general approach to Romance reflexives is an unaccusative analysis: the subject is an underlying object. However, it has also been argued for reflexives to be unergative entries.[12]

This latter approach corresponds to a view that there is a view that reflexives are derived from their transitive alternate by an operation reducing the internal argument (Reinhart and Siloni 1999).[13] Under this approach, the reflexive clitic is associated with the internal theta role in the lexicon, not in syntax. On the other hand, under the unaccusative analysis, the reflexive clitic is associated with the external theta role, and the reflexive verb is therefore an unaccusative verb, as its internal argument is the derived subject.[14]

As previously, we will not take a position as to the two different approaches, and limit ourselves to presenting the data at our hand.

The main interpretation of the reflexive pronoun *se* is anaphoric. The possibility of adding a strong reflexive together with a "intensive pronoun" (Rom. *pronume de întărire*) without altering the meaning is a reliable test.

---

[12] For instance, some research has shown that when reflexives are submitted to syntactic tests of unaccusativity, they systematically fail the tests in a variety of languages.

[13] Some of the proponents of this lexical approach are Grimshaw (1982), Wehrli (1986), Chierchia (1989) and Reinhart (1996).

[14] The research under the unaccusative approach argues either for the lexical absorption of the external argument (Bouchard 1984, Marantz 1984, Grimshaw 1990), or for an analysis according to which the external argument is present in syntax via *se* (Kayne 1988, Pesetsky 1995, Sportiche 1998).

(34) a. *Cu cǎniţa se-ncinje (âns sie / sine ânsuši)*

      with belt *relf* tie.3Sg he himself

      'He ties (her waist) with a belt.'

   b. *Brecu s-a oblatit.*

      dog-the *refl* have.3Sg dirtied

      'The dog got itself dirty.'

Other examples of reflexive (anaphoric) contexts provided by my informants are given below:

(35) a. *Če m- am   ponošeit bire ǎz?*

      Q *refl* have1.Sg behaved well today

      'Did I behave well today?'

   b. *Fečoru-i bur ke se ponošę mušǎt.*

      boy-the is good because *refl* behave nicely

      'A boy is good when he behaves well.'

(36) a. *Lu Marie   s   -a   rastagnit vestalja.*

      to Maria.Dat *refl* have.3Sg wrapped dress

   b. *Rastagnit –lj   -s -a   vestalja*

      wrapped herDatCL *refl* have.3Sg dress

      'Maria's dress got wrapped.'

Another interpretation of *se* is that of marking inalienable possession. It is thus used in constructions involving body parts, as illustrated below:

(37) a. *Io-m opilit lęmnele,   ši mi-am taljǎt žǎžetu.*

      I have.1Sg cut wood-the, and *refl* have.1Sg cut finger-the

      'I cut the wood, and I cut my finger.'

   b. *Brecu ši-a oblatit gloaca.*

      dog-the *refl* have3SG dirtied fur

      'The dog got his fur dirty.'

Further, a reflexive can bear a reciprocal interpretation:

(38)   *Gianni ši Maria se švades.*
Gianni and Maria *refl* argue
'Gianni and Maria argue (with each other)'.

As for the other uses of *se* (passive, middle construction and impersonal), they will be tackled in Section 4.2.

4.1.4. Meteorological verbs

Metereological verbs are intransitive with a *quasi-* argument (non-referential), which in non *pro-drop* languages (such as English) must be overt (the expletive pronoun *it* is inserted), but in languages that can drop the subject it is never expressed. This is the case in IR:

(39)   *Če ploià?.*   *Nu ploià.*
Q rain.3Sg   not rain.3Sg
'Is it raining? It's not raining.'

4.2.   Changes to argument structure

4.2.1. The passive voice

The passive voice is obtained by suppressing the external argument, which also corresponds to the impossibility of the verb to assign Accusative case to its object (cf. Burzio's Generalization). Thus, the internal argument moves to the subject position where it receives Nominative case.

As is the case for many oral languages, speakers of IR hardly ever use the passive. We have in fact attested no occurrences in our collected data. [15] We have found one example in Sârbu & Frățilă 1998 (in *Glosar*), which we give below:

---

[15] Elicitation of passive clauses and *se* middle constructions will have to be done in the future.

(40) *Čel'i miseţ zabraneno -åv fost de lovi.*
those months forbidden have.3Sg to hunt
'Hunting was forbidden in those months.'

4.2.2. The '*se*' middle construction

Similarly to other Romance languages, IR argument reduction may result in *se* middle constructions. Again, we provide for illustration an example from the literature:

(41) a. *Io nu ştiu cum iå uş-a rescl'is.*
I not know how she door have.3Sg opened
'I don't know how she opened the door.'
b. *Trei miseţ lovu s-åv rescl'is.*
three months hunt-the *se* have.3Sg opened
'The hunting (period) was open for three months.'

(examples from S&F 1998, *Glosar*)

4.2.3. Impersonal '*se*' construction

This construction is present in IR and rather frequently encountered. Verbs like *zice se*, *ganę se* ('it is said, (they) say')) can also select for a *ke* 'that' subordinate clause. [16]

(42) *Se zice / Se gane   k-å Maria čuda čiteit.*
se say3sg /*se* talk3sg that has Mary a lot read

4.3. Functional categories related to the verb

Let us schematically point out the main properties of verbal elements:

---

[16] This type of biclausal construction (*se* impersonal main verb and subordinate complement) appears extensively in my data. It exemplifies an intransitive *se* impersonal, whereas we have not elicited neither for transitive nor for unaccusative verbs.

- Verbs inflect for person and number (agreement with the subject)
- In compound tenses, it is the auxiliary that inflects.
- in passive voice, the auxiliary agrees for tense, person and number, the participle agrees for gender and number
- There is no object agreement.

In what follows, we will illustrate some classes of semi-functional and functional verbs in IR.

4.3.1. Modal verbs

Modals are functional verbs which lack argument structure. They combine with lexical verbs and form either monoclausal (if the embedded clause is infinitival) or biclausal constructions (if the subordinate verb bears a finite mood).

Most modal verbs can express both *epistemic* and *root* (or *deontic*) modality. Epistemicity deals with the evaluation that the speaker gives to the truth value of a proposition, based on logical inference. It is thus speaker-oriented, and it is used to express certainties, doubts, or guesses. Root or deontic modality is concerned with obligation and permission. It is discourse-oriented, rather than speaker-oriented.

In what follows, we will mainly deal with the constructions involving modal verbs as it emerged from the data that has been collected.

In Istro-Romanian, the constructions with modal verbs (*morę, treba* 'must', *putę* 'can', *vrę* 'will') display an inflected modal and an infinitival lexical verb (as in Italian). In Romanian, only *a putea* 'can' is allowed to optionally take a short infinitive instead of the much more largely used subjunctive form.[17] Examples are given below (also see Chapter 2, Section 2.5.4):

---

[17] We have seen in Chapter 2 that while Romanian patterns with Balkan languages with respect to the large substitution of infinitival complements by finite (subjunctive) ones, IR preserved some uses of the infinitives. Modal + infinitival lexical verbs is one of these syntactic contexts.

(43)   a.   *Putut-am   durmi časta nopte.*
             could-have.1Sg sleep.Inf this night
             'I was able to sleep last night'.

       b.   *Acmo moręim merindå.*
             now must.1Pl eat.Inf
             'Now we have to eat.'

       c.   *Vreţ čevå popí?*
             want.2Pl something eat.Inf
             'Do you want something to eat?'   (example from S&F 1998:116)

That we are dealing with a monoclausal construction can be tested by *clitic climbing*. The clitic can either climb (a restructuring context) or stay lower, linearly between the modal and the lexical verb. However, the (feminine Accusative, in the example below) clitic pronoun cannot stay in the lowest position after the lexical verb, which is the canonical position of the direct object, in which full NPs appear:

(44)   a.   *Io voi   {vo}   putę {vo} vedę {*vo}.*
             I will.1Sg {her} can {her} see {her}
             'I'll be able to see her'.

       b.   Io voi   (vo) putę vedę Lara.
             I will.1Sg her.Cl can see Lara
             'I'll be able to see Lara'

Note that in the (44.b) example above, the clitic pronoun is optional according to my informant. This optionality of clitic doubling is largely unclear to us at the moment, and we must leave it aside for now.

The behavior of the modals differs from that of the auxiliaries when it comes to ellipsis as an answering strategy (for auxiliaries, see Section 4.3.3 below). Consider the following examples:

(45)   a.   *Če poţi vedea Lara*?

Prt can.2sg see Lara?

'Can you see Lara'?

b.   *Poč vo vedę*   / *Poč.*   / **Poč vo.*

can.1sg her see.   /can.1sg   / can.1sg her

'I can see her' / I can.'

What we note is that either lexical verb and its internal argument are elided altogether, or they must both be overt.

4.3.2. Apectual verbs

As for aspectual verbs in IR, they combine with an infinitival lexical verb, arguably also giving rise to a monoclausal construction. In contrast, in Romanian the lexical verb is finite (subjunctive) and the construction is biclausal.

Some aspectual verbs are: *pošnę* 'start', *fini* 'finish', *provęi* 'try'. Some examples collected are given in what follows:

(46)   *Redento pošne lucrå na šapte pir la trei.*

Redento start.3Sg work.Inf at seven until three

'Redento starts working from seven until 3 o'clock.'

(47)   a.   *Pošni-voi prontivęi zeama.*

(I) start will prepare.Inf soup-the

b.   *Io voi pošni prontivęi zeama.*

I will start prepare.Inf soup-the

'(I) will start preparing the soup.'

c.   *Io voi vo pošni (*vo) prontivęi.*

I will it.Cl.Fem start prepare.Inf

'I will start preparing it.'

In example (47.c) above we observe that the feminine clitic pronoun undergoes obligatory clitic climbing, which supports the monoclausality hypothesis of these constructions. The previous two examples illustrate the raising of the aspectual verb in the initial clausal position if the subject is dropped. Presumably this takes place because of the enclitic nature of the future auxiliary.

4.3.3. Auxiliary verbs

Auxiliaries are functional verbs which, intuitively speaking, are employed in the formation of compound tenses. Their status in IR is somehow peculiar, as it will be seen below.

    Let us start with the paradigmes for indicative perfect and future (or restrictive) tenses:

(48)   a. *Indicative perfect tense: auxiliary + participle (of* fini *'to finish')*

		SG		PL	
SG	1	*io-m finit*	PL	*noi-åm finit*	
	2	*tu-i finit*		*voi åţ finit*	
	3	*je finit-a*		*čelj finit-a*	
		(*je a finit)		*(*celj a finit)*	

        'I / you / he / we / you / they finished'

    b. *Negative form*

SG	1	*io nu-m finit*	PL	*noi nu-m finit*	
	2	*tu n-ai finit*		*voi nu-ţ finit*	
	3	*je nu-a finit*		*čelj nu-a finit*	
		(*je nu finit-a)		*(čelj nu finit-a)*	

        'I / you / he / we / you / they didn't finish'

(49)   *Indicative future tense: auxiliary + infinitive*

SG	1	*io voi fini*	PL	*noi rem fini*	
	2	*tu ver fini*		*voi veţ fini*	
	3	*je vå fini*		*čelj vå fini*	

        'I / you / he / we / you / they will finish'

What we immediately note is that apparently the third person perfect auxiliary ('have') has a special behavior. This special behavior has also been encountered in spontaneous data, such as the following sentence: [18]

(50) *Vaca durmit-a tota nopta.*
      cow-the slept have.3Sg all night-the
      'The cow slept all night long'.

An interesting syntactic phenomenon that contrasts IR to Romanian is verb ellipsis. Romanian does not allow for verb ellipsis only, but the entire inflection phrase (thus including the auxiliary which bears tense and agreement features) can be elided. IR, however, displays VP-ellipsis.

(51) a. *Če-ai čiteit libru?*   *Am.*
          Q have.2Sg read book-the?   have.2Sg
          'Have you read the book? I have.'
      b. *Ai citit cartea?*   *Am citit-o. / *Am.* (Romanian)
          have.2Sg read book-the   have read it.Cl.Fem / have.2Sg

Moreover, the auxiliary can be disjoint from the verb in IR, whereas in Romanian the adjacency is a strict requirement (see Chapter 5, Section 5.3.1.).

The examples below illustrate different types of constituents which can intervene between the perfect auxiliary and the participle. All examples are taken from Sârbu (1992).

(52) *Aux >Subj >PastPart*
      a. *Cum åm io cuvintåt, av şi ie, Leca Morariu şi mul'åra.*
          how have.1Sg I spoken, have3.Pl also them, L. M. and wife-the
          'The manner in which I spoke, they also did, L.M. and his wife.'

---

[18] We will say more about this in a Chapter 5.

b.   *Mâre voi io učide.*

tomorrow will.1Sg I kill.

'Tomorrow I'll kill.'

(53)   *Aux >Obj >PastPart*

Ân cârca cu bręnta am åpa purtåt.

in back with bucket-the have.1Sg water brought

'I carried the watter on my back with the bucket.'

(54)   *Aux >Adv> PastPart*

a.   *Io-l voi acmo učide.*

I him.Cl will.1Sg now kill

'I will kill him now.'

b.   *Nu ştivu dupa câta vręme åv jos verit.*

not know.1Sg after how much time have.3Pl down came

'I don't know after how long they came down.'

(55)   *Aux >PP> PastPart*

a.   *Când am de mic fost…*

when have.1Sg of little been…

'When I was little…'

b.   *Pac s-åv cu tractoru učis.*

then *refl* have.3Pl with tractor-the killed

'Then they got killed by (means of) a tractor.'

Some examples have also emerged spontaneously in my own data. They are given bellow:

(56)   *Aux >Obj>PastPart*

Io-m tot vezut.

I have.1Sg all seen

'I saw it all.'

(57)   *Aux >Adv >PastPart*

    a.   *Tu-i   sigurno pročiteit čuda libri.*

        you have.2Sg   certainly read   many books

        'You have certainly read many books.'

    b.   *Ier-am   mušat cantat.*

        yesterday have.1Sg beautifully sung

        'Yesterday I sang beautifully.'

    c.   *Io-am   tota nopta durmit.*

        I have.1Sg all night-the slept

        'I slept all night long.'

4.3.4. Tense and aspect

Verbs in IR express the following categories morphologically or analytically:

    A. Tense: Present, Past, Future;

    B. Aspect: Perfective, Imperfective, Iterative;

    C. Mood: Indicative, Restrictive, Imperative, Subjunctive;

    D. Voice: Active, Passive;

    E. Person: 1st, 2nd, 3rd;

    F. Number: Singular, Plural.

While we have already briefly dealt with Voice in Section 4.2.1., Person and Number morphology has been illustrated extensively in the examples.

   As for Tense and Mood, Table 1 sketches the combinations of tense and mood markers, where "+" stands for bound morphemes (suffixes), and "&" stands for marking of person, number and mood on an unbound morpheme (the auxiliary), while the lexical verb surfaces either in its base form (infinitive) or as a participle.[19]

---

[19] "/" marks in the table the absence of future imperative and future perfect forms.

	Indicative	Restrictive	Imperative
Future	+	&	/
Present	&	+	&
Perfect (Past)	+	+	/

Table 1: Tense and Mood markers

The indicative present tense is formed synthetically with specific markers for person and number. Kovačec (1971: 136-137) notes that Old Romanian verbs (i.e. with Romanian etymology) are no longer productive. The word accent for these verbs falls on the root for all singular persons and 3rd person plural, and on the person and number marker for 1st and 2nd person plural. Productive verbs, on the other hand, (mostly of Slavic origin) always display the accent on the person and number marker, never on the root. Moreover, consonant and vowel alternation are only to be found within the non-productive class, never within the productive class, the members of which are borrowed either "frozen", as such.

Aspect is a category which has been highly subject to contact. IR has borrowed from Croatian the aspectual distinctions perfective-imperfective-iterative. Prefixation with Slavic aspectual morphemes on verbs of Romance origin is not rare, but it cannot be considered to be productive. This property distinguishes IR from Romanian, Aromanian and Megleno-Romanian. Some examples are: [20]

(58)　(imperf.) *durmi*　　/ (perf.) *zadurmi*　　'to (fall) sleep'

　　　　　mâncå　　/　　*namâncå*　　'to eat (up)'

The perfective-imperfective opposition can also be expressed by different conjugations, as exemplified below. Often, the meaning of the verbs may change:

(59)　(imperf.) *strili*　　/ (perf.) *strilęi*　　'to shoot / to shoot with a rifle'

　　　　　spovedi　　/　　*spovidęi*　　'to (finish) confess(ing)'

---

[20] Cf. Kovačec (1971:125) for more examples.

As for the imperfective-iterative opposition, the latter is to be considered morphologically marked with respect to the imperfective. Patterns usually emerge for Romance lexical entries, the two verbs in such an aspectual pair belonging to different conjugations. A couple of examples are provided:

(60) (imperf.) *cumpara* / (iter.) *cumparavęi* 'to buy'

　　　　　veri /　　　　　　*verivui*　　　'to arrive'

We must add that for other verbs of Latin origin, a suppletive Slavic form is used to express perfective or iterative. We provide a couple of cases:

(61) (imperf.) *bę*　　　 / (perf.) *popi*　　　'to drink (up)'

　　　　　torče　　/　　　*spredi*　　　'to (finish) spin(ning)'

By means of elicitated sentences in appropriate contexts, we have determined that IR does not use the imperfect for the 'reference time' event when two events are simultaneous.[21] It uses past tense for both events:

(62) *Cân -a ia verit, io-m　　čiteit.*
　　　when has she come, I have1SG read
　　　'I was reading when she came'.

---

[21] As for progressive, Kovačec (1971:123-4) states that it can be encountered, but it is very rare. My informants did not produce any elicitated example with the progressive. Instead they produced examples parallel to Romanian:

　　i.　　*Io sem trudna ši nu lucru nič.*
　　　　　I am tired and not work1.Sg nothing
　　ii.　　*Sunt obosit şi nu fac nimic.*　　　　　(Rom.)
　　　　　I am tired and not do.1Sg nothing
　　iii.　*Io sono stanco e non sto facendo nulla.*　　(Ita.)x
　　　　　I am tired and not be.1Sg doing nothing
　　　　　'I am tired and I'm not doing anything.' (context: now, not generically)

The Slavic-like imperfective marking and the lack of the function of the imperfect aspect as in Romance (cf. Romanian and Italian parallel examples below, but also the English translation) are likely not to be accidental.

(63)   a.   *Când a venit ea, eu citeam.*   (Rom.)
                when have.3Sg come she, I read.Imperf

        b.   *Quando è arrivata lei, io stavo leggendo.*   (Ita.)
                when be.3Sg come she, I be.Imperf reading

                'When she arrived, I was reading.'

4.3.5. Negation

There are two negative elements in IR: *nu* (and *nu-*, its phonologically reduced form) and *ne*. *Nu* expresses sentence negation, whereas *ne* negates constituents. *Nu* is a syntactic clitic (on a verbal element, be it auxiliary or the lexical verb), whereas *nu-* is a a syntactic and phonological clitic.

(64)   *(Io) nu lucru.*   */(Io) nu-am lucråt*
        I not work.1Sg   / (I) not have.1Sg worked
        'I don't work'   / I didn't work.

(65)   *Ier citeit-nu-am nič.*
        yesterday read-not-have.1Sg nothing
        'I didn't read anything yesterday.'

(66)   *Datu-mi-a*   *regalu, ne boca de vir.*
        Given me have.3sg   gift-the, not bottle-the of wine
        '(S)he gave me the gift, not the bottle of wine'

As a short answer to a Yes/No question *ne* is used:

(67) *Ce ploià? Ne. / Nu ploià.*

Q rains? no / not rains

'Is it raining? No (it's not)/ It isn't raining'.

Like Romanian, IR is a strict negative concord language: if one (or more) negative elements (negative polarity items) are present in a clause, sentence negation is obligatory. The difference between IR and Romanian, on the one hand, and Italian (a negative concord language), on the other hand, is that for the former, but not for the latter, a negative polarity item in subject position (preceding the verb) also requires sentence negation.

(68) *Ničur paninu nu-a poidit.*   (negative subject)

nobody sandwich-the not have.3Sg eaten

'Nobody ate the sandwich.'

Neanche ur panin nu-a poidit.   (negative object)

not.even a sandwich not have.3Sg eaten

'Not even a sandwich did (s)he eat.'

(69) *Åze nu lucra ničur.*

today not work.3Pl nobody.Pl

'Nobody works today'

(70) *Ničur *(nu) nič lucra.*

nobodyPl not nothing work.3Pl

'Nobody does anything'

(71) a.   *Nimeni *(nu) face nimic.*   (Rom.)

b.   *Nessuno (*non) fa nulla.*   (Ita.)

nobody not do.3Sg nothing

'Nobody does anything.'

Whether the negative object surfaces linearly pre- or postverbally does not make any difference for the requirement of sentence negation:

(72)   *{Nič} *(nu)-a {nič} čiteit.*

      {nothing} not have.3Sg {nothing} read

      '(S)he didn't read anything.'

Adverbial negative items have also been attested in the same context with other negative polarity items, such as in the example below:

(73)   *Tu nu nicad nič lucri.*

      you not never nothing work.2Sg

      'You never do anything.'

4.4. Ad interim conclusions

This chapter has been dedicated to an overall presentation of the verb phrase in Istro-Romanian. As with the noun phrase in Chapter 3, we first illustrated some classes of verbs based on argument structure (transitives, intransitives and unaccusatives, psychological verbs, reflexives and meteorological verbs), after which we dealt with data corresponding to the changes to argument structure (the passive voice, the middle and the impersonal constructions).

The second part of the present chapter took into account semi-lexical and functional verbs (modals, aspectual verbs and auxiliaries), and the realization of functional categories related to this category (tense, aspect and negation).

Chapter 5

Clitic placement and some syntactic implications

5.1. Types of clitics in Istro-Romanian. Data

Most of the data on clitic elements in Istro-Romanian was collected in a fieldwork trip in January 2010, mainly through elicitation and grammaticality judgements. In doing so, we followed the *Clitic Questionnaire* (created Riet Vos, Univ. Tilburg, e Ludmila Veselovská, Univ. Palackého Olomouc, as part of a NWO project *Clitics and UG: Bibliography and Questionnaire*), which has been applied to a fair number of languages between 1991 and 1995. The Questionnaire is available on the website of Max Planck Institute for Evolutionary Anthropology, Department of Linguistics, (http://www.eva.mpg.de/lingua/tools-at-lingboard/questionnaires.php), section Typological Tools for Field Linguistics, which is where we have downloaded it from.[1]

Two notes are in order: the *Clitic Questionnaire* was meant for informants who are also linguists, which was obviously not the case for Istro-Romanian. Moreover, I myself at that time had little knowledge of the IR grammar, thus felt the need to adapt: in certain circumstances, it was nearly impossible to answer a question (either because it was irrelevant for the IR, or because though being relevant, we lacked other data to make us see it).

The Questionnaire had already been answered both for Romanian (by Sandra Rîpeanu and Carmen Dobrovie-Sorin) and Serbo-Croatian (by Wayles Browne and Olga Tomić).[2] We will thus make reference to these two languages for comparative or contrastive data.

A note on ortography: I have chosen to employ the use of the hyphen to mark clitic relations (whether or not there is vowel reduction), corresponding to the Romanian ortography.

---

[1] http://www.eva.mpg.de/lingua/tools-at-lingboard/questionnaire/clitics_description.php

[2] The data on Italian was provided by Anna Cardinaletti and Denis Delfitto.

5.1.1. Some general properties of IR

To begin with, we will provide the data conforming to the initial part of the *Clitic Questionnaire*: General Questions.

1.1. The unmarked word order is VO:

(1)   *Redento čiteit-a libru.*

        Redento read have.3Sg book.the

        'Redento read the book.'

1.2. The language has prepositions:

(2)   *Mes-am ân besęrca cu fečoru.*   (*ân* spatial P, *cu* functional P)

        gone have.1Sg in church with boy-the

        'I went to church with my son.'

1.3. IR is a pro-drop language:

(3)   *Fost-a bolân / bolna.*

        been has.3Sg ill.MascSg/Fem.Sg

        'He was ill./She was ill.'

Table 1 summarizes the answers. For comparative reasons, Romanian and Serbo-Croatian (as per the original survey) are provided.

In the tables, the symbols are used as follows: [+] means that the language displays the property under examination, while [-] is used if it lacks the property. Unless otherwise specified (i.e. *1.2.* below), the [+/-] sign marks that the language may or may not have a certain property depending on the (syntactic) context. Further, [i] indicates that a certain property is irrelevant in the language (generally because it is related to a phenomenon which the language lacks, as it has been pointed out in a previous question), while [?] corresponds to either inconclusive or total lack of data.

	IR	**Rom.**	**SCr**
1.1. The unmarked word order is [+] VO, [-] OV	+	+	+
1.2. The language has [+] prepositions, [-] postpositions	+	+	+/-
1.3. Pro-drop [+] or [-] non-pro-drop language	+	+	+

Table 1. General Questions

5.1.2. Properties of clitic elements

We will now illustrate the properties of clitic elements, along the lines of sections 2-4 of the *Clitic Questionnaire* (Categories, Properties, Position: host clitics, fixed position clitics). We must underline again that it was not possible to find answers to all questions.[3] The data marked with *i* is considered to be irrelevant for the language in question.

5.1.2.1. Categories

Section 2 of the questionnaire deals with the categorial status of clitic elements.

2.1. Does your language have different series of weak and strong pronouns? [+] yes, [-] no

(4)   *Frane *miie / mi   -a dåt ur libru.*

     Frane   me.Dat /me.DatCL   has given a book

     'Frane gave me a book.'

(5)   *Dåt-a   miie   /*mi   ur libru.*

     given has me.Dat /*me.DatCL a book-the

     '(S)he gave me a book.'

2.2. If you answered (2.1) positively: can a weak pronoun function as: (a) subject, (b) indirect object, (c) direct object, (d) reflexive, (e) possessive. [+] yes, [-] no.

(6)   a. *Se ganę ke va Redento čitei čuda libri.*

      SE say that will.3Sg Redento read many books

      'It is said that Redento/he will read many books'

    b. *Frane ţi-a dåt libru.*

      Frane you.DatCL gave book-the

      'Frane gave you the book'

    c. *Frane m-a vezut.*

      Frane me.AccCL seen

---

[3] Section 5 of the *Clitic Questionnaire* dealt with syntactic contexts involving clitic elements which may or may not allow for movement operations. At the stage when I collected the data, it would have been very difficult to test for those contexts, with the risk of obtaining inaccurate data. We have thus preferred not to force our informants into giving judgements to highly complex elicitated data.

'Frane saw me.'

d. *Frane se spelavę.*

Frane *refl* wash.3Sg

'F washes himself'

2.4. Does your language have verbal clitics, like Czech být (be)? [+] yes, [-] no.

The auxiliary of indicative perfect can be proclitic or enclitic. The future marker is either free, or enclitic on the lexical verb.[4] The copula also has a clitic form. When two or more (verbal and/or pronominal) clitics are present, they cluster.

(7)　　a. *Io vezut-am...*

　　　　I seen have.1Sg

　　　　'I saw…'

　　　　b. *Io nu-am vezut…*

　　　　I not have.1Sg seen

　　　　'I didn't see…'

　　　　c. *Io l　　-am　　vezut.*

　　　　I him.AccCL have1SG seen

　　　　'I say him.'

(8)　　a. *Io cântå-voi.*

　　　　I sing will.1Sg

　　　　b. *Io voi mâre cântå.　　/ Mâre vor filj veri.*

　　　　I will.1Sg tomorrow sing / tomorrow will sons-the come

　　　　c. *Io-i mâre cântå.*

　　　　I will.1Sg tomorrow sing

(9)　　*Je-i mušat, nu-i grum.*

　　　he is handsome, not is ugly

　　　'He's handsome, not ugly.'

2.5. If you answered *2.4.* positively:

a) In which tense(s) may the verbal CL appear?

b) Which person features may the verbal CL possess?

---

[4] More data, including some patterns and asymmetries, will be given in Section 5.2.2.

As already illustrated, the auxiliary clitics can be encountered in the indicative perfect and paradigms.[5]

Verbal clitics appear in all persons and numbers. However, some asymmetries are observed. They will be discussed in Section 5.2.2.

2.6. Does your language have adverbial clitics, like Dutch *er*? [+] yes, [-] no.

We did not encounter adverbial clitics.

2.7. If you answered *2.6.* positively, are there (pro) nominal clitics in your language which have the same morphophonological form?

Irrelevant as far as the Istro-Romanian data at our disposal is concerned.

2.8. Does your language have adjectival clitics (APs)? [+] yes, [-] no.

If we are to interpret this question as it has been interpreted for Italian, then we could answer positively.

Denis Delfitto comments: "there are adjectival CL demonstratives in a substandard variety with a lexicalized equivalent phenomena in the standard language":

(10)   <subst.>   *sta casa, sto tipo*

                      this house this guy

        <stand.>   *stavolta, stamani*

                      this time this morning

Some (but not all) IR lexical entries are parallel to the Italian above. The same is true for Romanian though, which has been classified as a language without adjectival clitics by the respondents to the questionnaire.

Note that of the two examples below, one follows the Romanian pattern (in b), while the second one is likely to be a borrowing (or calque) from Italian.[6]

(11)   a. *åsta sęra,*   *stanopte*

       b. *astă seară,*   *astă noapte*   (Rom.)

---

[5] In traditional grammars, IR conditional is referred to as "restrictive mood". We have not elicitated for this mood, nor is it found in our spontaneous data.

[6] The comment is referred to the adjectival clitic *sta*, not to the lexical entry for 'night', which is an IR word.

this evening   this night

2.9. Does your language have prepositional clitics, like French *en* or *y* ? [+] yes, [-] no.

From the data, it appears that the answer is negative. However, since it was very difficult to elicit such constructions, we leave the answer open.

2.10. Does your language have particles functioning as clitics? [+] yes, [-] no.

a) negation PRT, b) question PRT, c) emphatic PRT, d) other e.g.

a) the negation is either a syntactic clitic, appearing before all other clitic elements in the cluster, or both a syntactic and phonological (pro)clitic, as illustrated below.

(12)   a. *Nu mi   -a verit listu   di la ničur.*

             not meDatCL   has come letter-the from nobody

             'I didn't receive the letter from anyone.'

        b. *Io nu-am vezut…*

             I not have1Sg seen

             'I didn't see.'

b) IR has a question particle *če* obligatory for wide scope Yes/No questions. It always appears in sentence-initial position, immediately followed by the other clitics, when present.

(13)   a. *Če-ăi fost ân čine?*

             PRT have2SG been in cinema

             'Have you been to the cinema?'

        b. *Če-l vezi?*

             PRT him.CL see.2Sg

             'Do you see him?'

        c. *Če nu vo vezi?*

             PRT her.CL see.2Sg

             'Don't you see her?'

We have no data as for emphatic particles or other types of particles.

    Table 2 gives the results for IR, alongside with Romanian and Serbo-Croatian.

	IR	Rom.	SCr
2.1. The language [+] has or [-] has not different series of weak and strong pronouns [7]	+	+	+
2.2.a. A weak pronoun can [+] or cannot[-] function as: SUBJ	+	-	-
2.2.b. A weak pronoun can [+] or cannot[-] function as: IO	+	+	+
2.2.c. A weak pronoun can [+] or cannot[-] function as: DO	+	+	+
2.2.d. A weak pronoun can [+] or cannot[-] function as: reflexive	+	+	+
2.2.e A weak pronoun can [+] or cannot[-] function as: possessive	?	+	+
2.3 The language has[+] or has not[-] nominal weak pronouns	?(-)	-	-
2.4. The language has [+] or has not [-] verbal clitics.	+	+/-	+
2.5.a. The verbal CL shows [+] or does not show [-]: tense features	+	i	+
2.5.b. The verbal CL shows [+] or does not show [-]: person features	+	i	+
2.6. The language has [+] or has not [-] adverbial clitics	?	-	-
2.7. CLs have [+] or have not [-] the same morphophonological form	?/i	i	i
2.8. The language has [+] or has not [-] adjectival clitics	+	-	-
2.9. The language has [+] or has not [-] prepositional CLs	-/?	-	-
2.10.a. The language has [+] or has not [-] particle CLs: negation PRT	+	+	+
2.10.b. The language has [+] or has not [-] particle CLs: question PRT	+	-	+
2.10.c. The language has [+] or has not [-] particle CLs: emphatic PRT	?	+	-
2.10.d. The language has [+] or has not [-] particle CLs: other	?	+	-

Table 2. Categories

---

[7] At the time the questionnaire was made, the tripartite distinction between clitic, weak and strong pronouns (Cardinaletti & Starke 1999) had not been clearly made. Thus, the terms "clitic" and "weak" are intended as referring to the same syntactic elements. We have chosen not to alter the original formulation of the questions, and keep both terminological items.

5.1.2.2. Properties

Moving on to Section 3 of the *Clitic Questionnaire*, we have inquired for properties of clitics, such as order and doubling.

3.1. The order of full NPs is [+] DAT - ACC or [-] ACC - DAT.

Just like Romanian and Serbo-Croatian, IR allows both word orders. It is a more delicate matter to establish which of the two orders is marked, since various factors may come into play (such as verb type, animacy, definiteness, information structure, etc.). We elicitated for following sentences in contexts involving contrastive focus.

(14) a. *Gianni dåt-a   listu   lu Mario (ne libru).*   ACC>DAT

       Gianni given -has letter-the to Mario (not book-the)

   b. *Gianni dåt-a lu Mario listu (ne lu Redento).*   DAT>ACC

       Gianni given -has to Mario letter-the (not book-the)

       'Gianni gave the letter to Mario (, not the book/not to Redento).'

3.2. CLs appear [+] or not [-] in the same order as NPs.

The clitic pronouns have a rigid order, DAT > ACC:

(15)   *Mi l-a dåt.*

       me.Dat.CL it.Acc.CL given

       'He gave it to me.'

3.3. CLs appear [+] or not [-] in conjunction.

No, clitic pronouns cannot appear in conjunction.

(16)   **mi ši ți l-a dåt.*

       me.Dat.CL and you.Dat.CL it.Acc.CL given

3.4. CLs appear [+] or not [-] in an environment of stress.

Just like the property in *3.3.* above, the ban on stress environments is a defining property of clitic elements.

(17)   a. ***MI l-a dåt, ne ţi.*

me.Dat.CL it.Acc.CL given you.Dat.CL

'(intended) He gave it TO ME, not to you.'

3.5. CLs appear [+] or not [-] after negation.

Yes, clitic elements (either pronominal or verbal, i.e. auxiliaries) can appear after negation. We offer two different orders involving a gender asymmetry with respect to the pronominal clitic:

(18)   a. *Io nu l-am vezut.*   Neg > CLmasc > AuxCL > PastPart

I not him.CL have.1Sg seen

b. *Io nu-am vo vezut.*   Neg > AuxCL > CLfem > PastPart

I not have.1Sg her.CL seen

'I haven't seen him/her.'

3.6. CLs appear [+] or not [-] in isolation.

Again, a defining property of clitics. The data confirms the expectations:

(19)   *Čire-ai vezut?*   **L. / *Vo.*

whom have.2Sg seen him.Acc.CL / her.Acc.CL

'Whom did you see? (intended) Him /Her.'

3.7. CLs can [+] or cannot [-] be modified.

We do not have a judgement on this property, which is however still a defining property of clitic elements, thus we expect the answer to be negative.

3.8. SUBJ CLs appear [+] or not [-] in a sentence initial topic position.

As seen above, impersonal *se* is a subject clitic, and it can appear sentence initially, but it is not a topic.

(20)   a. *Se ganę ke va Redento čitei čuda libri.*

SE say that will.3Sg Redento read many books

'It is said that Redento/he will read many books'

3.9. OBJ CLs appear [+] or not [-] in a sentence initial topic position.

We could not obtain a clear context. It is still very likely that a clitic element cannot be topicalized.

3.10. CL can[+] or cannot [-] be related to a quantifier.

In this respect, pronominal clitics are not different from nominal elements. The quantifier can be stranded in different positions, as shown by the paranthesis below:

(21) *Maria {toţ} lj   -a   {toţ} vezut {toţ}.*

     Maria {all} them.CL has {all} seen {all}

     'Maria saw them all.'

3.11. CLs allow [+] or not [-] a parasitic gap.

Differently from both Romanian and Serbo-Croatian, [8] IR seems to allow for parasitic gaps in relation to clitic objects. The relevant example (obtained by elicitation) is given below:

(22) Je -l   morę pure   __   ân cašeta de pošta, ši nu pure   __   pre scănd.

    he itCL must put.inf (GAP) in box of mail,   and not put.inf (P.GAP) on table.

    'He must put it into the mailbox and not put it on the table'

3.12. The language has [+] or has not [-] CL doubling.

A peculiarity of Romanian (and Spanish) is that clitic doubling is related to the animacy of the doubled object, which is introduced by a functional preposition (Romanian *pe*, Spanish *a*). In these two Romance languages, all animate direct objects, while in at least another Balkan language, namely standard Macedonian, doubling is obligatory irrespectively of animacy.

    This requirement seems to be optional in IR. It is worth mentioning that Accusative objects are not introduced by a functional preposition in IR.

(23)   a. *Io vedu   (*-l) mę frăte.*

       I see.1Sg (*himCL) my brother

---

[8] Italian does not allow for parasitic gaps related to clitics either.

b. *Vedu-l.   / Vedu-vo.*

    see.1Sg himCL /herCL

(24)   a. *Nu vedu   već zeče åń Livija.*

    not see.1Sg since ten years Livija

    'I haven't seen Livija in ten years.'

    b. *Nu   -am   (vo)   vezut Livjia već zeče åń.*

    not have.1Sg herCL seen L.   since ten years

    'I haven't seen Livija in ten years.'

My informants only allowed for the possibility of clitic doubling in (24.b). The two sentences above form a minimal pair. Apart from the compound tense in the second clause, the difference resides in the scrambled order between the adverbial adjunct and the direct object *Livija*. Still, the (24.b) example does not require a doubling clitic, but it allows it to appear.

    As for indirect objects, doubling is completely disallowed, whereas it is obligatory in Romanian. Two sets of data are provided. In (), the indirect object is a proper noun. In (25), it is a strong pronoun. The ban on doubling is observed in both contexts.

(25)   a. *Maria {*lj} dåie {*lj} libru lu Gianni.*

    Maria {he.DatCL} gives {he.DatCL} book-the to Gianni

    'Maria gives the book to Gianni.'

    b. *Maria dåie-lj libru.*

    Maria gives he.DatCL book-the

    'Maria gives him the book.'

(26)   a. *G. {miie} dåt-a {miie} libru.*

    G {me.Dat} given has {meDat} book-the

    b. *G. mi   -a   dåt libru.*

    G me.DatCL has given book-the

    c. **G.{miie} mi-a dåt {miie} libru.*

    G {me.Dat} meDatCL has given {meDat} book-the

3.13. (a) nominal,(b) pronominal,(c) prepositional,(d) verbal,(e) adverbial,(f) possessive CLs can [+] or cannot [-] double.

As seen above, the data at our hand indicates that pronominal clitics can optionally double.

We do not have data for the possessive clitic, while the other types are unlikely to double (just like in all other languages taken into account by the survey).

3.14. The double of the CL is (a) pronoun, (b) full definite NP, (c) full indefinite NP,(d) adverbial CL.

The data presented for the question *3.12* allows us to respond affirmatively to one of the categories listed, namely a full definite NP. Again, doubling in such contexts is optional.

3.15. CL function is: (a) SUBJ,(b) IO,(c) DO,(d) prepositional OBJ.

The only doubling clitic element is the direct object.

3.16. The CL can [+] or cannot [-] be present when questioned.

We have seen that the indirect object cannot be doubled. Interestingly, if questioned, a clitic element can optionally be present:

(27)　　a. *Lu cui dåt-a libru?*

　　　　　to whom given has book-the?

　　　　b. *Lu cui lj-a dåt libru?*

　　　　　to whom he.DatCL has given book-the

　　　　　'Whom did he give the book to?'

We do not have any data as for the direct object.

Table 3 shows the comparative answers for IR, Romanian and Serbo-Croatian.

	IR	Rom.	SCr.
3.1. The order of full NPs is [+] DAT - ACC or [-] ACC - DAT	+/-	+/-	+/-
3.2. CLs appear [+] or not [-] in the same order as NPs	-	-	+/-
3.3. CLs appear [+] or not [-] in conjunction	-	-	-
3.4. CLs appear [+] or not [-] in an environment of stress	-	-	-
3.5. CLs appear [+] or not [-] after negation	+	+/-	-
3.6. CLs appear [+] or not [-] in isolation	-	-	-
3.7. CLs can [+] or cannot [-] be modified	-	-	+/-
3.8. SUBJ CLs appear [+] or not [-] in a sentence initial	-	-	-

position			
3.9. OBJ CLs appear [+] or not [-] in a sentence initial position	?/-	-	-
3.10. CL can[+] or cannot [-] be related to a quantifier	+	+	+
3.11. CLs allow [+] or not [-] a parasitic gap	+	-	-
3.12. The language has [+] or has not [-] CL doubling	+/-	+	-
3.13.b Pronominal CLs can [+] or cannot [-] double	+	+	i
3.13.f Possessive CLs can [+] or cannot [-] double [9]	?	-	i
3.14.a The double of the CL is (a) pronoun	-	i	i
3.14.b The double of the CL is (b) full definite NP	+/-	i	i
3.14.c The double of the CL is (c) full indefinite NP	-/?	i	i
3.15.a CL function is: (a) SUBJ	-	-	i
3.15.b CL function is: (a) IO	-	+	i
3.15.c CL function is: (a) DO	+	+	i
3.15.d CL function is: (a) prepositional OBJ	-	-	i
3.16. The CL can [+] or cannot [-] be present when questioned	+/-	+	i

Table 3. Properties

5.1.2.3. Position

This part of the survey covers three topics: A. whether clitics attach to a host or not, B. the category if the host and the properties deriving from it, and C. fixed position clitics.

Subsection A involves one question only, which was mentioned above. The need for a host to attach to is a universal requirement for clitics, thus IR does have clitic elements that attach to hosts.

Let us now see the data corresponding to the questions in Subsection B.

4.B.1. The host is (a) verb,(b) noun, (c) adjective, (d) preposition, (e) complementizer, (f) adverb, (g) other.

We will provide examples both for auxiliary (or copula) clitics, and pronominal clitics. Prepositions cannot function as hosts for neither of the two, while all the other categories

---

[9] The other parts of question *3.13* (a) nominal, (c) prepositional,(d) verbal,(e) adverbial were considered irrelevant for all the languages surveyed.

have been documented as possible hosts, with the exception of pronominal cliticising on adjectives. Question particles are exemplified for (g).

(28)   a. *Dejan finit-a   lucră,   ši   va veri   ånca.*(aux. clitic)

Dejan finished =have.3Sg work.inf, and   will.3Sg come.inf here

'Dejan finished working, and he'll come here.'

a'. *Io l-am putut vedę   / Putut-l-am vedę.*   (pron. clitic)

I he.Acc.CL have.1Sg could see.inf   could he.Acc.CL have.1Sg see.inf

'I could see him.'

b. *Redento-i ur måre   om.*   (cop. clitic)

Redento =is a great man

b'. *Livija –l pote vedę.*   (pron. clitic)

Livija he.Acc.CL can see.inf

'Livija can see him.'

c. *Mušåta-i feata.*   (cop. clitic)

beautiful =is girl-the

'The girl is beautiful.'

c'. **Locu-i mušåt ej prope. / Lui locu mušåt ej prope.*   (pron. clitic)[10]

place =his beautiful is nearby. / his place beautiful is nearby

'His beautiful place is nearby.'

d, d'. *

e. *Se ziče k   -a   Redento {*a}   čuda čiteit.*   (aux. clitic)

SE say that =have3Sg R much have.3Sg   much read

'It is said that Redento read a lot.'

e'. *Misles   ke -l vedu.*   (pron. clitic)

believe.1Sg that he.Acc.CL see.1Sg

f. *Ier-a clemat Nela.*   (aux. clitic)

yesterday =have.3Sg called Nela

'Nela called yesterday.'

f'. *

---

[10] Cf. Romanian:

i.   *frumoasa-i ţară*

beautiful-the =his country

'his beautiful country'

g. *Če-i mę česta zelen mižul? Je.*   (cop. clitic)

   Q =is mine this green glass? is

   'Is this green glass mine? It is.'

g.' *Če-l vez?*   (pron. clitic)

   Q he.Acc.CL see.2Sg

   'Do you see him?'

4.B.2. CL precedes [+] or follows [-] its host.

In IR, both single clitics and clusters can either precede or follow their hosts. The minimal pair below includes one Dative and one Accusative pronominal clitic, and a perfect auxiliary. The order of the clitic elements is invariable, whether the cluster is enclitic or proclitic on its syntactic host.

(29)   a. *Dåt -ţi   -l   -am.*

       given you.Dat it.CL have.1Sg

  b. *Ţi   l   -am   dåt.*

    you.Dat it.CL have.1Sg given

    'I gave it to you.'

4.B.3. CL is [+] or is not [-] incorporated in its host.

No phenomenon of the type encountered in Spanish or Portuguese *mesoclisis* seems to be present in IR.

4.B.4. CLs are [+] or are not [-] at the same side of the host.

This question refers to whether or not more clitic elements in the same clause must cluster together, and thus appear on the same side of the host. Like in Romanian and Serbo-Croatian, IR clitics cluster. See the example in *4.B.2* above.

4.B.5/6. CLs have [+] or have not [-] fixed order.

Again, see the example in *4.B.2.*

In answer to *4.B.6* (which is the order in the cluster), the orders for pronominal clusters and pronominal + verbal clusters are given below (for compound tenses). While we will dwell more on the different positions of the Accusative pronominal clitics in Section 5.2.3.1., we provide examples for the three orders in (30.b).

145

(30)   a. DAT > ACC

    b. COMP >   **AUX** > SUBJ (> XP) > PAST PART

      COMP > SUBJ > **AUX** > PAST PART (> XP)

      *COMP > PAST PART > **AUX** (> SUBJ) (> XP)

    c. COMP > NEG > CL.Dat. > CL.Acc.masc > AUX > CL.Acc.fem > PAST PART

(31)   *Misles k-am*   *(io) ier*   *fos ân Umag.*

    believe.1Sg that =have.1Sg (I) yesterday   been to Umag

      /ke io-m   *fos*   *ier ân Umag.*

      that I =have.1Sg   been   yesterday to Umag

      */ * ke fost-am*   *ier ân Umag.*

      that been =have.1Sg yesterday to Umag

      'I believe I went to Umag yesterday.'

4.B.7. Function influences [+] or not [-] position in a cluster.

Case, for instance, is relevant (DAT > ACC). Also, gender (MASC > FEM), and category (pronoun, verb).

4.B.8. Case restricts [+] or not [-] position in a cluster.

As already mentioned and exemplified, case is a relevant feature in determining the position of a clitic in the cluster.

4.B.9. REFL CL in the CL cluster is (a) initial, (b) internal, (c) final), (d) depends upon its grammatical function.

Form the data at our hand, it appears that the reflexice clitic precedes other pronominal clitics.

(32)   a. *Gianni cumparåt-a ur mobitel. Gianni ši l-a cumparåt.*

    Gianni bought =has a mobile.phone. Gianni REFL =it.Acc.Masc =has bought

    'Gianni bought a mobile phone. Gianni bought it for himself.'

    b. *Gianni cumparåt-a ura jaketa. Gianni ši vo cumparåt-a.*

    Gianni bought =has a jacket. Gianni REFL =it.Acc.Fem bought =has

    'Gianni bought a jacket. Gianni bought it for himself.'

CL attached to a verb is pre [+] or post [-] verbal in:

4.B.10. - a declarative main clause.

While the masculine pronominal clitic has a fixed position, clustering with the auxiliary, its feminine counterpart may surface in two different positions, pre- and post-verbally, independently from the auxiliary, which in this case obligatorily enciliticizes on the past participle.

(33) a. *Gianni l-a vezut.*   */ Gianni vezut-l-a.*

Gianni he.Acc.CL = has seen   / Gianni seen he.Acc.CL =has

b. *Gianni {vo} vezut-a {vo}.*

Gianni {her.Acc.CL} seen = has {her.Acc.CL}

'Gianni saw him/her.'

4.B.11. - a YES/NO question.

Pronominal clitics immediately follow the question particle *če*, thus preceding the verb.

(34) a. *Če-l vezi?*

Q he.Acc.CL see.2Sg

b. *Če vo vezi?*

Q her.Acc.CL see2SG

'Do you see him/her?'

(35) *Če l-åi čiteit?* [11]

Q itAcc.Masc.CL =have2Sg read

'Did you read it?'

4.B.12. - a main clause wh-question [12]

(36) *Čire ni l-a dåt?*

who we.Dat.CL it.Acc.Masc.CL =has given

'Who gave it to us?'

---

[11] We do not have examples with the feminine accusative clitic for this construction.

[12] *idem.*

(37)   *Cân l-ai čiteit?*

when it.Acc.Masc.CL =have.2Sg read

'When did you read it?'

4.B.13. - a positive imperative.

As for imperatives, IR pronominal clitics behave exactly like their Romanian counterparts, with no gender distinction.

(38) *Mârânca-l!*   *Mârânca-vo!*

eat.2Sg.imp it.Masc.CL   eat.2Sg.imp it.FemCL

'Eat it!'

4.B.14. - a negative imperative.

(39) *Nu-l mâncǎ!*   *Nu vo mâncǎ!*

not it.Masc.CL eat.inf   not it.Fem.CL eat.inf

'Don't eat it!'

Table 4 shows the comparative answers for IR, Romanian and Serbo-Croatian.

	IR	Rom.	SCr.
4.A.1. CL attach [+] or not [-] to a host	+	+	+
4.B.1.a The host is (a) verb	+ ,+	+	+
4.B.1.b The host is (b) noun	+ ,+	+	+
4.B.1.c The host is (c) adjective	+ , ?	+	+
4.B.1.d The host is (d) preposition	- , -	+	-
4.B.1.e The host is (e) complementizer	+, +	-	+
4.B.1.f The host is (f) adverb	+ , -	-	+
4.B.1.g The host is (g) other	+ , +	-	+
4.B.2. CL precedes [+] or follows [-] its host	+/-	+/-	-
4.B.3 CL is [+] or is not [-] incorporated in its host	-	+/-	-
4.B.4 CLs are [+] or are not [-] at the same side of the host	+	+	+
4.B.5/6 CLs have [+] or have not [-] fixed order	+	+	+
4.B.7 Function influences [+] or not [-] position in a cluster	+	+	+/-

4.B.8 Case restricts [+] or not [-] position in a cluster	+	i	+
4.B.9.a REFL CL in the CL cluster is (a) initial	+	-	-
4.B.9.b REFL CL in the CL cluster is (b) internal	-	+	-
4.B.9.c REFL CL in the CL cluster is (c) final	-	-	+
4.B.9.d REFL CL in the CL cluster is (d) depends upon its grammatical function	-	+	-
4.B.10 CL attached to a verb is pre [+] or post [-] verbal in - - a declarative main clause,	+/-	+	+/-
4.B.11 - a YES/NO question	+/?	+	i
4.B.12 - a main clause wh-question	+/?	+	i
4.B..13 - a positive imperative	-	-	+/-
4.B.14 - a negative imperative	+	+	+/-

Table 4. Hosts (B, part I)

The second part of Subsection B deals with the verbal environments which may determine the position of the clitic (proclitic or enclitic on the verbal element).

We have already illustrated in the previous part some of the points in this subsection. Let us go through the quesits.

CL attached to a verb is pre [+] or post [-] verbal in:

4.B.15. In an 'absolute construction' e.g. gerund, infinitive.

We do not have data for these constructions.

4.B.16. In Aux + participle construction a CL precedes [+] or follows [-] (a) Aux, (b) participle.

We have seen (cf. the examples in *4.B.10* above, for instance) that the masculine direct object clitic precedes both the auxiliary and the participle, while its feminine counterpart follows it.

4.B.17. In Aux + gerund construction a CL precedes [+] or follows [-] (a) Aux, (b) gerund.

Like in Romanian, this construction is unavailable in IR.

4.B.18. In Aux + infinitive construction a CL precedes [+] or follows [-] (a) Aux or (b) infinitive.

Although there is a past infinitive in IR, we could not elicit relevant data.

CL attached to a verb is pre [+] or post [-] verbal in:	IR	Rom.	SCr.
4.B.15. In an 'absolute construction' (gerund, infinitive,…)	?	+/-	i
4.B.16.a In Aux + participle construction a CL precedes [+] or follows [-] (a) Aux	+ , -	+	i
4.B.16.b In Aux + participle construction a CL precedes [+] or follows [-] (b) participle.	+ , -	-	i
4.B.17.a . In Aux + gerund construction a CL precedes [+] or follows [-] (a) Aux	i	i	i
4.B.17.b .In Aux + gerund construction a CL precedes [+] or follows [-] (b) gerund	i	i	i
4.B.18.a In Aux + infinitive construction a CL precedes [+] or follows [-] (a) Aux	?	+	i
4.B.18.b In Aux + infinitive construction a CL precedes [+] or follows [-] (b) infinitive.	?	-	i

Table 5. Table 4. Hosts (B, part II)

Subsection C investigates the position of the clitics to the extent it is fixed in the target language. Romanian is clearly not a relevant case study, whereas Serbo-Croatian displays "clitic-second". It is interesting to compare and contrast the syntactic contexts in which IR seems to resemble Serbo-Croatian.

The premise is that although in some constructions, clitic elements (in particular auxiliaries, but also pronominal clitics when in a cluster) may obligatorily surface in the second position (following the first word/constituent), IR does not display a rigid "clitic-second" requirement. See Section 5.2. below for further considerations.

Some of the points in Subsection C have already been illustrated in the previous subsection. We have chosen to mark with *i/+* answers *4.C.1.a-e,* since IR is not strictly speaking a clitic second language. At most, it is a "non-clitic first" language, but not in all contexts, as we will see further on. What is important to bear in mind is that clitic elements cluster together, and clusters can be phonologically independent, as is the cases below, for instance:

(40) a. *Če m-   am   ponešeit bire åz?*   *(V-) Ai.* [13]   */ N-ai.*

    Q REFL.1Sg have.1Sg behaved well today   (*v*) have.2Sg   not =have.2Sg

     'Have I behaved well today? You have. / You haven't.'

  b. *Io n-am nicad vezut divli porc.*

    I not =have.1Sg never seen wild pig

     'I've never seen a wild pig.'

Based on the above (but see also the next section), we provide the final table related to the Position part of the *Clitic Questionnaire.*

	IR	Rom.	SCr.
4.C.1.a The 'fixed position' of the CLs in the language is: - sentence initial, behind the first word	i/+	i	+
4.C.1.b - sentence initial, behind the first constituent	i/+	i	+
4.C.1.c - between the complementizer and the subject	i/+	i	+
4.C.1.d - between the subject and before sentence adverbs	i/+	i	-
4.C.1.e - another	i/+	i	-
4.C.2. CLs are pre [+] or post [-] verbal in: - a declarative main clause	+/-	i	+
4.C.3. - a yes/no question	+	i	+
4.C.4. - a main wh- question	+	i	+
4.C.5. - a positive imperative	-	i	+/-
4.C.6. - a negative imperative	+	i	+/-
4.C.7. - an 'absolute construction'	?	i	+/-
4.C.8.a A CL can [+] or cannot [-] be separated from its adjacent element by (a) subject	-	i	-
4.C.8.b A CL can [+] or cannot [-] be separated from its adjacent element by (b) an adverb	-	i	-

Table 6. Fixed position clitics

---

[13] My two informants disagreed on the necessity of the epenthetic consonant as a phonological support for the second person auxiliary. In fact, one of the informants produced it more often than not, indicating that for her 1st and 2nd person perfect auxiliaries are also phonological clitics, on a par with 3rd person auxiliary. This didn't seem to hold for one of her sisters, who would accept and produce those auxiliaries in isolation with no epenthetic support.

5.2. Clitic placement in Istro-Romanian

5.2.1. Contact-induced phenomenon? Some considerations

Language contact has influenced all modules of the IR grammar, to various degrees. "But this influence manifests itself in different ways, being reflected by numerous borrowings, and also by Istro-Romanian innovations, following some Croatian patterns, and still remaining acquisitions of this idiom, as alloglotic elements, adapted to the Istro-Romanian linguistic system" (Sârbu 1992:33).

The position of verbal and pronominal clitic elements in IR, a well-described phenomenon in Serbo-Croatian, seems to indicate that contact is at stake in certain environments.

Two major types of clitics are currently assumed in the literature: 1) verb-adjacent clitics (found in Romance languages, for instance), and 2) second position or Wackernagel clitics (found in Slavic languages such as Serbo-Croatian, Bulgarian, Czech, or in other unrelated languages such as Cypriot, etc.). There is much debate as to whether the second position is determined by syntax or by interface relations with prosody/phonology (cf. Browne 1974, Progovac 1993, Schütze 1994, Radanović-Kocić 1988, Bošković 2000, 2004, Werle 2009, among many others). Facts are however clearly classified in the literature: there is *2D placement* (in Halpern's 1995 terms), where the element in the first position is a syntactic daughter of the clause in which the clitic appears, and *2W placement*, where the first element is a prosodic word (*Prosodic Inversion*, Halpern 1995). We will illustrate these two types of clitc placement with a minimal pair from (Serbo-)Croatian (Penn 1999:2, 1).

(41)   a. *Taj čovek je video Mariju.*   *2D*
             that man Cl.3Sg saw Mary.Acc
       b. *Taj je čovek video Mariju*   *2W*
             that Cl.3Sg man saw Mary.Acc
             'That man saw Mary'.

This distinction is not perfectly mirrored in the IR system. Although 'clitic second' appears to be a constraint in some environments, IR and SCr second postion clitics do not share all properties.

We have illustrated in the first part of the present chapter the main properties of clitics (whether perfect auxiliary or pronominal clitics) in IR. The syntactic properties are listed

again below, together with the relevant examples, with an eye on whether there is a requirement as for the position in the clause (rather than with respect to the host, as seen so far).

Auxiliary clitics may require a host in initial position in main clauses, which may either be the subject, or the participle, but not both.

(42) a. *(Io-) am   tota   nopta   durmit.
        I   = have.1Sg   all   night   slept
     b. (*Io) Durmit -am   tota nopta.
        slept   have.1Sg all   night
       'I slept all night long'

If two or more clitics co-occur, they surface in a rigid order. The cluster may be sentence-initial, or enclitic on the participle, in which case it seems to be necessary for it to appear in second position.

(43)   *Prezânteit -mi   -l   -aţ   voi,   ne åt.*
      introduced =I.Dat =he.Acc. =have.2Pl you.Pl, not others
      'You introduced him to me, not someone else'.

The 'non-clitic first' requirement that we have seen above for 1st and 2nd person auxiliaries seems to hold for 3rd person auxiliaries as well:

(44)   a. *Livija mi l-a prezânteit.*
        Livija I.Dat =he.Acc =has introduced
     b. *(*Livija) Prezânteit-mi-l-a.*
        Livija introduced I.Dat =he.Acc =has
     c. *Prezânteit-mi-l-a Livija, ne åt.*
       introduced =I.Dat =he.Acc =has Livija, not someone.else
       'Livija introduced him to me (,not someone else.)'

However, while this may be the right observation for non-3rd person perfect auxiliaries (whether singular or plural), it is not the strongest requirement in act for 3rd person.

(45)   *Fečori   sigurno   čuda   čiteit -a.*
        boys-the certainly much   read have.3Pl
        'The boys have certainly read a lot'.

Interestingly, the paradigm for perfect indicative, as provided by the speakers themselves, is as given below. [14]

(46)   a.   *Io-am finit*   *Noi-am finit*
                tu-i finit.   *Voi aţ finit.*
                Je finit-a.   *Čelj finit-a.*
                'I/You/He have/s finished; We/You/They have finished.'
        b.   **Je a finit*   **Čelj a finit.*

Note that while my informants obligatorily place the 3rd persona auxiliary enclitic on the participle, [15] the paradigm in Kovačec (1971) does not capture this asymmetry.

(47)   *(Io) am*   *(Noi) am*
        (Tu) ai /Tu-i   *(Voi) aţ*
        (Ie) a(v)   =   *(Iel') a(v)*

                            +   *cântåt / avut / avzit / copęit / plâns / putuit*
                                   sung / had / heard / digged / cried / travelled

                                            (Kovačec 1971:147)

However, Kovačec himself notes that:

"Cînd se întrebuinţează pronumele personal ca indicator redundant al persoanei şi numărului, atunci verbul auxiliar se află în urma pronumelui personal, iar participiul în urma verbului auxiliar. Daca pronumele personal lipseşte, iar perfectul se află după o pauză, verbul stă, de obicei, după participiu. În vorbirea afectivă, în raport cu ncecesităţile comunicării, această ordine poate fi schimbată."

(When the personal pronoun is used as a redundant indicator of person and number, the auxiliary verb follows it, and the participle follows the auxiliary. If the personal

---

[14] We have rechecked it on several occasions, and no hesitation or variation was encountered.

[15] When no pronominal clitic is present.

pronoun is not overt, and the perfect is used after a break, the [auxiliary] usually stands after the participle. Function of the purposes of communication, this order can be changed in affected speech.)

Kovačec's observation does not make any person distinction of the type we have encountered. We could speculate either that in four decades since Kovačec's grammar was written, *a* has been subject to a change in status (as will be tentatively suggested in Section 5.3.2 in the light of Embick & Noyer's 2001 distinction between syntactic and post-syntactic clitics), or that the Southern variety differs from Žejanski in this respect.

The aforementioned property indicates that IR is not 'clitic second' as (Serbo-)Croatian is.

We have seen that in declarative main clauses, clitics need not be 'second'. Instead, they seem to necessarily surface in second position in interrogative and exclamative root clauses (48), and in embedded clauses (49).

(48) *Če -a   čelj finit lucru?*
      Q have.3Pl they finished work-the
      'Did they finish the work?'

(49) *Se ziče k   -a   Maria čuda čiteit.*
      SE say that have.3Sg Maria much read
      'It is said that Maria read a lot.'

The status of the 'first element' in these contexts a syntactic head (i.e. the question particle *če*, the complementizer *ke* undergoing phonological reduction to *k-*).

However, as indicated in Kovačec's note above, information structure makes this requirement a rather flexible one. See the examples below:

(50) a. *Misles   ke   čaista fost-a   ier ân Umago.*
         believe.1Sg that this.one been =has yesterday to Umago
         'I believe that he went to Umago yesterday.'

      b. *Misles   ke   tu-i fos(t) ier ân Umago.*
         believe.1Sg that you have.2Sg   been yesterday to Umago
         'I believe that you went to Umago yesterday.'

The observed phenomena may suggest that IR is a 'language in transition' (cf. Kroch 1989). Two major reasons that go in this direction are that standard Croatian is much more pervasive in the community than it used to be five decades ago (cf. Kovačec 1971, Sârbu 2003, and others), and that, presently, all speakers are bilingual. Accordingly, two (competing) grammars are active in the speech community (also cf. Shokeir 2006 for a similar proposal for clitic placement in present day Serbo-Croatian).

Syntactically, this line of reasoning translates into capturing facts that on the surface appear to be parallel to Romanian or to Croatian. Under a more careful look, one could be put forth an account for the different properties of perfect auxiliary clitics, 1st and 2nd person vs 3rd person (which in Serbo-Croatian occupy different positions in the cluster hierarchy), and of other auxiliaries which only optionally require Long Head Movement of the verb (cf. Rivero 1991 and further work). Other distribution facts will also follow, such as the data concerning fronted constituents which appear not to intervene with clitic placement (heavy fronting does not intervene in Serbo-Croatian, either), or the non-obligatory adjacency between the auxiliary and the verb (which in Romanian can only be interrupted by a very small class of 'semiadverbs', see Section 5.3.1). We must however leave this for future research.

5.2.2. More on auxiliaries

We have already seen elsewhere that in IR auxiliaries can be detached from the lexical verb. Thus, they do not form a complex head with the verb (as has been claimed for Romanian, see Section 5.3.1. below.). Let us illustrate this again:

(51) a. *Mâre   voi   io učide.*
          tomorrow will.1sg I   kill

    b. *Ân cârca cu   brẹnta   åm   åpa   purtåt.*
        in   back   with bucket-the have.1Sg   water(-the) brought

    c. *Nu ştivu   dupa câta vrẹme åv   jos   verit.*
        not know.1Sg after how much time have.3Sg/Pl down came.

Appearing in isolation is a property that clitic elements lack. Interestingly, 1st and 2nd person auxiliaries may stand alone. The first example below is taken from Sârbu (2009), while (51.b-c) in which the person perfect auxiliary surfaces as an answer to a yes/no question, were provided by my informants. Notice again the asymmetry between the 1st

(51.b), and 3rd person (51.c) auxiliaries. We must note however that the vowel quality of the 1st person auxiliary changes from *a* to the round middle vowel rendered by the grapheme *å*. This may play an important role in the behavior of these elements (1st and 2nd person), but we were unable to obtain a paradigm.

(52)   a. *Cum åm io cuvintåt, av şi ie, Leca Morariu şi muljåra.*
        how have1Sg I spoken, have.3Pl also them, L. M. and wife-the
        'The way I spoke, so did they, Leca Morariu and his wife.'

    b. *Če -ai durmit ier?   Durmit -am   / Åm.*
        Q have.2Sg slept yesterday   slept have.1Sg   /have.1Sg
        'Did you sleep yesterday? Yes, I slept. / I did.'

    c. *Če mes-a az Goran ân Trst?   Mes -a.   / *A.   / V-a.*
        Q gone have.3Sg today G. inTriest gone have.3Sg/*have.3Sg /v have.3Sg
        'Did Goran go to Triest today? Yes, he went.'

On the other hand, the 3rd person auxiliary *a* cannot appear in isolation, it needs a host. The insertion of the epenthetic consonant *v* suffices as a host to a phonological clitic, as illustrated in (52.c).

Unless other pronominal clitics are present, 3rd person auxiliaries are enclitic on the participle. We provide one more example below. Contrast (53.a) with (53.b), where the presence of the pronominal subject seems to function as a host the 1st person auxiliary.

(53)   a. *Dejan mes -a ân beserica.*
        Dejan gone =has in church

    b. **Dejan a mes ân beserica.*
        Dejan has gone in church
        'Dejan went to church.'

(54)   a. **Io mes -am ân beserica.*
        I gone =have.1Sg in church

    b. *Io -am mes ân beserica.*
        I =have.1Sg gone in church
        'I went to church.'

(55)   a. *Ţi l -am dåt.*
        you.Dat.CL him.Acc.CL =have.1Sg given

'I have it to you.'

    b. *Mi l -a dåt.*

      me.Dat.CL him.Acc.CL =have.3Sg given

      'He gave it to me.'

(56)   c. *Nu mi l -a dåt.*

      not me.Dat.CL him.Acc.CL =have.1Sg given

      'He didn't give it to me.'

In main declarative clauses, 3rd person auxiliary *a* can also attach to an element in the left periphery. In embedded clauses, the complementizer appears to be triggering the raise of the auxiliary.

(57)   *Ier -a Redento čuda čiteit.*

      yesterday =has Redento much read

      'Yesterday Redento read a lot.'

(58)   *Se ganę k -a Redento čuda čiteit.*

      SE says that=has Redento much read

      'It is said that Redento read a lot.'

It thus appears that the 3rd person (singular/plural) auxiliary *a* must attach to the past participle, *or* to a complementizer, *wh-* elements, Top/Foc (i.e. temporal adverbials, quantified items), all of which surface in the left periphery of the clause. However, this observation of surface facts needs more data in order to pursue a line of analysis. For the time being, let us bear in mind that a the syntactic and phonological verbal clitic *a* behaves differently from the phonologically more complex members of the paradigm. We will turn to this again later in this chapter.

5.2.3. More on pronominal clitics [16]

5.2.3.1. The masculine/feminine asymmetry

We have seen so far that the positions that the masculine accusative clitic pronoun *l* can occupy are different from its feminine counterpart *vo.*

---

[16] The approach we adopt in the second part of this chapter to clitic elements deals with linearization rules. For cartographic studies, cf. for instance Cardinaletti (2008), Cardinaletti & Shlonsky (2004), among many others.

This reminds us of the Romanian facts, presented here briefly in (59), and to which we will come back in Section 5.3. Note that the phonetic realization of Romanian feminine accusative pronoun is the vowel *o*, whereas the IR one is *vo*.

(59) a. *îl   / o   văd*
      he.Acc.CL / she.Acc.CL see
      'I see him/her'

    b. *l-am văzut.*   */ *am văzutu-l.*
      he.Acc.CL =have.1Sg seen   / have.1Sg seen =he.Acc.CL

    c. **o -am văzut.*   */ am văzut-o.*
      she.Acc.CL =have.1Sg seen   / have.1Sg seen =she.Acc.CL
      'I saw him/her.'

In simple tenses, neither Romanian ((59) above), nor IR (60) display any gender asymmetry. The two languages do however differ in the position of the pronominal clitics with respect to the main verb: in Romanian, they are preverbal, whereas in IR they must surface postverbally in the absence of a negative item or of another pronoun, i.e. Dative clitic or subject pronoun (60.b), with which the accusative can cluster).

(60) a. *Vedu -l.   /vo*
      see.1Sg he.Acc.CL   / see.1Sg she.Acc.CL

    b. *Io   -l   /vo   vedu.*
      I   he.Acc.CL   / she.Acc.CL see.1Sg

The data contrasts when the we turn our attention to compound tenses. [17] Note the different position of the masculine pronominal and the feminine one.

(61) a. *Nona vegljat -a Dejan/Lara.*
      grandma-the looked.after =has Dejan/Lara

    b. *Nona l -a vegljat.*
      grandma-the he.Acc.CL =has looked.after

---

[17] As we have seen, Romanian also displays different orders for *l* and *o* for perfect indicative. However, as it will be clear from the examples, the IR positions of *vo* do not mirror its Romanian counterpart.

c. *Nona vo vegljat -a.*

  grandma-the she.Acc.CL looked.after =has

  'The grandmother looked after Dejan/Lara/him/her.'

In the absence of a subject (or of another initial element), the auxiliary *a* needs to attach to the participle. What we see in (61.a-b) is that masculine *l* also encliticizes on the participle (preceding the auxiliary), whereas the feminine *vo* surfaces postverbally following the auxiliary, too. It appears to be phonologically independent.

Recall however that 1st and 2nd person auxiliaries have a less restricted behavior than the 3rd person one. Such context allows us to check whether the different positioning of the masculine/feminine is maintained in the absence of the restrictions due to the phonological nature of *a*. In (62) we see that indeed *l* and *vo* do not surface in the same position: while *l* does not display a different behavior, *vo* is allowed to (and in fact must) intervene between the auxiliary and the past participle.

(62) a. *Vegljat   -l   -a.*

         looked.after he.Acc.CL =has

    b. *Vegljat   -a   vo.*

         looked.after =has she.Acc.CL

(63) a. *Io l   -am   vegljat.*

         I he.Acc.CL have.1Sg looked.after

    b. *Io -am   vo   vegljat.*

         I have.1Sg sheAcc..CL looked.after

If other elements, i.e. adverbials, also intervene between the auxiliary and the participial verb, *vo* surfaces immediately following the auxiliary.

(64) a. *Io nu l-am   nicad vezut.*

         I   not he.Acc.CL =have.1Sg never   seen

         Neg + CLmasc   + AuxCL > Adv > PastPart

    b. *Io nu -am   vo   nicad vezut.*

        I not have1sg   she.Acc CL never seen

       **Neg + AuxCL   + CLfem   ** > Adv > PastPart

The data presented so far seem to indicate that the different positions of the two accusative pronominal elements may be a consequence of their different status, as will be suggested below.

In what follows, let us look at two more syntactic contexts: coordination and restructuring.

Coordination facts indicate a further asymmetry. Recall that while the masculine clitic clusters with the 3rd person auxiliary, this is not verified for the feminine (61) repeated here as (65)).

(65)   a. *Nona l -a vegljat.*

     grandma-the he.Acc.CL =has looked.after

    b. *Nona vo vegljat -a.*

     grandma-the she.Acc.CL looked.after =has

     'The grandmother looked after Dejan/Lara/him/her.'

If two verbs are coordinated, the direct object of which are the same, two facts are worth noticing: first, the cluster *l –a* in (66.a) needs not be repeated for the second event. Secondly, the feminine pronoun can take scope over the two coordinated verbs (participle + enclitic auxiliary *a*), whether it precedes (66.b) or follows them (66.c).

(66)   a. *Nona l -a [vegljat ši låsat slobodan].*

     grandma-the he.Acc.CL =has looked.after and left free

    b. *Nona vo [vegljat -a ši låsat -a slobodna].*

     grandma she.Acc.CL looked.after =has and left =has free

    c. *Nona [vegljat -a ši låsat -a] vo slobodna.*

     grandma looked.after =has and left =has she.Acc.CL free

     'The grandmother looked after him/her and let him/her free.'

As for restructuring, *vo* can stand both high and in the intermediate position, where Italian clitics (be they masculine or feminine) are banned (see (67)). *L*, on the other hand, cannot appear in these contexts. Instead, the weak pronoun *îl* appears, but it is only allowed in its highest occurence.

(67)   a. *vo poč vedę*   / *îl poč vedę*
        she.Acc.CL can.1Sg see.inf   he.Acc can.1Sg see.inf

        b. *poč vo vedę*   / **poč îl vedę*
        can.1Sg she.Acc.CL see.inf   can.1he.Acc Sg see.inf

        'I can see her/him'

(68)   a. *lo/la posso vedere*   (Italian)
        (s)he.Acc.CL can.1Sg see.inf

        b. *posso veder-lo/la*   / **posso lo/la vedere*
        can.1Sg see –(s)he.Acc.CL   can.1Sg (s)he.Acc.CL see

        'I can see him/her'

The same asymmetry is observed when the restructuring context involves an aspectual verb:

(69)   a. *Io voi pošni prontivei paninu / zema.*
        I will.1Sg start prepare.inf sandwich.Masc-the / soup.Fem-the

        b. *Io -l voi pošni prontivei.*
        I he.Acc.CL will.1Sg start prepare.inf

        c. *Io voi vo pošni prontivei.*
        I will.1Sg sheAcc.CL start prepare.inf

   To sum up, data shows that the Masculine and the Feminine Accusative clitics are occupy different positions. As such, they subject to different linearization rules. Moreover, the placement of *vo* is clearly not phonologically driven (as with Romanian *o*, see Section 5.3).

   In the next section, we will draw our attention to the different typology of clitics as proposed by Embick & Noyer (2001), and used to account for Romanian *o* by Giurgea (2010). We will eventually see what determined the differences between the feminine accusative clitic pronouns in IR (*vo*) and Romanian (*o*), and suggest that the IR 3rd person perfect auxiliary *a* corresponds to the PF-clitic in Embick & Noyer's classification of clitic elements.

5.3. Syntactic clitics vs PF-clitics (Embick and Noyer 2001)

5.3.1. Romanian verbal clitic cluster

It is a well known empirical fact that the Romanian verbal elements (in (70) below), from the mood markers to the past participle) cluster together, following a rigid order:

(70) [$_{CP}$că/ca…..[$_{M(ood)P}$să/a   [$_{NegP}$nu[$_{CliticP}$[$_{AgrS+TP}$Aux/Vfin[$_{AspP}$(fi) (Adv) Vpast.part[$_{vP}$]]]]]]
　　　that　　that.subjunctive not　　　　　　perfective
　　　　　/to.infinitive

The theoretical point under debate is whether the obligatory adjacency is a result of the lack of specifiers of all projections in the string, or that all these elements for a complex head in syntax. The first approach is entertained in Avram (1999), Alboiu (2002), who argue that the head-initial order inside the cluster is a proof that it is not a complex X°, and that adjacency can only be explained by the fact that the projections lack Specs.

However, as Giurgea (2010) points out, there are several problems with this account.

Problem 1: If aspectual adverbs are in AspP, why is it only the restricted class of 'clitic ad-verbs'(*cam, mai, prea, şi, tot*) and not any other adverbs are allowed in this position (but rather, if present in Spec,MoodP)?

The second problem comes from coordination facts. No subset of projections can be coordinated (in the case of two events):

(71)　　a. *Cred că [va pleca şi va uita].*
　　　　　　believe.1Sg that will.3Sg leave and will.3Sg forget
　　　　　'I believe that he will leave and forget.'
　　　　b. **Sper să [plece şi uite].*
　　　　　　hope.2Sg SUBJ leave.3Sg and forget.3Sg
　　　　　'I hope that he leaves and forgets.'
　　　　c.**va [pleca şi uita].*
　　　　　　will.3Sg leave and forget
　　　　d. **a [plecat şi uitat].*
　　　　　　have.3Sg left and forgotten

Giurgea (2010) correctly points out that there are some exceptions, i.e. if the two events have a semantic relation (are logically sequential), but he does not try to give an account

for these facts:

(72) a. *Nu [mânca şi vorbi] în acelaşi timp.*

        not eat.2Sg and speak.2Sg in same time

        'Don't eat and speak at the same time.'

    b. *Voi citi şi corecta toate temele.*

        will.1Sg read and revise all homework.Pl-the

        'I will read and revise all homework.'

    c. *Le-am citit şi corectat cu atenţie.*

        they.Acc.CL have.1Sg read and revised with attention

        'I read and revise them all with care.'

However, no intervening material is allowed between the two coordinated lexical verbs:

(73) a. **Voi citi mâine şi corecta toate temele.*

      will.1Sg read tomorrow and revise all homework.Pl-the

    b. **Le -am citit cu atenţie şi corectat.*

      they.Acc.CLhave.1Sg read with attention and revised

    c. ?(?) *Nu vorbi cu alţii şi mânca in acelaşi timp.*

        not talk.2Sg to others and eat.2Sg in same time

        'Don't talk to others and eat at the same time.'

Moreover, the complements/adjuncts must belong to both verbs:

(74) **Au [intrat şi salutat cu jumătate de gură].*

     have.3Pl entered and saluted with half of mouth

     'They entered and said hello unwillingly.'

From the above, Giurgea (2010) concludes that what is at stake is X° coordination, which means that the verbal cluster involves a complex X° of sorts, which can comprise two coordinated heads.

Let us now turn to the third Problem, namely the position of the feminine singular Acc clitic *o*, which we have already mentioned as being 'quirky' with respect to the other pronominal clitics in the paradigm:

(75)   a. *m-/ te-/ l-/   ne-/ v-/ /i   /le-   a(u)   cumpărat*

        1Sg 2Sg 3Sg.Masc 1Pl   2Pl 3Pl.Masc/Fem   =have/has bought

    b.   **o a cumpărat*

        3Sg.Fem has bought

    c.   *a cumpărat -o*

        hasbought 3sg.Fem

The feminine singular accusative clitic is realized as one vowel, *o* .[18] On the surface of the facts in (75) above, there seems to be a ban for *o* to precede the vowel-initial auxiliary. Note that while *o* can precede the future marker *va* (75.d), it can also precede a lexical verb starting with a vowel (75.e). Thus, the linearization rule to which *o* is subject cannot be *[VOWEL].

    d. *o va lua*

        she.Acc.CL will.3Sg take

    e. *o ia*

        she.Acc.CL takes

According to Embick & Noyer (2001), there are two types of PF rules:

a)   rules applying before Vocabulary Insertion, which can make reference to structural information (thus to the notion of 'head of the complement') -> *PF head-clitics* (i.e. English Affix Hopping);

b)   rules applying after Vocabulary Insertion, which only sensitive to linear adjacency and the grouping of morphemes into words -> *PF linear-clitics*.

At Vocabulary Insertion, linearization applies: hierarchical information is lost, except for the grouping of morphemes into morphosyntactic words. Displacement rules applying at this stage, called 'local dislocations', target adjacent words or morphemes (they may reverse the linear order between adjacent morphemes inside an m-word, or between a morpheme and the following m-word).

Giurgea (2010) builds on Embick & Noyer's (2001) 'local dislocation' operation. He argues for a diplacement PF rule which affects a clitic consisting of a vowel (in this case,

---

[18] The masculine plural accusative is realized as one vowel, as well. However, it patterns with the rest of the clitic elements in the paradigm, which makes the PF rule advocated for by Giurgea (2010) item-specific. The question why is *o* the only 'quirky' element in the paradigm remains unanswered.

o) , in the context [Aux VOWEL]. This rule applies at or after Vocabulary Insertion. In the view of Embick & Noyer (2001), the target of such PF rules cannot be phrases, but only morphemes or words. Such an account explains the surface position of the feminine accusative singular clitic: *o* is displaced at the end of a complex X° (the whole cluster in () above).

Coming back to the obligatory adjacency of mood makers, the negative marker, pronominal clitics, auxiliary, the perfective marker and the lexical verb, if it is a complex head, it cannot have been a consequence *PF linear* rules, since obviously structural information is relevant. Thus, the complex X° may have been formed by rules that create *PF head-clitics.*

Comparing Romanian the data to French subject clitics, which are shown to *PF head-clitics* by several tests, Giurgea (2010) concludes that the string of elements in around the Romanian verbal phrase are not *PF head-clitics* either.[19] Note the contrast involving coordination of phrasal constituents below:

(76)   *Il [[leva les yeux] et [commenca à parler]].*   (French)
         he raised the eyes and started to talk

(77)   **A [[ridicat ochii] şi [început să vorbească]].*   (Romanian)
         has raised eyes-the and started SUBJ talk
         'He raised his eyes and started talking.'

The third possibility for a complex X° formation is that it takes place prior to PF, i.e. in syntax. If so, syntactic operations (i.e. movement) should apply to the entire complex head.

One piece of evidence for syntactic movement that Giurgea (2010) puts forth comes from thetic clauses, which in Romanian linearize as VS(O). Thetic sentences have narrow focus either on the verb or on the subject (Giurgea & Remberger 2009).

---

[19] French subject clitis are not a complex head in syntax, since they do not take part in movement (ii) (as does the negative element or object clitics, in French). Subject-clitic inversion leaves *il* in a postverbal position in iii below:

  i. *Il (*deja /*apparemment /*probablement...) est venu.*
      he already apparently probably is come

  ii. **Il, je suppose, est venu*
      he I suspect is come

  iii. *Ne le ferait il pas?*
       not he.Acc.CL do.COND he not

(78) a. *A STIUT el cum să obţină slujba.*

     has known he how SUBJ obtain job-the

     'He DID know how to get the job.'

  b. *L -a IUBIT Maria pe Ion.*

     he.Acc.CL has loved Mary *pe* Ion

     'Maria DID love John.'

As the capital letters in the example above indicate, the verbal cluster has undergone focus fronting, a syntactic operation that affects not only the lexical verb itself (which bears the focus), but also the other elements adjacent to it (in the examples above, the pronominal clitic and the auxiliary). [20]

The second piece of evidence comes from Hill's (2006) Subject-Pronoun Inversion Construction: Hill (2006) proposed the raising of the verbal cluster to a projection in the CP layer:

(79) a. *(Cât despre) Maria, (ea) precis (ea) a calculat (ea) deja (?*ea) datele.*

     as for Maria, she surely she has computed she already she data-the

  b. *Maria, (*ea) a CALCULAT *(ea) precis (*ea) deja (*ea) datele.*

     Maria she has compued she surely she already she data-the.

The phenomena above lead Dobrovie-Sorin (1994, 2001), Matushansky (2006), Giurgea (2010) to conclude that the Romanian verbal cluster forms a complex head in syntax. This leads to the question of how to derive the head-initial linear order of the elements in the cluster.

Giurgea (2010) assumes that the head-final order is default setting of the parameter in Indo-European languages, and it is a rule operating at Vocabulary Insertion. The head-initial order, on the other hand, is obtained by specific linearization instruction which overrides the default rule in languages where prefixation is predominant. Accordingly, Romanian Negation, Auxiliaries, Mood and perfectivity markers bear a syntactic feature which triggers head-movement, and a morphological feature which he calls 'Pref' that triggers prefixation. The proposal is schematized below:

---

[20] presumably to a position higher than T°, in a sisterhood relation H-H between the complex head and T°.

(80) *linearization of head-final complex X°* (default rule):

    a.   [X° Y°] 9 → Y° > X°

    b.   *mai citesc*

        Adv read.1Sg

    c. $[_{T°}[_{Asp°}Asp°[_{v°} V°]]] → [[_{Asp°}Asp°[_{v°} V°]] > T°]$

(81) *linearization of head-initial complex X°*:

    a.   [X°+Pref Y°] → X° > Y°

    b.   *am mai citit*

        have.1Sg Adv read.PastPart

    c. $[_{T°+Pref}[_{Asp°} Asp° [v° V°]]] → [_{T°} T° > [_{Asp°} Asp° [v° V°]]]$

From (80-81) above, it results that the final position of the feminine Acc clitic *o* results, the only element in the cluster that does not occupy the position where it is expected for it to be found, is a result of the application of default rule in:

(82) $[[_{CL} o][_{T}[_{T}Vowel-...]]] → [[_{T}T...] >[_{CL} o]]$

    The solution above seems to capture most fact in an elegant way. However, the objection that can be raised to Giurgea (2010)'s proposal concerns is the following: if the position masculine clitic *l* (preceding the Aux) is obtained because of the [+pref] feature on T, should he claim that T does not have this feature when *o* is merged? If yes, this would imply an undesirable optionality. Even more so since "this feature [...] is a morphological feature of individual vocabulary units" (p.13), thus each item realizing T should bear [+pref] or [-pref] irrespectively of the presence of a pronominal clitic.

5.3.2. Implications on Istro-Romanian clitics

The linearization operations as have been applied to account for the order observed in the Romanian verbal cluster (made of clitic and clitic-like elements) may be used to also explain the different position of clitic items in IR.

    At a speculative level, we note that while the behavior of the feminine accusative pronominal clitic *vo* differs from Romanian *o* (also the phonetic quality of the element is different). However, the absence of the [+pref] feature may account for the position of the third person (singular and plural) perfect auxiliary *a* in IR, as documented throughout

Section 5.2.3.1. However, a careful analysis along these lines is needed, which we must leave as part of a future endeavour.

The final section will shed some light on interesting facts from an older stage of Romanian.

5.4. Old Romanian

A diachronic perspective over the phenomena we have dealt with in this chapter reveals some interesting facts.

The data from Old (Common/Ancient/Primitive) Romanian is taken from Zamfir (2007) and Uşurelu (2009), who have collected it from historical documents ranging from the 14th century to the 18th century.

In Old Romanian, auxiliaries do not form a complex syntactic head with the verb. We observe this by the fact that it allows other constituents to intervene. This property recalls Istro-Romanian, and it diverges from modern Romanian – which must have developed the adjacency requirement in more recent times.

(83) a. *Prince ţie   au Domnedzeu toate acestea arătat.*   (subject)

　　　because you.Dat has God   all these shown

　　　'Because God showed all that to you.'

　　　　　　　　　　　　　　　　　　　　　　　(Palia de la Orastie, 1581, XLI:39)

　　b. *Cum aţi astadzi aşa curând venit.*   (adverb)

　　　how have.2Pl today so soon arrived

　　　'Since you came so soon today.'

　　　　　　　　　　　　　　　　　　　　　　　　(ibid, Exodul, II:18)

　　c. *[…]am eu vindut*   (subject)

　　　have I sold

　　　　　　　　　　　(Documenta Romaniae Historica A, 1626-1628, XIX:28)

It has been noted that in the 14th-17th centuries, such scrambled orders was very frequently found in texts (Zamfir 2007, Uşurelu 2009).

The scrambled order of clausal constituents as illustrated above reminds us of verb-second languages. The inflected verbal element appears high, possibly in the complementizer domain. Modern Germanic languages are V2, but also Old Germanic. While Old Scandinavian had a "rigid" V2 requirement in root contexts, allowing only one

169

constituent in preverbal position, Old Romance (Fischer and Alexiadou 2001), in particular Old Italian, had a "relaxed" V2, i.e. one or more than one consitutents could precede the inflected verb (cf. Franco 2009 and references therein).

This line of research would require a more careful analysis of the data, which we will not pursue here. It is however meaningful to notice the similarities between Old Romanian and Istro-Romanian, to the exclusion of modern Romanian.

Another apparently productive phenomenon in Old Romanian was inversion (syntactically Long Head Movement). See two examples below, also involving pronominal clitics which still precede the inflected verbal element (the auxiliary *ai* in (.a) and the future marker *voi* in (.b)), as they would do if the lexical verb would linearly follow the clitic cluster.

(84)   a. *din iadul scosu -ne -ai*

    from hell.the got.out us.CL have.2Sg

    'You took us out of hell.'

    (M. Ieud, 1391-1392)

   b. *închina -mă -voi cătră beserica svântă a Ta*

    pray me.Acc.Cl will.1Sg to chirch holy of yours

    'I will pray to your holy Church.'

    (Varlaam, 1645)

As for pronominal clitics, they generally seem to surface postverbally. See the examples below for the indicative (present, past simple and compound):

(85)   a. […] *celuia ce toate silele ceriului slujescu -i*

    the.one.Dat which all creatures-the sky.Gen serve.3Pl him.CL

    'to the one to whom all the creatures in the sky pray'

    (M. Ieud)

   b. […] *pentru că ce vădzu Cain toate lucrurele frățâne -său, ucise -l*

    for that what saw.3Sg Cain all things-the brother his killed.3Sg him.CL

    'because of all his brothers's belongings which Cain saw, he killed him. '

    (Floarea darurilor, 1480?, 128)

   c. *am săgetatu -le cai*

    have.1Sg darted them.Acc.CL horses

'I threw arrows at their horses'

Coming now to the gender asymmetry observed for Istro-Romanian third person singular pronominal clitics, Old Romanian displayed two different stages, both of which worth mentioning for a comparison.

Feminine Accusative clitic *o* always precedes the auxiliary, even in the phonological context [Vowel-]. This contrasts with modern Romanian, where, as we have seen, *o* is obligatorily enclitic on the participle verb.

(86)   a. *şi calea cea dreaptă mi -o au tremisu*

and path-the the right to.me.CL her.CL has sent

'and he has sent me the right path'

b. *o au chiemat de o a întrebat*

her.CLhave.3Pl called toher.CLtohas asked

'they called her in order to ask her'

In constructions with Modal + Subjunctive and Aspectual + Infinitive, the pronominal clitic appeared on each verb.

(87)   a. *că nu -l poate să -l înveţe nimica*

that not him.CLcan.3Sg SUBJ him.CLteach.SUBJ nothing

'that he can't teach him anyhing'

b. *începutu -l a -l întreba*

begun him.CL to him.CL ask.INF

'he began to ask him'

Later, in the 17th century, pronominal clitics appeared twice (proclitic and enclitic) in other contexts too, clearly monoclausal:

(88)   a. *de te voi cerca -te cu faţa curată*

     if you.Acc.CL will.1sg search you.Acc.CL with face-the clean

     'if I shall look for you with a clean face'

(Dosoftei, Psaltirea, 179)

    b. *după ce l -au slobodzitu -l turcii*

    after what him.Acc.CL released him.Acc.CL Turks-the

    'after the Turks set him free'

(Neculce, 66)

    c. *scrisoarea o au dat -o lui Staico*

    letter-the 3sg.Fem.CLhave.3plgiven3sg.Fem.CLto Staico

    'they gave the letter to Staico'

(Anonimul, Brâncovenesc, 301)

The Old Romanian data may be taken as an indication that the different position of the masculine versus feminine singular accusative clitics may have been the result of different choices that took place at a certain diachronic stage. Although both varieties have developed a gender-based asymmetry, this process has lead to different results in Romanian and Istro-Romanian.

5.5. Ad interim conclusions

The present chapter consisted of two main parts.

The first part had a descriptive purpose. We have presented the data which we collected by using a typological tool, namely a *Clitic Questionnaire* which has already been applied to many other languages. We were thus able to compare the IR data with the answers given for Romanian and (Serbo-)Croatian.

As for the second part (from Section 5.2 onwards), we have concentrated on clitic placement, function of significant asymmetries. Firstly, the different position and syntactic properties of perfect auxiliaries (1st /2nd versus 3rd person), and secondly, pronominal accusative singular clitics (masculine versus feminine). Finally relevant data from Old Romanian is presented.

The data presented in this chapter may stand as basis for further studies on dialectal variation: Daco-Romanian, Istro-Romanian, Megleno-Romanian, Aromanian. Diachronic variation is currently under study at the University of Bucharest (cf. Uşurelu 2009).

Furthermore, a more in depth look a clitic elements in IR may result in important implications for the theory on clitics and (long) head movement (in the sense of Kayne 1991, Lema and Rivero 1991 and further work).

6. Concluding remarks and future research

Historical linguists and Romanian dialectologists traced the splitting of the four major present-day Eastern Romance dialects (Daco-Romanian, Aromanian, Megleno-Romanian and Istro-Romanian) to somewhere between the second half of the first millennium CE and the 12th century. While it is clear that Daco-Romanian, Aromanian and Megleno-Romanian (to which greater attention has been paid even in recent generative approaches to language study, cf. Friedman 1994, Tomić 2004, Campos and Stavrou 2004, Campos 1995, 2005, 2006, etc.) developed from Common Romanian as "sisters", IR appears to have a closer relation to Daco-Romanian than with the two other varieties (Puşcariu 1926). Hence it is believed (although not largely accepted) that IR is "daughter" to Daco-Romanian. Moreover, IR is considered to be the "least Balkan" of the four, i.e. it displays the least number of properties in common with the other languages belonging to the Balkan Linguistic Area (cf. Joseph 1999 for a Romanian as a Balkan language, Lindsted 2000 for an overview of Balkan features).

Starting from the above general premises, I have attempted to answer the following **two research questions in Chapter 2**: 1) what does IR have in common, and in what respects does it differ from Daco-Romanian (henceforth DR), Aromanian (Ar) and Megleno-Romanian (MR)?, and 2) what specific syntactic phenomena can be identified to belong to the Balkan Linguistic Area properties? A third, yet not answered research question (to be approached at in future work) would be: what language variation theory (i.e. contact, prolongued bilingualism of the entire community, etc.) can best explain the relevant IR syntactic phenomena which have been identified.

The **methodology** can be largely described along the following lines: collecting empirical data and analysis of the data in the generative linguistics framework (Chomsky 1957 onwards), which places the present study out of the lines of previous monographs. The most important **findings** are have been organized in three chapters: **Chapter 3** deals with the nominal expression and definiteness, i.e. the development of the (definite and indefinite) article system in Eastern Romance (from Latin) dates prior to the split into the four present day idioms. What can be evinced from my own data (properly described) about how (in)definiteness is realized in IR differently from DR, Ar and MR was explained by a theory of grammaticalization, combined with some speculations on the effects of

language contact. Also, modification in the nominal expression plays an important role in how referentiality is expressed.

Chapter 4 is concerned with the Verb Phrase, and is mainly descriptive in nature, i.e. no major claims are put forth.

The behaviour of (pronominal) clitics in IR, already identified as "between Slavic and Romance" was carefully described and analyzed in **Chapter 5**, which underlined not only some microcomparative potentiality of clitic elements (i.e. Slavic "clitic second" phenomenon), but also broader implications for syntactic theory (i.e. "verb second" in Germanic languages). However, this was but an initial state of some research that needs to be carried on to obtain conclusive results.

Last but not least, I would suggest that the current work is an open door to some interesting *intra*disciplinary discussions: historical linguistics (what can be inferred about the approximate time of the split of IR from Common Romanian?), language contact (between Slavic and Romance, given that IR developed in isolation from all other (Proto-)Romanian varieties for at least five centuries), typological consideration (the contribution of research on a previously understudied language), and also preservation of severely endangered languages (by providing a formal description).

Appendix 1: Other data

The pieces of data attached in the appendix are not systematic, either because they occurred in spontaneous speech, or because they occurred while I was aiming at different phenomena. Due to lack of time, we could not investigate further on these pieces of data. I shall present them as I have transcribed them at the respective time (January / June 2010).

(1)   *Te piažę Zagrebu.   / Zagrebu je mušåt gråd.   /Mes-am ân Zagreb.*
        youAcc like Zagreb-the   /Zagreb-the is beautiful city /gone have1SG in Zagreb
        'You like Zagreb. Zagreb is a beautiful city. I went to Zagreb.'

(2)   *Conosc doi madit''.   Ur âi din Zagreb, si ur âi din Umago.*
        Know1SG two young.men.   One is from Zagreb, and one is from Umago
        'I know two you man. One is from Zagreb, and one is from Umago.'

(3)   *Čela din Zagreb mai mund mi -e dråg.*
        the.one from Z more much meDatCL is dear
                      / Mije -i mai dråg čela din Zagreb.
                       meDat   is more dear the.one from Z
                              Lu Lara-i mai dråg čela din Umag.
                              to Lara is more dear the.one from Umag
        'The one from Zagreb is more dear to me.
                              To Lara, it is the one from Umag who is more dear.'

(4)   *Mę dorę gerunclju.*
        meAcc hurt1SG knee-the
        'My knee hurts.'

(5)   *Io-m opilit lęmnele,   ši   mi-am   taljåt žåžetu /ur žåžet.*
        I have1SG cut wood-the, and meDat have1SG cut finger-the /a finger
        'I was cutting the wood and I cut this finger / a finger.'

(6)   *žåžetele de desna mâr /* mâr desna*
        fingers-the of right hand / *hand right

177

(7)

dupa scånd	'behind table'
pre	'on'
su	'under'
ântra	'in front (of)'
dispre	'from, off' (source directional)
prin/scros	'through'
ân	'in'
cu (+N+art)	'with'
na	'at (/?to)'
la	'to' (/?at)'

(8)   *Mižulu cazut-a dispre scånd.*
      glass-the fallen have3SG off table
      'The glass fell off the table.'

(9)   *Roba âi ân picåbit.*
      stuff-the is in drawer
      'The stuff is in the drawer.'

(10) *pičoru de scånd = pičoru lu scåndu*
      leg-the of table   leg-the ofGen table-the

(11) *Mes-am ân besęrca cu fečoru / cu ur fečor.*
      gone have1SG in church with boy-the /with a boy
      'I went to the church with my son / with a son.'

(12) *Trecut-am prin šcur tunel.*
      passed have1SG through dark tunnel
      'We went through a dark tunnel.'

178

(13)   *Mežen la Nela / ân Trst.*
Go.1PL to Nella / in Triest
'We're goint to Nela / to Triest.'

(14)   *Din Hrvatska priste Slovenije mežen ân Italjie.   (Slovenija âi mušåta).*
from Croatia over Slovenia go1PL in Italy.   SloveniaNom is beautiful
'From Croatia, we go through Slovenia to reach Italy. (Slovenia is beautiful.)'

(15)   *Nåradu męrę ân besęrica.*
people-the go3SG in church
'People go to church.'

(16)   *Ân vęra męrę nåradu la måre.*
in summer go3SG people at sea
'In the summer, people go to the seaside'

(17)   *Ân vęra se zležescu čuda cåpre.*
in summer refl born3PL many goats
'In the summer many goats are born.'

(18)   *Breči scu påmetne životinje.*
dogs-the are intelligent animals
'Dogs are intelligent animals.'

(19)   *Måia båte čeljadu   / čeljada.   /ur čeljad*
mother-the beats child-the   /children-the   /a child
'The mother scolds the child / children / a child.'

(20)   *Čiteit-ai čuda libri?*
read have2SG many books
'Did you read many books?'

179

(21) *Če-ai mâncåt?*

Q have2SG eaten

'Have you eaten / What have you eaten?'

(22) **Če mâncåt -ai?*

Q eaten have2SG

(23) *Če -i Lara mušåta fęta / fęta mušåta?*

Q is Lara beautiful girl /girl beautiful

'Is Lara a beautiful girl?'

(24) *Če -i mušåt Dejan?*

Q is handsome D

'Is Dejan handsome?'

(25) *Če-i Lara cåsa?*

Q is Lara home

'Is Lara home?'

(26) *Če je vrur cåsa?*   *Voi vedę se-i vrur casa.*

Q is anybody home   will1Sg see if is anybody home

'Is anyone homw?   I'll see if anyone's home.'

(27) *(*Če) Pre telefonin te-a clemat?*   = *Če te-a clemat pre telefonin?*

(*Q) on phone youAcc have3SG called?   (*Q) youAcc have3SG called on phone?

'Did she call on the mobile phone?'

(28) *Ier-a clemat Nela?*   *Ier   t-a clemat Nela?*

yesterday have3SG called Nela?   yesterday youAccCL have3SG called Nela

'Was it yesterday that Nela called?'   'Was it yesterday that Nela called you?'

(29) *Če t-a   ier clemat Nela?*

Q youAccCL have3SG yesterday called Nela

'Did Nela call you yesterday?'

180

(30) *Če Nela ier clemat-a?
Q Nela yesterday called have3SG

(31) Če-i mę česta zelen mižul? Je.
Q is mine this green glass? is
'Is this green glass mine? Yes it is.'

(32) Če -a čelj finit lucru?
Q have3PL they finished work
'Did they finish the work?'

(33) Če-ai vezut fečori?   Vezut-am fečori.   / Åm.
Q have2SG seen boys-the   seen have1SG boys-the   /Have1SG
'Did you see the boys?   I saw the boys.   /Yes I did.'

(34) Če -ai durmit ier?   Durmit-am   / Åm.   / N-am.
Q have2SG slept yesterdat   slept have1SG / have1Sg   not have1SG
'Did you sleep yesterday?   I slept.   / Yes I have. / No I haven't.'

(35) Če m-   am   ponešeit bire åz?   (V-) Ai.   / N-ai.
Q refl1SG have1SG behaved well today   (expl) have2SG   not have2SG
'Did I behave today?   / Yes you have. / No you haven't.'

(36) Če mes-a az Goran ân Trst? Mes-a.   /V-a   /*A.
Q gone have3SG G in Triest gone have3SG   /expl have3SG   / *have3SG
'Did Goran go to Triest today? He went (to Triest). / Yes he has.'

(37) Cân ai fos la Vanda? Fost-am ieri. / *Ier fost-am / Ier am fos. / *Ier am.
when have2SG been at Vbeen have1SG yesterday
'When did you go to Vanda? I went yesterday.'

181

(38)   *Čire-a telefoneit?*   *Čire telefoneit-a? (only for Nela)*
       who have3SG called   who called have3SG
       'Who called?'

(39)   *Čire cumparåt-a regålu?*
       who bought has present-the
       'Who bought the present?'

(40)   *Čire čåsta cumpårat-a?*
       who this bought have3SG
       'Who bought this?'

(41)   *Če-a zis?   Nela zis-a   ke va veri   ne vedę.*
       what have3SG said   Nela said have3SG that will3SG come   usAcc see.inf
       'What did she say? Nela said she'll come see us.'

(42)   Če-a Nela cumparåt? /*Če Nela cumparåt-a?   /Če -ai cumparåt? / *Če cumparåt-ai?
       what haves N bought /what N bought has /what have2SG bought / what bought have2SG
       'What did Nela buy?   / What did you buy?'

(43)   **Čire če-a cumparåt?*
       who what have3Sg bought

(44)   *Če-a vrur cumparåt?*
       what have3SG somebody bought
       'What did someone buy?'

(45)   *Lu   cui ai dåt regålu?*   *Lui   / L'ei   = Lu iå.*
       prepGen whom have2SG given gift-the   heGen /sheGen   prepGen she
       'Who did you give the present to?   To him / to her'

(46)   *Cân-a telefoneit?   *Cân telefoneit-a?*
       when have3SG called when called have3SG
       'When did she call?'

(47) *Iuv-ai mes?*   *Mes-am ân beşęrica.*   *Io-m mes ân beşęrica.*

where have2SG gone gone have1SG in church   I have1SG gone in church

'Where did you go? I went to church.'

(48) *Zǎc- ai mes ân beşęrica?*   *Ke mę piažę męrę se pomoli.*

why have2SG gone in church   because meDat like go refl pray.inf

'Why did you go to church?   Because I like going to pray'

(49) *Zǎc nu-ai*   *mes ân beşęrica?*   *Ke nu-am vut lezna.*

why not have2SG gone in church   because not have1SG had time

'Why didn't you go to church?   Because I didn't have time.'

(50) *Cum-ai scuheit zęma?*

how have2SG cook soup-the

'How did you cook the soup?'

(51) *Cât   -ai {*åpa} pus {åpa}?*

how.much have2SH {*water} put {water}?

'How much water did you put?'

(52) *Cât́ ań (v)-åre?*

how many years (expl) have3SG

'How many years does she/he have?'

(53) *Cât́ fečor ai vezut?*   *Vazut-am činč.*   / *Činč fečor am vezut.*

how many boys have2SG seen   seen have1SG five   / five boys have1SG seen

'How many boys did you see?   I saw five.   / I saw five boys.'

Different Writing Systems Used to Write Vlaški/Žejanski

(Vlashki/Zheyanski, also, Istro-Romanian)

Modified Romanian Orthography	Mixed Orthography	Modified Croatian Orthography	Example[1]
Morariu (1928)	Kovačec (1998)	Vrzić (2009)	
a	a	a	acmo / acmo / akmo
a (boldface type)	å	å	**a**b / åb / åb
ă	â	â	cănd / când / kând
b	b	b	bât
c (+ a, å, â, o, u, consonant)	c (+ a, å, â, o, u, consonant)	k	c**a** / cå / kå
ce, ci, cĭ (word final)	č	č	cir**a** / čirå / čirå
d	d	d	dintru / dintru / dintru
-	(ʒ)	(dz)	- / ʒero / dzero (very rare)
e	e	e	f_ecior / f_ečor / f_ečor
e (boldface type)	ę	ę	s**e**ra / sęra / sęra
f	f	f	fini / fini / fini
g	g	g	grumb / grumb / grumb
gh (+ i, e, ę, consonant)	g	g	bl**age** / blågę / blågę
g	γ	g	grije / γriže / griže
-	(ǧ)	(đ)	- / ǧoventu / đoventu (very rare)
h	h	h	hiti
i	i	i	inche / inkę / inkę
ĭ, i	ʌ	j	ĭo / ʌo / jo
î	â	â	în / ân / ân
ch (+ i, e, ę)	k (+ i, e, ę)	k	che / ke / ke
ch (+ i, e, ę), chi (+ a, å, â, o, u)	t"	ć	chi**a** / t"å / ćå
l	l	l	la / la / la
l'	l'	lj	l'epur / l'epur / ljepur
m	m	m	m**a**nchie / månt"e / månće
n	n	n	n**e**vu / nęvu / nęvu
n'	ń	nj	n'ive / ńive / njive
o	o	o	opt / opt / opt
p	p	p	per / per / per

---

[1] A word, as it would be written following the Modified Romanian Orthography, is given first. After a slash, the same word is written using the Mixed Orthography and then the Modified Croatian Orthography.

r	r	r	r**a**da / r**å**da / r**å**da
s	s	s	s**a**re / s**å**re / s**å**re
ş	ş	š	şti / şti / šti
t	t	t	tari / tari / tari
ţ	ţ	c	ţire / ţire / cire
u	u	u	ur / ur / ur
ŭ, u	u̯	u	g**ŭa**nti / γu̯**å**nti / gu**å**nti
v	v	v	vir / vir / vir
z	z	z	zice / ziče / ziče
j	ž	ž	juc**a** / žuc**å** / žuk**å**

References

Abels, K. (2003). *Successive cyclicity, anti-locality, and adposition stranding.* Ph.D. Diss, University of Connecticut.

Abney, S. (1987). *The English Noun Phrase in Its Sentential Aspect.* Ph.D. Diss. MIT Cambridge, Mass.

Adger, D. (2003). *Core Syntax.* Oxford, OUP.

Alboiu, Gabriela (2002) The features of movement in Romanian. Bucharest, Editura Universităţii Bucureşti.

Alexiadou, A. & C. Wilder (1998). 'Adjectival Modification and Multiple Determiners', in A. Alexiadou and C. Wilder (eds.) *Possessors, Predicates and Movement in the DP,* Amsterdam, John Benjamins, 303-332.

Aljović, N. (2002). 'Long Adjectival Inflection and Specificity in Serbo-Croatian'. *Recherches Linguistiques de Vincennes* 31, 27-42.

Aljović, N. (2005). Syntactic positions of attributive adjectives: the case of Serbo-Croatian long and short adjectives. Paper presented at the workshop JE "Adjectifs", Paris, 28 September 2005 (handout available at: http://peternad.club.fr/aljovic/home.htm).

Avram, Larisa (1999) Auxiliaries and the structure of language. Bucharest, Editura Universităţii Bucureşti.

Baker, M. (1988). *Incorporation: A Theory of Grammatical Function Changing,* University of Chicago Press.

Belleti, A. & L. Rizzi (1998). 'Psych Verbs and θ - Theory', in *Natural Language and Linguistic Theory*, 6, pp.291-352.

Birnbaum, H. (1965). 'Balkanslavisch und Südslavisch: Zur Reichweite der Balkanismen im südslavischen Sprachraum', in *Zeitschrift für Balkanologie* 3 (1-2), 33-48.

Bošković, Z. (2000). 'Second position cliticisation: Syntax and/or phonology?', in F. Beukema and M. den Dikken (eds), *Clitic phenomena in European languages*, Amsterdam: John Benjamins, 71–119.

Bosković, Z. (2004). 'Clitic placement in South Slavic'. *Journal of Slavic Linguistics* 12, 37-90.

Bouchard, D. (1984). *On the Content of Empty Categories*. Foris, Dordrecht.

Bowers, J. (2002). 'Transitivity'. *Linguistic Inquiry* 33, 183-224.

Bresnan, J. (2001). *Lexical-functional syntax*. Oxford: Blackwell.

Browne,W. 1974. 'On the problem of enclitic placement in Serbo-Croatian. In R.D. Brecht and C.V. Chvany (eds), *Slavic transformational syntax*, Michigan Slavic Materials, vol. 10, Ann Arbor, 36-52.

Burge', L. (1996). 'Demonstrative Movement in Spanish: A Comparative Approach', *University of Venice Working Papers in Linguistics 6*, 1-53.

Burge', L. (2002). 'The positions of Demonstratives in the Extended Nominal Projection', in G. Cinque (ed.), Functional Structure in the DP and IP. The Cartography of Syntactic Structures, volume 1, Oxford, UK, Ed. Oxford Unversity Press, 16-80.

Burzio, L. (1986). *Italian Syntax,* Dordrecht: Kluwer.

Campbell, L., T. Kaufman & T. Smith-Stark (1986). 'Meso-America as a Linguistic Area'. *Language* 62: 3, 530-570.

Campos, H. (1995). 'Full and reduced clitics in Megleno-Romance'. In *Probus* 17.3, 247-278.

Campos, H. (2005). 'Noun modification, pseudo-articles, and last resort operations in Arvantovlaxika and in Romanian'. In *Lingua* 115.3, 311-347.

Campos, H. (2006). 'Possessive Constructions in Greek Aromanian'. In B. Fernández-Soriano and I. Laka (eds.), *Andolin gogoan. Essays in honor of Professor Eguzkitza*. Bilbao: Euskal Herriko Unibertsitateko Argitalpen Zerbitzua, 205-220.

Campos, H. and M. Stavrou. (2004). 'Polydefinite constructions in Modern Greek and in Aromanian'. In Tomic, O. (ed.), *Topics in Balkan Sprachbund Syntax and Semantics*, Benjamins, Amsterdam, 147-186.

Cantemir, T. (1959). *Texte istroromâne*. Editura Academiei Republicii Socialiste Române.

Cardinaletti, A. (2008). 'On different types of clitic clusters'. In C. De Cat & K. Demuth (eds.), *The Bantu-Romance Connection. A comparative investigation of verbal agreement, DPs, and information structure.* Amsterdam, John Benjamins, 41-82.

Cardinaletti, A. & U. Shlonsky (2004). 'Clitic positions and restructuring in Italian'. *Linguistic Inquiry* 35 (4), 519-557.

Cardinaletti, A. & G. Giusti. 2006.T he Syntax of quantified phrases and quantitative clitics. In M. Everaert & H. van Riemsdik (eds) *The Blackwell Companion to Syntax*. (vol. V). Blackwell. London, 23-93.

Carstens,V. (2000) . 'Concord in minimalist theory'. *Linguistic Inquiry*. 31, 319-355.

Chierchia, G. 1989. *A semantics for unaccusatives and its syntactic consequences*. Ms, Cornell University (published in Alexiadou et al (2004), *The Unaccusativity Puzzle*, OUP).

Chomsky, Noam (1981) – *Lectures on Governemnt and Binding*, Dordrecht, Foris.

Chomsky, N. (1986). *Barriers*, Linguistic Inquiry Monograph 13, MIT Press, Cambridge, MA.

Chomsky, Noam (1995) – *The Minimalist Program*, Cambridge, Massachusetts, MIT Press.

Chomsky, N. (2001). 'Derivation by Phase'. In Kenstowicz, M. (ed.), *Ken Hale. A Life in Language*. Cambridge Mass., MIT Press.

Chomsky, N. (2005). On phases. Ms.

Cinque, G. (1990). 'Agreement and Head-to-Head movement in the Romance Noun Phrase'. Paper presented at the XX Linguistic Symposium on the Romance Languages, University of Ottawa.

Cinque, G. (1994). 'On the evidence of partial N movement in the Romance DP', in G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi & R. Zanuttini (eds.), *Paths Towards Universal Grammar. Studies in Honor of Richard S. Kayne.* Washinghton (D.C.), Georgetown University Press, 85-110.

Cinque, G. (2010). *The Syntax of Adjectives. A Comparative Study*, Cambridge, MA, MIT Press.

Coene, M. (1999). *Definite Null Nominals in Romanian and Spanish. A Generative Approach to the Internal Syntax of DP*. Doctoral dissertation, Universiteit Antwerpen.

Cornilescu, A. (1992). 'Remarks on the determiner system of Rumanian: the demonstratives *al* and *cel*'. In *Probus* 4, 189-260.

Cornilescu, A. (2009).'On the construction of determiner phrases: changes in the pattern of

definiteness checking. In *Grammaticalization and pragmaticalization in Romanian,* 2009.

Cryssman, Berthold (1997) Cliticization in European Portugues using parallel morpho-syntactic constraints. In Butt, M. and Y. Halloway King (eds), Porceedings of the LFG97 Conference, CSLIPublications, availableat http://csli-publications.stanford.edu/LFG/2/crysmann-lfg97.pdf.

Dahmen, W.(1989). 'Rumänisch: Areallinguistik III. Meglenorumänisch', in *LRL*, III. 423-435.

Dimitrova-Vulchanova, M. & V. Vulchanov (2008). 'An article evolving: The case of Old Bulgarian'. In Jonas, Diane & Steve Anderson. Proceedings of DIGS 8. OUP. Oxford .

Dobrovie-Sorin, C. (1987). *Syntaxe du Roumain. Chaines Thématiques*, Thèse de Doctorat d'État, Université de Paris 7.

Dobrovie-Sorin, C. (1994). *The Syntax of Romanian*, Berlin, Mouton de Gruyter. Dobrovie-Sorin, C. & I. Giurgea (2006). 'The Suffixation of Definite Articles in Balkan Languages'. In *Revue roumaine de linguistique*, LI, 1, 73–104.

Dobrovie-Sorin (2001)

Du Nay, A. *The Origins of the Rumanians: The Early History of the Rumanian Language*; Matthias Corvinus Publishing, Toronto.

Dragăn, R. (2002). 'Issues with Adjectival Participles in English and Romanian'. *Bucharest Working Papers in Linguistics* IV (1), 107-112.

Embick, D. & R. Noyer (2001). 'Movement Operations after Syntax'. *Linguistic Inquiry* 32, 555-595.

Filipi, G. (2002). *Istruruminsjki lingvistički atlas / Atlasul Lingvistic Istroromân / Atlante Linguistico Istrorumeno*. Pula/Pola: Mediteran Pula.

Fischer, S. & A. Alexiadou (2001). 'On Stylistic Fronting: Scandinavian vs. Romance'. *Working Papers in Scandinavian Syntax* 68, 117-145.

Franco, I. (2009). *Verbs, Subjects and Stylistic Fronting. A comparative analysis of the interaction of CP properties with verb movement and subject positions in Icelandic and Old Italian*. Unpublished doctoral dissertation, University of Siena.

Frăţilă, V. (2009). 'The origin of the Istro-Romanians: certainties and conjectures', abstract of the talk given at *The Days of the Vlåski and Žejanski language,* Žejan.

Friedman, V. (1994). 'Variation and Grammaticalization in the Development of Balkanisms'. *CLS 30 Papers from the 30th Regional Meeting of the Chicago Linguistic Society, Volume 2: The Parasession on Variation in Linguistic Theory.* Chicago: Chicago Linguistic Society, 101-115. (Reprinted in Studia Albanica, Vol. 32, 1995-99, 95-110).

Giurgea, Ion (2010) The Romanian Verbal Cluster and the Theory of Head Movement. LSRL, Bucharest, Editura Universităţii Bucureşti.

Giurgea, I. & E. Remberger (2009). 'Postverbal subjects in Romance null-subject languages. Informationstructural conditions and variation'. Talk given at Going Romance, Nice, December 3-5.

Giusti, G. (1992). 'Enclitic artiche and double definiteness: a comparative analysis of nominal structure in Romance and Germanic'. In *University of Venice Working Papers in Linguistics* 3, Nr. 1, 83-94.

Giusti, G. (1993). *La sintassi dei determinanti.* Unipress, Padova.

Giusti, G.. (1995). 'A Unified Structural Representation of (abstract) Case and articles. Evidence from Germanic'. In Haider, Hubert, Susan Olsen, Sten Vikner (eds.) *Studies in Comparative Germanic Syntax.* Kluwer, 77-93.

Giusti, G., (1997). 'The categorial status of determiners', in Haegeman, L. (Ed.), The New Comparative Syntax, Longman, London, 95–123.

Giusti, G. (1998). 'The rise of a functional category. From Latin ILLE to the Romance article and personal pronoun'. *University of Venice Working Papers in Linguistics* 8 (2), 53-71.

Giusti, G. (2001). 'The birth of a functional category: From Latin ILLE to the Romance article and personal pronoun'. In Cinque, Guglielmo and Giampaolo Salvi (eds.) *Current studies in Italian Syntax. Studies in Honor of Lorenzo Renzi*, 157-171.

Giusti, G. (2006). 'Parallels in clausal and nominal periphery'. In Frascarelli, Mara (ed.). *Phases of Interpretation*. Mouton de Gruyter. Berlin, 151-172.

Giusti, G. (2008). 'Agreement and Concord in Nominal Expressions', in C. De Cat and K. Demuth (eds.), *The Bantu-Romance Connection. A comparative investigation of*

verbal agreement, DPs, and information structure, John Benjamins, Amsterdam, 201-237.

Giusti, G. & N. Leko (2004). '*T*he categorial status of quantity expressions'. In Leko N., *Linguisticki Vidici,* Sarajevo, Forum Bosniae, 121-184.

Grimshaw, J. (1982). 'On the Lexical Representation of Romance Reflexive Clitics". In: Bresnan (ed.) 1982. 87-147. Hale, Kenneth - Keyser, Samuel J. (eds.) 1993. The View from Building 20. Cambridge - London : The MIT Press.

Grimshaw, J. (1990). *Argument Structure*, Cambridge, Mass: MIT Press.

Grohmann, K. & P. Panagiotidis (2004). 'Demonstrative Doubling in Greek'. *University of Maryland Working Papers in Linguistics* 13, 109-131.

Grohmann, K. (2003). *Prolific Domains. On the Anti-Locality of Movement Dependencies.* Amsterdam. John Benjamins.

Grosu, A. (1994). *Three Studies in Locality and Case.* New York: Routledge.

Hale, K. & S. J. Keyser (1993). *Argument Structure and Lexical Expression of Syntactic Relations*, in Hale, K. and S.J. Keyser (eds.) *The View from Building* 20, pp. 53-100, MIT Press.

Hale, K. & S. J. Keyser (2002). *Prolegomenon to a Theory of Argument Structure*, MIT Press.

Halpern, A. 1995. *On the Placement and Morphology of Clitics*. Stanford, CA: CSLI.

Hawkins, J. A. (2004). *Efficiency and Complexity in Grmmars*, Oxford, OUP.

Heine, B. & T. Kuteva, (2006). *The Changing Languages of Europe*, Oxford, OUP.

Hill, Virginia (2006) Stylistic inversion in Romanian. Studia linguistica 60(2), 156-180.

Hock, H. H. (1991). *Principles of Historical Linguistics,*.2nd edition, Mouton de Gruyter

Horrocks, G. & M. Stavrou. (1987). 'Binding theory and Greek syntax: evidence of wh-movement in NP'. *Journal of Linguistics*. 23, 79-108.

Halle, M. & Marantz, A. (1993), "Distributed Morphology and the Pieces of Inflection", *The View from Building 20* (Cambridge, MA: MIT Press), 111–176.

Iordan, I. & M. Manolriu (1965). *Introducere în lingvistică romanică,*Bucureşti, Editura Didactică şi Pedagogică.

Joseph, B. (1999). 'Romanian and the Balkans: Some Comparative Perspectives'. http://www.ling.ohio-state.edu/~bjoseph/publications/1999roma.pdf.

Joseph, B. (2001). Is Balkan Comparative Syntax Possible?, in Rivero, M. and A. Ralli (eds) *Comparative Balkan Syntax*, Oxford University Press.

Kayne, Richard (1981) Romance clitics, Verb movement, and PRO. Linguistic Inquiry 22: 647-686.

Kolliakou, D. (1995). *Definites and possessives in Modern Greek: An HPSG syntax for noun phrases.* Unpublished Ph.D. dissertation, University of Edinburgh.

Kovačec, A.(1962). 'Notes sur des formes de cas en istroroumain'. *Studia Romanica et Anglica Zagrabiensia* 13-14, Zagreb, 74-84.

Kovačec, A. (1971). *Descrierea istroromânei actuale,* Ed. Academiei Republicii Socialiste România, Bucureşti.

Kovačec, A. (2009). 'Some facts about Istro-Romanian', handout of the talk given at "The Days of the Vlaški or Žejanski language", Žejane.

Larson, R. (1988). 'On the Double Object Construction'. *Linguistic Inquiry* 19, 335-391.

Lekakou, M. & K. Szendrői (2007). 'Eliding the noun in close apposition, or Greek polydefinites revisited'. In Breheny, R. & N. Velegrakis (eds.), *UCL Working Papers in Linguistics* 19, University College London.

Lema, José and Maria-Luisa Rivero (1991). Types of Verbal Movement in Old Spanish: Modals, Futures, and Perfects. *Probus* 3, 237-278.

Levin, B. and M. Rappaport Hovav (1995) *Unaccusativity*, Cambridge, Mass.: MIT Press.

Lombard, A. (1974). *La Langue roumaine, une présentation*. Bibliothèque française et romane, A-29. Klincksieck, Paris.

Longobardi, G. (1994). 'Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form', *Linguistic Inquiry* 25, 609-665.

Lyons, C. (1999). *Definiteness.* Cambridge: CUP.

Marantz, A. (1984). *On the Nature of Grammatical relations*. MIT Press, Cambridge Mass.

Mathewson, L. (2005). 'On the Absence of Tense on Determiners'. *Lingua* 115(12):1697-1735.

Matushansky, O. (2006) Head Movement in Linguistic Theory. *Linguistic Inquiry*, 37: 1, 69–109.

Morariu, L. (1928). Lu frati noştri: libru lu Rumeri din Istrie. Ed. Făt Frumos, Bucuresti.

Montoliu, C. & J. Auwera (2004). 'On Judeo-Spanish conditionals'. In O. Tomic (ed.), *Balkan syntax and semantics*, 461-474. Amsterdam: Benjamins.

Nicolae, A. (2009). 'Aspects on the grammaticalization of Romanian articles'. In Funcţionare discursivă şi gramaticalizare în limba română veche, Universitatea din Bucureşti, http://www.unibuc.ro/ro/nicolae2009_ro.

Panagiotidis, P. (2005). 'Determiner spreading as DP-predication', ms. Cyprus College.

Perlmutter, D. M. (1978). "Impersonal passives and the Unaccusative Hypothesis". *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society*. UC Berkeley, 157–189.

Pesetsky, David (1995). *Zero Syntax – Experiencers and Cascades*, Cambridge, Mass.: MIT Press.

Progovac, L. 1996. 'Clitics in Serbian/Croatian: Comp as the Second Position'. In: A. Halpern and A. Zwicky (eds), *Approaching Second: Second Position Clitics and Related Phenomena.* Stanford, CA: CSLI, 411-428.

Progovac, L. (1998). 'Determiner Phrase in a Language without Determiners', *Journal of Linguistics* 34(1), 165-79.

Puşcariu, S. (1906), and coll. M. Bartoli, A. Belulovici and A. Byhan. *Studii istroromâne*, vol. I, Cultura Naţionala, Bucureşti.

Puşcariu, S. (1926), and coll. M. Bartoli, A. Belulovici and A. Byhan. *Studii istroromâne*, vol. II, Cultura Naţionala, Bucureşti.

Puşcariu, S. (1929), and coll. M. Bartoli, A. Belulovici and A. Byhan. *Studii istroromâne*, vol. III, Cultura Naţionala, Bucureşti.

Radanović-Kocić, V. (1988). *The Grammar of Serbo-Croatian Clitics: A Synchronic and Dyachoric Perspective.* PhD dissertation, University of Illinois, Urbana-Champaign.

Reinhart, T. (1996). 'Syntactic effects of lexical operations: Reflexives and unaccusatives', OTS *Working Papers*, Utrecht University.

Reinhart, T. & T. Siloni (1999). 'Against the Unaccusative Analysis of Reflexives'. In A. Alexiadou et al (eds.) *The Unaccusativity Puzzle: Studies on the syntax-lexicon interface*. Oxford University Press.

Rivero, Maria-Luisa (1994). Clause Structure andV-movement in the Languagesof the Balkans. Natural Language and Linguistic Theory 12, 63-120.

Roberts, I. & Roussou, A. (2003). *Syntactic change: a minimalist approach to grammaticalization.* Cambridge: Cambridge University Press.

Rudin, C. (1988). 'On multiple questions and multiple Wh-fronting.*' Natural Language and Linguistic Theory* 6: 445-501.

Sârbu, R. & V. Frăţilă (1998). *Dialectul istroroman. Texte si glosar.* Ed. Armacord, Timişoara.

Sârbu, R. (1992). *Texte istromâne cu un studiu introductiv "Istroromâna – azi" şi un glosar*, 2nd edition, Ed. Universităţii din Timişoara, Timişoara.

Sârbu, R. (2009). *Noi Rumunski cuvantam*, ms., Universitatea de Vest, Timişoara.

Schütze, C. (1994). 'Serbo-Croatian second position clitic placement and the phonology-syntax interface'. *MIT Working Papers in Linguistics* 21, 373–473

Shokeir, V. (2006). 'Two incompatibile grammars define second position in Serbo-Croatian'. In M. Radisic and C. Gurski (eds), *Proceedings of the 2006 Annual Conference of the Canadian Linguistic Association*. Available at: http://ling.uwo.ca/publications/CLA2006/Shokeir.pdf.

Sobolev, A. N. (2004). 'On the areal distribution of syntactic properties in the languages of the Balkans'. In Tomić, O. *Topics in Balkan Sprachbund Syntax and Semantics*, Benjamins, Amsterdam, 61-100.

Sobolev, A. N. (2008). On some Aromanian grammatical patterns in the Balkan Slavonic dialects. In Sikimić, Biljana and Tijana Ašić (eds) *The Romance Balkans,* Serbian Academy of Sciences and Arts, Institute for Balkan Studies 103, Belgrade, 113-121.

Solta, R. G. (1980). 'Einführung in die Balkanlinguistik: Mit besonderer Berücksichtigung des Substrats und des Balkanlateinischen'. Darmstadt: Wissenchaftliche Buchgesellschaft.'

Sportiche, D. (1998). 'Pronominal Clitic Dependencies'. In H. van Riemsdijk (ed.), *Language Typology: Clitics in the European Languages*. Mouton de Gruyter, Berlin

Szabolcsi, A. (1994). 'The Noun Phrase', in F. Kiefer & K.E. Kiss (eds.), *The Syntactic Structure of Hungarian [Syntax and Semantics 27],* Academic Press: New York, 179-274.

Tagliavini, C. (1969). *Le origini delle lingue neolatine*, Bologna, Pàtron, 293-298.

Terzi, A. (1992). *PRO in Finite Clauses: A Study of the Inflectional Heads of the Balkan Languages*. Doctoral Dissertation, City University of New York.

Tomić, O. (2004). The Balkan Sprachbund properties: An Introduction to *Topics in Balkan Sprachbund Syntax and Semantics*, Benjamins, Amsterdam, 1-58.

Tomić, O. (ed.) (2004). *Topics in Balkan Sprachbund Syntax and Semantics*, Benjamins, Amsterdam.

Ungureanu, M.-L. (2009). *A Syntactic Map of Dps in Romanian. Structure and Movement in the Higher and Lower Domains.* VDM Verlag Dr. Müller, Saarbrücken.

Uşurelu, Camelia (2009) Comportamentul cliticelor in limba romana veche. Universiteatea din Bucureşti, available at http://www.unibuc.ro/ro/usurelu2009b.

Van Valin, R. D. (1990). 'Semantic parameters of split intransitivity'. *Language* 66(2), 221-60.

Vrzić, Z. (2009). 'How to write Vlaški or Žejanski', talk given at "The Days of the Vlaški or Žejanski language", Sušnjevica. Pdf downloadable at www.vlaski-zejanski.com.

Werle, A. 2009, *Word, phrase and clitic prosody in Bosnian, Serbian and Croatian*. Ph.D. dissertation, University of Massachusetts Amherst.

Zamfir, Dana (2007). 'Morfologia verbului în dacoromâna veche', II. Bucharest, Editura Academiei Române.