

Paul Compensis

Definiteness in Molise Slavic

Opera Slavica Coloniensia

Edited by Daniel Bunčić and Jörg Schulte

Volume 23

Cologne 2017

Opera Slavica Coloniensia

Slavisches Institut
Universität zu Köln
Weyertal 137
50931 Köln, Germany

ISSN: 2748-6060

This thesis was supervised by Professor Daniel Bunčić and accepted as a master's thesis by the Faculty of Arts and Humanities of the University of Cologne.

The thesis was awarded the **Reinhold Olesch Prize** for papers on Slavic linguistics (<https://slavistik.phil-fak.uni-koeln.de/studium/auszeichnungen>).

Published with minor corrections
on the Cologne University Publication Server (KUPS)
of the University and City Library of Cologne (<https://kups.ub.uni-koeln.de/>)
in May 2022

© Paul Compensis 2022

This work, including its parts, is protected by copyright. Any use without the author's consent is prohibited. This applies in particular to electronic or other reproduction, translation, distribution and making available to the public.

Znadahu dobra ka,
vana do pjaneti velke,
kana Kaša, Džove, Marte, Venere,
kijimi su dal jena jiman,
jesu stotine do drugihi,
ka jesu koju votu naka male
ka je difičil hi vit s teleskopjam.

Table of contents

List of abbreviations.....	3
List of tables and Praat pictures	4
1. Introduction.....	5
1.1. Aims of the study	5
1.2. Hypotheses.....	7
1.3. Notes on research and corpus	9
2. Complex definites in Molise Slavic	10
2.1. Demonstratives.....	10
2.2. Personal pronouns	15
2.2.1. Meaning and function.....	15
2.2.2. Clitic doubling	18
2.3. Possessives	22
2.3.1. Genitive case	22
2.3.2. do-construction.....	24
2.3.3. Possessive pronouns	25
2.4. Numerals and quantifiers	30
2.5. Adjectives	35
2.6. Case distinction.....	38
3. The article system of Molise Slavic	41
3.1. Simple indefinite: jena	41
3.2. Simple definite: zero article.....	47
4. Topicality and zero articles: The interplay of information structure and definiteness.....	50
4.1. Theoretical deliberations: Referentiality and information structure	50
4.2. Topicality in Molise Slavic	55
4.2.1. The linguistic concept of <i>topic</i>	55
4.2.2. Topic tests	59
4.2.3. Sentence-initial position	65
4.2.4. The role of intonation.....	76
4.2.5. Topicalization strategies	80
5. Conclusion	82
Appendices	85
Appendix 1: Tables of forms.....	85
Appendix 2: Topic analyses	88
Appendix 3: Intonation analyses.....	93
Literature	102

List of abbreviations

!	intonational downstep
#	intonational upstep
%H	high word boundary
%L	low word boundary
>	delayed peak of a pitch
Ø-	global downstep within an intonational phrase
0 (break index)	lack of word boundary (clitic+word)
1 (break index)	word boundary
1 (glossing)	first person
2 (break index)	intonational phrase boundary
2 (glossing)	second person
3 (glossing)	third person
ACC	accusative
ART	article
BCS	Bosnian/ Serbian/ Croatian
DAT	dative
def	definite
DEM	demonstrative
DIST	distal
F	feminine
GEN	genitive
H%	high intonational phrase boundary
H*+L	falling pitch accent
HiF0	highest pitch in an intonational phrase
IMP	imperative
INDF/indef	indefinite
INF	infinitive
INS	instrumental
INTERJ	interjunction
IPFV	imperfective aspect
IPRF	imperfective
IS	information structure
L%	low intonational phrase boundary
L*+H	rising pitch accent
M	masculine
MED	medial
MS	Molise Slavic
N	neuter
NOM	nominative
PFV	perfective aspect
PL	plural
PROX	proximal
PRS	present

PTCP	participle (active)
PTCP.PASS	participle (passive)
REFL	reflexive
REL	relative
SAEL	Standard average European languages
SC	sub-ordinate clause
SG	singular
spec	specific
ToBI	tones and break indices (annotation)
WALS	World atlas of language structures online

List of tables and Praat pictures

Table 1	Types of givenness categories	75
Table 2	Molise Slavic Orthography	85
Table 3	Corpus frequencies of demonstratives	85
Table 4	Full paradigm of personal pronouns	86
Table 5	Full paradigm of possessive pronouns	87
Text 1	Topic analysis <i>The little prince</i>	88
Text 2	Topic analysis <i>The Russian tractor</i>	90
Text 3	Topic analysis <i>The hare and the crayfish</i>	91
Praat picture 1	Pitch accent of the numeral jena	94
Praat picture 2	Pitch accent of the indefinite article	94
Praat picture 3	Declarative sentence with focus (<i>vaza</i>) and post-focal pitch reduction	95
Praat picture 4	Declarative sentence with focus (<i>bolje</i>) and post-focal pitch reduction	95
Praat picture 5	Topical focus (<i>zec</i>) with gradual pitch reduction	96
Praat picture 6	Intonation of the narrative inversion	96
Praat picture 7	Bridge intonation (shifted topic)	97
Praat picture 8	Bridge intonation (listing)	97
Praat picture 9-10	Declarative sentences at the beginning of discourse	98
Praat picture 11	Declarative sentence with pronominal topic	99
Praat picture 12	Declarative sentence with subject drop	99
Praat picture 13-14	Declarative sentence with right dislocation	100
Praat picture 15	Declarative sentence at the end of discourse	101

1. Introduction

1.1. Aims of the study

Slavic languages (except Bulgarian and Macedonian) fall within the context of Standard Average European Languages (SAEL) for not using articles for the marking of (in)definiteness on a nominal phrase. Areal-typologically, they belong to:

the eastern area without articles including West Slavic, East Slavic, Baltic, and Balto-Finnic standard languages, as well as Standard Slovene and Standard Serbian and Croatian. (Schroeder 2006: 585).

Leaving aside discussions in Slavic linguistics, none of these languages has a proper article. It is striking that at least for one Slavic language, namely Molise Slavic, claims are made that it displays both definite and indefinite articles. For the purpose of this study, I understand *definiteness* in its simplest way based on the status of referents in communication. I take up the view of Trenkic, who defines the difference between the two concepts regarding who has previous knowledge of the entity:

The crucial difference that separates definite from specific [...] is to whom something is identifiable: to both the speaker and the hearer (definite), or just the speaker (specific). (Trenkic 2004: 1406).

This notion is refined in more detail in chapter 4.1. In example 1.1-1, Peter introduces a discourse referent he knows well. This referent is therefore specific at the moment of the utterance. The specificity of the referent is also clear to the second interlocutor, Mary. Once the referent is introduced to the discourse and both hearers have some knowledge of him, he can be considered definite. This is expressed overtly by the use of a demonstrative (*this friend*) or a personal pronoun (*him*):

- 1.1-1 Peter: A friend of mine is going to New York next week.
 Mary: Where do you know this friend from?
 Peter: I know him from university.

Molise Slavic is a micro-language of Slavic descentance spoken only in three villages in the province Molise in Southern Italy. Belonging to the štokavian varieties, the language is related to nowadays Standard Croatian and Standard Serbian, but differs from them due to three main factors. Whereas the modern standards are based on štokavian varieties with the two accepted forms -ijekavian and ekavian – Molise Slavic is ikavian. Furthermore, čakavian features can be found in this variety. Additionally, the long-lasting contact with Italian varieties led to contact-induced convergences at all levels of the language system. The language was first described

extensively by Rešetar (1911) and later by Walter Breu. Breu researched the neighbouring Albanian villages in Molise in the late-1970s and later directed his attention at the Slavic communities of this area. He recorded a large amount of spoken material and published a dictionary in collaboration with Giovanni Piccoli, mainly covering the variety of the largest Molise Slavic village, Acquaviva Collecroce.

One aspect not covered sufficiently in the literature is the question of how definiteness and indefiniteness can be expressed in Molise Slavic. Indefiniteness is expressed with an article, while no overt article exists for the expression of definiteness. This is not problematic from a genetic perspective, because Serbian and Croatian also do not display a definite article. By contrast, the contact language Italian uses both definite and indefinite articles. In contact linguistics, contradicting claims are made about how prone to borrowing the category of definiteness is in general. From a typological perspective, it is untypical in Europe to find an overt marking of indefiniteness without the necessity to specify definiteness on a nominal phrase. Breu postulates the existence of a fully-grammatised zero article in Molise Slavic to account for this problem. However, the concept of zero morphemes is problematic and examples can be found in which his interpretation is not consistent. Eventually, we are touching upon a complex field in linguistics to which the case of Molise Slavic can provide important insights:

One of the most vexed questions in all of pragmatics is the question of what it means for a noun phrase to be definite. (Birner 2013: 121).

Definiteness is both a grammatical as well as a semanto-pragmatic category. While the former is a grammatical category that some languages have and others do not, the latter is probably universal across all languages of the world:

[...] definiteness as a meaning category is an element of interpretation in all languages: it can be part of lexico-semantic content of certain expressions, and it can also be inferred through some general principles of goaloriented behaviour. (Trenkic 2004: 1402).

Considering this aspect, the question of this study is not whether Molise Slavic has the right means to express definiteness at all, but rather if it has specific markers and strategies overtly indicating whether a given nominal phrase is definite or not. The aim of this study is two-fold. On the one hand, I aim to provide a systematic overview of those modifiers that can impose definiteness on a nominal phrase. These markers

are also investigated regarding typological peculiarities. It is clear from comparative research on other Slavic languages that there is a huge variety of strategies how to express definiteness:

[...] article functions in Slavic, so far as they are expressed at all, are rendered by demonstratives or the numeral/ indefinite pronoun *ONE*. They also coincide in part with secondary functions of the theme-rheme structure or with verbal aspect." (Breu 2012: 275).

More abstract categories like word order and intonation are also sometimes claimed to impose a definite reading. In the second part, definiteness is seen as a semanto-pragmatic category that can be expressed by other means than morphosyntactic ones. I investigate particularly whether topic positions and topicalization strategies in Molise Slavic can enforce a definite interpretation of a NP in a sentence without necessarily being marked in the morphological sense. This study is the first attempt to describe aspects of Molise Slavic information structure. Indefiniteness is only secondary to the study. We are largely dealing with an issue concerning the nominal part of a sentence. The question of whether verbal aspect can impose a definite reading on its arguments and other features of the verbal system are not discussed.

1.2. Hypotheses

In my analysis, I follow the distinction by Lyons (1999) into simple definites (articles) and complex definites (modifiers, pronouns). On the one hand, an account of the complex definites is first provided. The second part of the study comprises a more theoretical discussion of the question whether Molise Slavic really makes use of a definite article in the form of a zero article or if the concept is redundant because the category under discussion can be explained by other factors, in particular information structural categories. This study follows the main hypothesis (H1) that the definiteness of a referent can be explained either by the presence of modifiers or by information structure.

H1 The definiteness of a nominal phrase is expressed either by using (definite) modifiers or by information structural processes (word order, intonation, topicalization).

If this hypothesis is incorrect, it means that other factors not covered in this study are responsible for the definite status of a NP. In this case, Breu's assumption of a zero article is a likely candidate for a better explanation. Since two distinct possibilities are summarised in the main hypothesis, two additional hypotheses are formulated.

Hypothesis 2 predicts that certain modifiers exist that can trigger a definite interpretation of a nominal phrase. It is not expected that the whole hypothesis is true or false, but rather that it is true only in case of some modifiers. A number of modifiers are therefore tested in chapter 2 regarding the question whether they can impose definiteness on a NP. At some points in the exploration, corpus analyses are used. Most features are discussed qualitatively with examples from the corpus. Where necessary, quantitative measures of frequencies are added.

H2 Certain modifiers are responsible for the definiteness of a nominal phrase.

The third hypothesis aims to find an alternative explanation concerning how bare nouns without further marking or modification can be understood as being definite. I claim that information structure might be responsible for this, because this effect can be observed in other Slavic languages.

H3 A non-modified (bare noun) NP can be definite as a result of information structural processes, in particular the role of topic-comment structure.

This is tested from three different perspectives. Topic marking may occur as a result of word order (H3a), intonation (H3b) or other topicalization strategies (H3c).

H3a The responsible process mentioned in hypothesis 3 is achieved by word order.

H3b The responsible process mentioned in hypothesis 3 is achieved by intonation.

H3c The responsible process mentioned in hypothesis 3 is achieved by other topicalization strategies.

Hypothesis 3 is tested in the second part of this study. A topic test is developed in chapter 4.2.2 and applied to three Molise Slavic texts. A topic position is identified and its relation to definiteness is discussed. Additionally, a quantitative approach to further controlling this hypothesis is chosen. An intonational analysis is presented to control hypothesis 3b. Other topicalization strategies were identified during the exploration of the corpus, although due to limitations in space not every aspect could be investigated to the same extent. Some features and functions could not be investigated sufficiently due to the lack of data. Other claims need to be confirmed by experimental tasks. In summary, I present a systematic overview of modifiers that can be used for the expression of definiteness in Molise Slavic. Furthermore, I present the first account of the information structure of Molise Slavic, which remains sketchy. For the description of intonational aspects of information structure, a first proposal for a

ToBi annotation convention for Molise Slavic is developed. In this regard, this study can serve as a starting point for research in three directions.

1.3. Notes on research and corpus

Molise Slavic is a Slavic micro-language in Italy. It is closely related to istrian-ikavian dialects (in southern and western parts of Istria and few villages in Čićarija). Molise Slavic is positioned between štokavian and čakavian (Breu 1990: 39), but exhibiting mainly štokavian-ikavian features with some material of čakavian origin. The štokavian origin is apparent in the development of the phonemes [st] and [sk] > [št] and [žd] (and not [šč] and [žď]) in the typical contexts (see Rešetar 1911: 141–142). Čakavian features are the occasional development of [dj] into [j] and the preservation of proto-slavic reduced vowels or alternatively the realisation as [a] as well as single words like *crikva* ‘church’ instead of the regular štokavian *crkva*. The pronunciation is strongly influenced by Italian. Stressed syllables are heavily stressed, while unstressed syllables are reduced. This change has led to complications with the (unstressed) case endings (see already Rešetar 1911: 142; Breu 1990: 47). All examples in the present paper are given in the dialectal form of Acquaviva Collecroce.

This study uses data from three sources. Breu, Piccoli (2011 and 2012) are printed editions of text collections, recorded by the two in Molise. The first volume – containing material from Acquaviva Collecroce – was copied into a text file for further investigation with the text editor *notepad ++* and the corpus analysis tools *AntCon* and *AntPCon*. The corpus comprises 60,219 tokens and 7,023 types (each case variant of a noun is counted as an own token). This corpus is mainly used for the exploration of modifiers in chapter 2. Every example from this corpus is quoted with reference to the chapter of the text and the number of the sentence. For instance, [3.1.4]-35 refers sentence 35 in chapter 3.1.4. For another analysis in chapter 4, the corpus is separated into single sentences and reduced in size. In addition, the text collection Breu (2017) is used for the analysis of information structure. Two texts from this source are analysed regarding their information structure. All texts from Breu (2017) are available online in the Pangloss collection, an open-access repository for linguistics field data (<http://lacito.vjf.cnrs.fr/pangloss/index.html>). All recordings and morphosyntactic

analyses by Breu can be downloaded there. The phonetical analysis tool Praat is used to investigate pitch values in selected utterances from the texts. Pitch settings were only adjusted to the speakers' voice, whereas no further changes of the standard settings were carried out. All Molise Slavic examples are written in the orthography developed in Breu, Piccoli (2000, 2011 and 2012) and Breu (2017), which is mainly based on Croatian orthography. The orthography and possible phonetic realisations of its letters are provided in table 1 in the appendix. For the glosses, I follow the glossing rules applied by Breu (2017), based on the Leipzig Glossing Rules. Occasionally, examples from other languages are used in this study. All examples are given in either their original orthography or scientific transliteration. When I refer to linguistic features present in the Standard languages based on what was formerly known as Serbo-Croatian, I simply speak of BCS. All examples of BCS are given in their Croatian form for practical reasons.

2. Complex definites in Molise Slavic

2.1. Demonstratives

Our exploration begins with demonstratives because they undisputable constitute a definite category by its core meaning:

[...] [+ Dem] is always accompanied by [+ Def], apparently because demonstrativeness is semantically incompatible with indefiniteness. (Lyons 1999: 25).

It is sometimes even argued that languages that do not have a definite article use demonstratives more frequently. Molise Slavic has three different demonstrative forms, all marked for case, gender and number. Their main function lies in specifying distance, while definiteness is a secondary meaning. Breu assumes a three-term, person-based system in the domain of demonstratives. The first form (*ovi* 'this') denotes referents that are closer to the speaker, the second form (*ta* 'this/that') is used for referents closer to the listener, while in contrast to both *oni* 'that' signals entities far from both the speaker and hearer (see Breu, Piccoli 2000: 405-206). In comparison to the contact language – Italian – the system would be in line with the distinction between *questo* 'this' and *quello* 'that', a two-term distinction based on person, in which the position relative to the interlocutor does not seem to play a major role (see

Reumuth, Winkelmann 2001: 73). This system presumably reflects the three-term distinction in BCS into *ovaj* 'this', *taj* 'this/that' and *onaj* 'that', although it is assumed here that the focus of orientation in the demonstrative system rather lies on distance than person. The distance-based view is described – for example – in the language textbook by Javarek, Sujić (1963). This view is disputed, given that other accounts also identify a person-based system; for example, Kunzmann-Müller (2002: 154). A compromise is taken up in Alexander, who refuses a strict separation of the two:

The relationship among them can be viewed in two different ways – either in terms of distance from the speaker, or in terms of physical space. (Alexander 2006: 12).

Hereby, distance can be viewed from both a speaker-oriented and a physical space perspective, depending on the exact context. In general, the two are not always separable and it is difficult to test the exact point of orientation:

Even with such three-term systems, however, it is not always clear whether distance or person is the principle involved, and some languages may mingle the two. (Lyons 1999: 109).

The distinction is also rather difficult due to the two functions that demonstratives can fulfil. They can be used for deictically for reference to entities in the immediate surrounding of the interlocutors. In example (2.1-1), the little prince is referring to the drawing of a sheep, which is in the immediate proximity of both the speaker and the hearer.

2.1-1 Ti misliš ka **ovu ovcu** ča ju servit čuda trave?

ti.NOM think-IPRF-PRS.2SG REL **DEM.PROX-ACC.SG.F** **sheep-ACC.SG.F** want-PRS-3SG
she.ACC need.IPRF-INF much grass-ACC.SG.F

‘Do you think that **this sheep** needs a lot of grass?’

(Saint-Exupéry 2009: 14)

If reference is made to an entity that was already mentioned in the discourse or is somehow accessible to both the speaker and hearer at the discourse level, demonstratives can facilitate anaphoric reference to this entity. Distance can be expressed at both the anaphoric and deictic levels. In the deictic use, it is oriented in the relative distance of the speaker or hearer. In the anaphoric reference, the concept of distance is rather abstract. The common claim is that the distal forms are used for an entity that was mentioned some time before, while the proximal form refers to something just mentioned or reactivated in the awareness of the interlocutors:

The choice of anaphoric expressions depends on the recency of the antecedent [...]. (Krifka, Musan 2012: 3).

Example (2.1-2) illustrates this difference with reference to the same entity:

2.1-2 È, **oni** ka mi proda mutor men, **ovi** ka mi ga proda je umbra.

INTERJ **DEM.DIST-NOM.SG.M** REL I.DAT sell.PFV.PTCP.SG.M motor.ACC.SG.M I.DAT
DEM.PROX-NOM.SG.M REL I.DAT he.ACC sell.PFV.PTCP.SG.M be.PRS.3SG
die.PFV.PTCP.SG.M

‘**That one** who sold me the motor, **this one** who sold it to me, died.’

(see *The Russian tractor* in the appendix)

The first NP (including a relative clause) introduces the referent into the discourse. The distal form is used to indicate that the referent was mentioned some time ago¹. Since he is now active in the awareness of the hearer, the second reference to the same referent is made with the proximal form.

In the data, we still find examples of every form, although the middle form *ta* is relatively rare. This could provide a hint that this form is about to disappear and the demonstrative system might rather follow the Italian model. Diachronically, a reduction of three-term systems to two-term systems can often be found in cross-linguistic comparison, while the latter also being more common in general (see Lyons 1999: 110-111). This can also be seen in Feature 41A of the World Atlas of Language Structures (WALS), with about 54% of the languages in the sample having a two-way contrast in the demonstrative system (see Diessel 2013). While this reduction took place in the Romance languages, it did not affect the South Slavic varieties. The idea whether Molise Slavic follows this trail cannot be tested here. Alternative explanations for the scarce use of the medial form are possible. The distal form is used the most in the corpus, probably because it is also used as demonstrative for anaphorical reference to something mentioned earlier before or being more distant to the speech time. Perhaps the medial form is simply not suitable for anaphoric use or only appears rarely due to its meaning. All frequencies of occurrences of every demonstrative form can be found in table 2 in the appendix.

Regarding the position within the nominal phrase, languages that exhibit (obligatory) definite articles can be distinguished along another parameter regarding

¹ To be more precisely, the actual referent was not mentioned at all, but the speaker talked about buying a tractor. The referent can be inferred from this.

demonstratives. In some languages, the demonstrative replaces the article (or takes the determiner phrase position in syntactical terms), while in other languages both can co-occur (see Lyons 1999: 121). In case of a zero article, we have no indication where to assume the position of a definite article in the NP. The demonstrative always occurs at the beginning of a NP in Molise Slavic:

2.1-3 Je veča lipa **ova vizita** [...]

be.PRS.3SG more nice-NOM.SG.F **DEM.PROX-NOM.SG.F visit-NOM.SG.F**

‘**This visit** was nicer.’

(Saint-Exupéry 2009: 43)

2.1-4 Jesu kačatura zgora **onga pjaneta**?

be.PRS-3PL hunter-NOM.PL.M on **DEM.DIST-GEN.SG.M planet-GEN.SG.M**

‘Are there hunters on **this planet**?’

(Saint-Exupéry 2009: 68)

This position is also preserved when no noun is present in the phrase (2.1-5) or when other modifiers co-occur (2.1-6):

2.1-5 [...] kaka je **ovi moj**, ma bit pur **ta tvoj** !

how be.PRS-3SG **DEM.PROX-NOM.SG.M my-NOM.SG.M** but be-INF also **DEM.MED-NOM.SG.M your.NOM.SG.M**

‘How **mine** is, must also be **yours**!’

(Breu 2011: [1.2.2]-381)

2.1-6 Je rispunija **oni drugi karbunir** [...].

be.PRS-3SG respond.PFV-PTCP.SG.M **DEM.PROX-NOM.SG.M other-NOM.SG.M policeman.NOM.SG.M**

‘**This other policeman** responded [...].’

(Breu 2011: [3.1.4]-35)

Demonstratives can also be used as free-standing forms in predicative use replacing a whole nominal phrase:

2.1-7 **Ova** je dža čuda razboljana.

DEM.PROX-NOM.SG.F be.PRS.3SG already much sick-NOM.SG.F

‘**This (one)** is already very sick.’

(Saint-Exupéry 2009: 14)

2.1-8 **To** je mi surtila veramend [...] !

DEM-MED.NOM.SG.F be.PRS.3SG I.DAT happen.PFV-PTCP.SG.F really

‘**This** really happened to me!’

(Breu 2017: [3.3.6]-31)

There are shortened versions of *ovi* with an omission of the initial vowel (cf. Breu, Piccoli 2000: 405-406). This is illustrated in example (2.1-9):

2.1-9 Jesa žedan do **ve vode**.

be.PRS-1SG thirsty-NOM.SG.M for **DEM.PROX-GEN.SG.F water-GEN.SG.F**

'I am thirsty for **this water**.'

(Saint-Exupéry 2009: 80)

In the literature, sometimes the claim is made that demonstratives occur more frequently in languages that do not display a definite article (see Trenkic 2004). The idea behind this claim is that the form somehow takes over some of the functions that the article would express. Trenkic (2004) conducted a small corpus analysis with a small sample of 50,000 words each in English and Serbian, whereby both corpora were extracted from internet newspapers. She found out that demonstratives occurred at a frequency of 0.43% in Serbian and 0.23% in English, contrasted to 7.32% occurrences of the definite article in English. Admittedly, her corpus is not large for those two languages and the text type chosen rather specialised. If we compare our data set with her findings, we must take into consideration the different text types underlying the corpora. Nonetheless, a first rough estimate should be possible to falsify the claim. In the MS corpus, all demonstratives together occurred with a total frequency of 3.95% (2,381 out of 60,219 tokens). It seems that the demonstratives occur more frequently than in the Serbian corpus, but still much less than the definite article in English. This requires more detailed study of similar text types as well as more fine-grained analyses of actual occurrences and functions of definite articles and demonstratives. For now, at least we can conclude that the demonstratives are neither extraordinarily frequent nor obligatory and it does not yet seem that they are about to grammaticise into a definite article. Regarding position and main function, the demonstratives in Molise Slavic are quite comparable to those in other European languages without extraordinary features. In preserving the three-term system, it also reflects the same system known from other South Slavic languages.

2.2. Personal pronouns

2.2.1. Meaning and function

Personal pronouns are typically not used as modifiers², but are used in place of a full NP. They do not have a fixed lexical meaning. In her account, Burkhardt (Schumacher) notes that:

[p]ronominal entities [...] are referentially dependent elements, as their interpretation is not sufficiently determined by their lexical content [...] and as a consequence, pronominals must rely on a dependency with another entity (i.e. antecedent) to obtain referential content or as is the case of indexicals (e.g. I, you), they select a discourse referent from the context of utterance. (Burkhardt 2005: 1).

In some reference works, the forms of 1st and 2nd person pronouns are called *deictics* because they can only be defined in context and never refer to a unique entity. They are determined by who is the speaker and who is the hearer of an utterance. Hence, the roles may change several times in discourse. 3rd person pronouns are anaphorical by their meaning, because they depend on contextual factors. Both types of personal pronouns share the feature that they can be considered definite and are therefore occasionally even called definite pronouns in the literature (see Lyons 1999: 26). This derives from the fact that they must refer to an antecedent that was already mentioned in the discourse or locally present referent in the case of deictics. As Lyons put it, they can be considered the “pronominal counterpart of definite articles, and sometimes of demonstratives” (Lyons 1999: 134). This view goes back to a syntactical account by Postal (1970), who “proposes to account for the definiteness of personal pronouns by deriving them transformationally from definite articles” (Lyons 1999: 26-27). He argues that (in English) 3rd person pronouns can never be used pre-nominally. In this, the anaphorical personal pronouns is syntactically similar to definite articles by having only one context in which they can be used. In the case of the latter, it can only function as determiner in a NP, while the former only occurs pronominally (Lyons 1999: 26-27). Lyons modifies this analysis and argues that “the English personal pronouns are forms both of the definite article and of the [-Prox] demonstrative”

² Although in English 1st and 2nd person plural pronouns can be used pre-nominally in phrases such as “We Europeans” or “you politicians” (Lyons 1999: 27).

(Lyons 1999: 28). In my view, the definite reading of pronouns derives from the pragmatic role that they play in context.

The system of personal pronouns in Molise Slavic is quite comparable to the system in modern BCS and from a typological perspective the MS personal pronouns system reflects the typical pattern found in most European languages, distinguishing three persons and two numbers, without further distance distinction. The dual as number – as can be found in Slovenian and Old Church Slavonic – is lost and gender is only marked in 3rd person singular and only in a binary sense, due to the loss of grammatical neuter forms. The full paradigm is given in table 3 in the appendix. Just like in other South Slavic languages, there are two full sets of pronouns: one paradigm comprises long, stressed forms and the other is made up by phonetically weakened, short clitics. Both forms may occur alone (2.2-1, 2.2-2, 2.2-3) (see also Breu 2017: 39) or together (2.2-4).

2.2-1 **On** biša moj mekanik! **Ja** bihu manager [...].

he.NOM be-IPRF-3SG my-NOM.SG.M mechanic-NOM.SG.M/ I.NOM be-IPRF-1SG manager-NOM.SG.M

‘**He** was my mechanic. **I** was (the) manager [...].’

(Breu 2011: [3.4.1]-30)

2.2-2 Bihu... one dva soldate ka bihu kurta **njega**.

be-IPRF-1SG DEM.DIST-NOM.PL two.M soldiers-NOM.PL REL be-IPRF-1SG beside **him-GEN.SG.M**

‘There were ... those two soldiers who were next to **him**.’

(Breu 2011: [1.2.1]-106)

2.2-3 **Ga** zovahu profesor.

he.ACC call.IPFV-IPRF.3PL professor.NOM.SG.M

‘They called **him** profesor.’

(Breu 2011: [1.2.1]-304)

2.2-4 Ja što mahu **ti** jimat čit **teba**?

I.NOM what can-IPRF-1SG **you**.ACC have.IPFV-INF make-INF **you**.ACC

‘What should I do for **you**?’

(Breu 2011: [3.6.4]-21)

In contrast to BCS (but similar to Slovenian and Macedonian), the clitic forms can appear sentence-initially. They seem to follow slightly different cliticisation rules than

their counterparts in BCS. While the sentence-initial position of clitics is excluded in BCS, it is well known from Southern Italian dialects (2.2-5b).

2.2-5a **Ga** vidahu Matija.

he.ACC see.PFV-IPRF-1SG Matija.NOM.SG.M

'I saw Matija.'

(Breu 2011: [3.5.4]-15)

2.2-5b **L'** aggiu visto a Giuanne. (Napolitanian)

he.ACC have-PRS.1SG see-PTCP.SG.M PREP Giuanne.ACC

'I have seen John.'

(Roberts 2006: 789)

In general, this requires more detailed study. The clitic forms never occur sentence-finally; rather, only in a few constructions when they appear as a one-word-utterance. In all cases, the clitic occurs pre-verbally. The long form appears mainly sentence final or later in the sentence. Both observations only apply to case forms other than nominative since this case does not have clitics. The clitics tend to be used more often than the full forms. For instance, *njoju* only appears 6 times in the corpus, while *ju* can be found 209 times. In the case of *njega*, only 30 instances were found, while its clitic counterpart *ga* shows up 500 times. One could argue that the full form occupies the normal position of the direct or indirect object since it pronominally replaces it. However, there are also examples in which the clitic form takes the same position as other object forms. In example 2.2-6a, this is represented by a demonstrative:

2.2-6a A **ono** držaša napo moj sfekar.

and **DEM.DIST-ACC.SG.N** keep.IPFV-IPRF-3SG shared my-NOM.SG.F father-in-law.NOM.SG.F

'My father-in-law farmed **this** proportionately.'

(Breu 2011: [1.2.2]-283)

2.2-6b E **ju** držaša moj sfekar.

and **it.ACC** keep.IPFV-IPRF- father-in-law.NOM.SG.F

'And my father-in-law farmed **it**.'

(Breu 2011: [1.2.2]-283)

Here, it is important to note that example 2.2-6 does not reflect the neutral word order, but rather a type of inversion because the object appears before the subject in both cases. The issue of topicalization and highlighting will be discussed in further detail in the second part of this study. We conclude that there is a functional and positional

difference between the two forms. This is in so far expectable that clitics are usually characterised by differing from their long forms in terms of phonetical properties and syntactical position:

Object clitics are unstressed forms of pronouns in direct object function. They are ‘special clitics’ [...] in that they generally occupy a position distinct from that of a stressed object pronoun or a nonpronominal direct object (cf. §48.2.1). (Roberts 2016: 786)

2.2-4 was a first example of clitic doubling, a construction that is rare in BCS (except some dialects), but very frequently used in Italian as well as Bulgarian and Macedonian. This construction can also shed light on the function of both the clitic and the long form of pronouns.

2.2.2. Clitic doubling

Clitic doubling (sometimes called *object doubling*) is a construction in which a direct or indirect object is doubled with a clitic form of a pronoun. This may occur typologically in two forms: sometimes the full noun is doubled with a clitic, while in other languages the long form of the pronoun is accompanied by a clitic. Some languages exhibit both strategies, while others only have one or the other type of clitic doubling. In some languages, this might be an obligatory process, dependent on definiteness (Macedonian), topicality (Bulgarian) or else, while in other languages it can be a facultative construction for emphasis (see Tomić 2008). In addition, there is sometimes a large difference (at least in some languages) in doubling a direct object or an indirect object with a clitic. For the sake of brevity, no further discussion of the theoretical issues with the construction can be given here. There are different factors causing the doubling in different languages. In the South Slavic context, two determining factors are known: while in the written standard of Macedonian doubling is mandatory with direct objects, in Bulgarian it is largely driven by information structure, in particular the topicality of a referent. It is one of the features of the Balkan sprachbund but it only affected the BCS language area in the periphery (see Tomić 2008: 462-464). In the case of Molise Slavic, it cannot be explained by Balkan linguistic convergence. Clitic doubling is also prominent in Romance languages. The extensive clitic system even characterises this language sub-family:

In many respects, the complex and intricate pronominal clitic systems of medieval and modern Romance constitute a ‘signature’ morphosyntactic property of this family, such clitics being absent or radically different in typical distribution patterns in the other branches of Indo-European.” (Roberts 2016: 786)

Italian – and particularly Southern Italian dialects – make extensive use of clitic doubling. Breu notes this at several points in his extensive work on Molise Slavic and highlights that the construction in MS perfectly replicates the Italian or Southern Italian dialectal model:

Besonders hochfrequent ist das - wie auch allgemein im Italienischen - bei der Stellung des Objekts vor dem Verb. Verdopplung kommt aber auch bei der Normalwortstellung Subjekt-Verb-Objekt vor, was im Standarditalienischen nicht der Norm entspricht, aber durchaus zu den Verhältnissen in den süditalienischen Regionalvarietäten stimmt. (Breu 2017: 39)

In Italian, the standard pattern for clitics is OV and the clitic being in enclisis to finite verb (see Roberts 2016: 786). The last examples show that the doubling might cause the object to move to a pre-verbal position. This is an indicator that topicalization is connected to the process of doubling. Rešetar already mentions instances of clitic doubling in his corpus (Rešetar 1911: 233):

Ziemlich häufig wird im Satze ein durch ein Pronomen ausgedrücktes Objekt wiederholt, indem neben einer vollen Form auch eine enklitische erscheint: *ńěga su-ga-ūbil; ònôj maše ju-priséć, ju-je-bušila ńôj* B.26, was auf italienischen Einfluß zurückzuführen ist [...]. (Rešetar 1911: 233)

The occurrence of clitic doubling is very frequent in Molise Slavic, whereby both types can be found (see also Breu 2017: 41). 2.2-7 and 2.2-8 illustrate the doubling of full nouns that are doubled with a clitic and 2.2-9 and 2.2-10 exemplify the doubling of long pronouns.

- 2.2-7 Pa, one vadu nu vodu, **nu vodu** ti **ju** hitaš.
 then DEM.DIST-NOM.PL eject.IPFV-PRS.3PL DEM.DIST-ACC.SG.F water-ACC.SG.F
DEM.DIST-ACC.SG.F water-ACC.SG.F you.SG.NOM **she.ACC** pour.away.IPFV-PRS-2SG
 ‘Then they release their water, you pour **this water** away.

(Breu 2017: 75)

- 2.2-8 [...] ka **tvoju nevistu ju** ne moram či nišča [...].
 [...] REL **your-ACC.SG.F daughter-in-law-ACC.SG.F she.ACC** NEG can-PRS.1SG make-INF
 nothing
 ‘[...] that I cannot do anything to your **daughter-in-law** [...].’

(Breu 2011: [2.5.1]-17)

2.2-9 e **njihi** maš **hi** šuši utra na... na strofinač
 and **they-ACC** must-PRS-2SG **they.ACC** dry-INF inside ART.INDF-ACC.SG.M ART.INDF-ACC.SG.M cloth.ACC.SG.M
 and you have to dry **them** with a ... a cloth

(Breu 2017: 75)

2.2-10 Ti, pa, onu vodu maš hitit e **njihi** maš **hi** lipa šušit!
 you.NOM then DEM.DIST-ACC.SG.F water-ACC.SG.F must-PRS-2SG pour.away.PFV-INF
 and **they-ACC** must-PRS-2SG **they.ACC** nicely dry-INF
 'You then have to throw that water away, and you have to dry **them** nicely!'

(Breu 2017: 76)

Therefore, investigating clitic doubling in Molise Slavic holds relevance for the study of both definiteness and information structure, due to the correlation of clitic doubling with definite NPs in comparable languages. Nonetheless, Breu emphasises that the doubling is optional (see Breu 2017: 81), because it is easy to find contrasting pairs in which no doubling occurs. The same was argued for long time in the context of Bulgarian clitic doubling (see Tomić 2008: 457). It was mentioned as an optional construction, in contrast to Macedonian, where it is obligatory with every definite NP. Further research on Bulgarian made clear that CD in Bulgarian is by no means random. It rather follows clear conditions and has a pragmatic function that triggers its occurrence. Given that language is a dynamic process, it might not always be possible to explain every single occurrence of a construction, although it should be possible to identify a core meaning for a construction, especially when it occurs that frequently in a language. In case of Bulgarian, it is sometimes emphasised that specificity seems to be the crucial prerequisite for clitic doubling:

Identification of the item's uniqueness by the speaker is central, since non-specific DPs may not be duplicated. (Tomić 2008: 473).

Although we expect this given that we consider pronouns to be inherently definite, we must refine this in the case of clitic doubling. Example 2.2-7 is interesting, because the water to which the speaker is referring is not really specified because he generally explains the process of producing a dish with peppers. Nonetheless, within the context the water might be considered definite at an abstract level because it is the water that was just mentioned. In this case, definiteness refers to an entity in the discourse world, rather than making reference to something in the actual outer world.

When the doubling is made to a full form pronoun, we are dealing with definiteness, in deictic use as well as anaphorical use:

2.2-11 Andz, ču **ta** či bijat pur **teba** prije!

even will.PRS-1SG **you.ACC.SG** make.INF send.PFV-INF also **you.SG-ACC** before

‘I’ll even let **you** go first!’

(see *The hare and the crayfish* in the appendix)

In examples 2.2-11, the deictic reference to the hearer enforces definiteness. Therefore, this precondition is probably necessary in every language in which the clitics derive from the pronoun (as long as they can still be used in this function). Nonetheless, the precondition does not explain the motivation that a speaker has to use the doubling.

I mentioned before that in the South European context two motivations are typical and therefore need to be checked. The obligatory use with every definite NP (as in Macedonian) can be excluded. The other option is to assume that clitic doubling fulfils some type of topicalization. Examples 2.2-12 and 2.2.13 indicate this interpretation. In both examples, a referent is mentioned right before and then reference is taken up again in the second clause in which the doubling occurs.

2.2-12 [...] Karluč je osta nondeka, nisu **ga** vazal **njega**.

Karluč.NOM.SG.M be.PRS.3SG remain.PVF-PTCP.SG.M there NEG-be.PRS.3PL **he.ACC** take.PFV-PTCP.SG.M **he.ACC**

Karluč remained there, they did not take **him** with them.

(Breu 2011: [1.2.2]-5)

2.2-13 [...] one kisa što s’ **mu** čil **njemu**!

DEM.DIST-NOM.PL ADV what with **he.DAT** make.PFV-PTCP.PL **he.DAT**

‘Who knows, what they did with **him**!’

(Breu 2017: [1.2.2]-392)

The same happens in examples 2.2-14, although here it is interesting to note that the subject changes, but topichood remains on the boy. Nonetheless, reference to this topic is made by doubling pronominal reference:

2.2-14 [...] na dičalj tijaša nu divojku a ova divojka pur **ga** tijaša **njega**.

ART.INDF-NOM.SG.M boy.NOM.SG.M want.IPFV-IPRF.3SG ART.INDF-ACC.SG.F girl-ACC.SG.F and DEM.PROX-NOM.SG.F girl-NOM.SG.F also **he.ACC** want.IPFV-IPRF.3SG **he.ACC**

‘A boy liked a girl, and this girl also liked him.’

(Breu 2017: [2.4.1]-50)

It seems that clitic doubling is related to topicality. This needs further investigation with experiments focussing on this issue.

2.3. Possessives

2.3.1. Genitive case

In his seminal work, Lyons mentions possessive determiners (*my, their*) and genitive forms (possessive forms with 's in English; *Fred's*) under the heading of *possessives* (cf. Lyons 1999: 22), taken together as one category based on syntactical deliberations (cf. Lyons 1999: 124). In my account I do not want to discuss the second case in length. Without further ado, I understand proper nouns as being necessarily definite by their lexical meaning. The NP *Džuva'* in example (2.3-1) is definite, as can be seen by the first NP referring to the same referent. Examples like (2.3-2) raise the question of whether the respective NP should be interpreted as being specific only rather than being definite, given that the discourse entity is only known to the speaker but cannot necessarily be identified by the interlocutor. For the purpose of communication, it might be sufficient for the interlocutor to understand that there is an entity called Giovanni even acknowledging that there are several other Giovanni in the world.

2.3-1 [...] **Džuva'**, je reka familju ka ovi je bi umbra [...]!

Giovanni, be.PRS.3SG say.PFV-PTCP.SG.M family-ACC.SG.F REL DEM.PROX-NOM.SG.M
be.PRS.3SG PART die.PFV-PTCP.SG.M

'**Giovanni**, he told the family that this one had died!'

(Breu 2011: [1.2.2]-244)

2.3-2 Je reka drugi čeljada, sa zovaša **Džuvan**.

be.PRS.2SG say.PFV-PTCP.SG.M another-NOM.SG.M person.NOM.SG.M REFL call.IPFV-IPRF.3SG **Giovanni**

'He told another person, that he was called **Giovanni**.'

(Breu 2011 : [3.1.3]-7)

This problem is also described in Lyons (1999:21):

But a common view is that we use proper nouns as if they were absolutely unique, corresponding more closely to inherently unique definites (like the sun) than to possibly contextually unique definites (like the man). [...] It is clear that the uniqueness of reference of proper nouns is what aligns them with definites, though it may be added that this very uniqueness will generally ensure the identifiability of their referent. (Lyons 1999: 21).

One could challenge this view in questioning how this identifiability should be understood, since it seems that very often a minimal identifiability – namely that the hearer just accepts that there is a person called Giovanni about whom the speaker is talking – is sufficient. Since this aspect will occur several times in this study I will call it a *requirement of minimal knowledge* about the referent. For the sake of brevity, I cannot proceed deeper into this discussion. For our purposes, it is sufficient to take proper nouns as being inherently definite in the semanto-pragmatic sense (for alternative views and explanations, see Lyons 1999: 21-22). In Italian, the definite article is sometimes added to proper nouns in certain contexts. Due to the absence of an grammaticised overt definite article in Molise Slavic, this plays no role here. In addition to the points just mentioned, genitive forms of proper nouns are also definite. In example 2.3-3, the Little Prince must be considered a proper noun and therefore the possessor is definite due to this proper noun status. In example 2.3-4, both the possessor and the possessed are indefinite:

2.3-3 [...] ovi fat zakrati do života **maloga kraljiča**.

DEM.PROX-NOM.SG.M story-NOM.SG.M secret-NOM.SG.M of life-GEN.SG.M **little-GEN.SG.M prince-GEN.SG.M**

‘[...] this secret story of the **Little Prince’s** life.’

(Saint-Exupéry 2009: 27)

2.3-4 Ovi štarič je bi poša furt utra portun, **na veliki portun do jene hiže boghate**.

DEM.PROX-NOM.SG.M basket.NOM.SG.M be.PRS.3SG PART go.PFV-PTCP.SG.M end.PFV-NOM.SG.M in door-ACC.SG.M **ART.INDF-NOM.SG.M big-NOM.SG.M door-NOM.SG.M of ART.INDF-GEN.SG.F house-GEN.SG.F rich-GEN.SG.F**

‘This basket finally came into the door, **a large door of a rich house**.’

(Breu 2011: [3.6.7]-10)

The definiteness of *život* ‘life’ in 2.3-3 is not necessarily explicitly marked. This referent is definite by world knowledge, since it is clear that one person (treating the uminous Little Prince as [+human]), only has one life; therefore, fulfilling the uniqueness condition. In contrast to pronouns or demonstratives, “possessives [...] are not inherently definite or semantically incompatible with indefiniteness in the noun phrase they modify” (Lyons 1999: 130). It depends on the exact interplay of the possessor and the possessed as well as world or situational knowledge if a possessor or a possessed should be considered as being specified as definite.

2.3.2. *do*-construction

The special contact situation with Italian seems to trigger some unusual patterns in the expression of possessives. With full-noun possessors, we find a specific pattern that reflects some of the genetic inheritance from Slavic origin as well as some influence of the donor varieties leading to a somewhat special system. Molise Slavic has two strategies for the expression of possession without possessive pronouns: besides the sole use of genitive case marking on the possessor, the preposition *do* ‘of/by’ (analogous to Italian *di*) can be used. It would appear that in contrast to Italian (and other varieties of BCS), Molise Slavic has three ways for expressing possession, which differs from the surrounding language because:

Italian has only pronoun-derived possessive determiners (*mio* ‘my’ etc.); possession with full-noun phrases is expressed prepositionally (*il libro di Carlo* (the book of Carlo) ‘Carlo’s book’) – a structure which does not seem to impose definiteness in any language.” (Lyons 1999: 24).

Here, it is important to note Lyons’ claim on the definiteness of prepositionally expressed possession. As we saw in example (2.3-4), the use of the preposition itself has no effect on the definiteness of the NP within its prepositional phrase. In (2.3-4), the indefinite article is used to overtly mark the indefinite reading of the NP *hiža* ‘house’. In (2.3-5), no indefinite marking is used:

2.3-5 Ova žena, **spodarica do hiže**, je rekla: [...].

DEM.PROX-NOM.SG.F woman-NOM.SG.F **landlady-NOM.Sg.F of house-GEN.SG.F**
be.PRS.3SG

‘This lady, **the lady of the house**, said [...].

(Breu 2011: [3.6.7]-35)

In contrast to possession expressed with a pronoun (see below), in the *do*-construction possession is separate from definiteness. Examples 2.3-4 and 2.3-5 show instances of possession with the *do*-construction with indefinite articles and without.

In possessive constructions with bare genitive, only examples were found with definite possessors. Possessors marked overtly with the ART.INDF.GEN.SG or PL (*jenoga/jenga*) were not found; rather, all instances of these forms occurred either as an accusative form of an animate referent or in combination with a preposition requiring the genitive case:

- 2.3-6 Ona, žena **Karbunirina**, Džuva', ona je bi sa udala drita [...].
 DEM.DIST-NOM.SG.F woman-NOM.SG.F **policeman-GEN.SG.M** Giovanni she-NOM
 be.PRS.SG PART REFL marry.PFV-PTCP.SG.F straightforward
 'This one, the wife **of the policeman**, she got married for good.'

(Breu 2011: [1.2.2]-310)

- 2.3-7 [...] Biša tata **pokonja Paskvala Dzakarina**.
 Be.IPRF.3SG father.NOM.SG.M **deceased-GEN.SG.M Pasquale-GEN.SG.M Dzakarini-**
GEN.SG.M
 'He was the father **of the deceased Pasquale Dzakarini**'.

(Breu 2011: [1.2.4]-15)

- 2.3-8 Ja sa skanja solda **moga tata**! [...].
 I.NOM be.PRS.1SG change.PFV-PTCP.SG.M money.ACC.SG.M **my-GEN.SG.M**
father.GEN.SG.M
 'I changed the money **of my father**'.

(Breu 2011: [3.4.2]-30)

The question remains in which situations a pure genitive is sufficient for the expression of possession and when the preposition *do* is (or must be) used. One could hypothesise that definiteness plays a role here given that we could not find any instance of indefinite possessors as bare noun genitives. However, without further elicitations this question cannot be solved. A possible alternative explanation for having three types of expressing possession is perhaps a partitive meaning of possession expressed with the *do*-construction. At a broader level, it might relate to a differentiation into alienable and inalienable possession (see Lyons 1999: 128). For instance, in example 2.3-4 parts of house the suggest an intrinsically held possession while the sold of the father in example 2.3-8 is only related to him, not an intrinsic part.

2.3.3. Possessive pronouns

In the case of possessive pronouns, the definiteness of the possessor derives from the pronominal being which again depends on contextual factors, as we have seen in the previous chapter. For full-noun possessors, it rather depends on context, although the exact relation is language-specific: "[...] that in some languages a possessive induces definiteness in the matrix noun phrase while in other languages it does not" (Lyons 1999: 24). In example 2.3-9, the uniqueness of the (abstractly) possessed entities (*sfekar*

‘father-in-law’, *žena* ‘woman/wife’) derives from the knowledge that within the context of European culture we only expect one father-in-law as well as one wife.

2.3-9 E **moj sfekar, moja žena** aš pokonja Pipina sijahu žita [...].

And **my.NOM.SG.M father-in-law.NOM.SG.M my-NOM.SG.F woman-NOM.SG.F** and deceased-NOM.SG.M Pipina sow.PFV-IPRF.3PL wheat-ACC.SG.M

‘And **my father-in-law, my wife** and the deceased Pipina sowed the wheat [...].’

(Breu 2011: [1.2.2]-279)

2.3-10a Je sa sprobudija pur **moj dita**.

be.PRS.3SG REFL get-up.PFV-PTCP.SG.M also **my-NOM child.NOM.SG.M**

‘**My child** woke up, too.’

(Breu 2011: [3.2.7]-13)

2.3-10b Eine meiner Töchter...

2.3-10c Meine eine Tochter...

By contrast, in example 2.3-10a it is not necessarily clear whether the speaker has only one child or several. Therefore, reference is ambiguous (even if we assume Breu’s claim about a definite article to be correct). One could imagine that constructions as in German (2.3-10b and c) – which help to narrow down the proposition to one interpretation – are possible, but were not found in the corpus. Thus far, we can posit that possessives may trigger a definite interpretation of one or the other constituent of the possession, but in the case of possessive pronouns we must assume definiteness at least on side of the possessor. Possessive pronouns are similar to genitive possessors regarding their function (and often position) in the NP, because “[t]hese possessive forms of noun phrases occur as modifying expressions within other noun phrases” (Lyons 1999: 23). The system of possessive pronouns is comparable to the respective forms in nowadays BCS standard languages³, although there is some reduction in the system, probably due to the influence of Italian (cf. Breu 2017: 43): In general, Molise Slavic also exhibits the two-fold system of expressing person and number of the possessor in the word stem and the possessed with the case marking (*moj grad* ‘my village’, *njegov grad* ‘his village’, *njegova čita* ‘his city’). In contrast to standard BCS (*njegov* ‘his’ vs. *njezin* ‘her’), *njegov* in Molise Slavic is used for both male and female possessors. A distinct plural form (*njihov* ‘their’) yet exists. For the sake of

³ The role of tone is not discussed here, but see footnote 39 in Breu 2017: 42.

completeness, the full paradigm is listed in table 3 in the appendix. In the accusative of masculine singular possessed, the animateness category is mainly preserved (see Breu 2017: 42 and example 2.3-11), but we can find counter-examples in which the difference between animate and non-animate nouns seems to be confused (2.3-12):

2.3-11 Su pol nondeka, su pital **moga cila**: [...].

be.PRS.3PL go.PFV-PTCP.PL there be.PRS.3PL ask.PFV-PTCP.SG.M **my-ACC.SG.M uncle-ACC.SG.M**

‘The went there and asked **my uncle** [...].’

(Breu 2011 : [3.1.5]-8)

2.3-12 E za po di biškup, je ga ponija **njev brat**, s mulam.

And for go.PFV-INF to forest-ACC.SG.M be.PRS.3SG carry.PFV-PTCP.SG.M **his.ACC.SG.M brother.ACC.SG.M** with mule-INS.SG.F

‘And for going to the forest, he carried **his brother** on the mule.’

(Breu 2011: [3.1.6]-3)

The difference between the three variants of Molise Slavic seems to be minor. There is some variation in the system regarding the stems: reduced forms exist for most forms and it seems that they are used more frequently (cf. Breu 2017: 43). More detailed research should be conducted on the question whether this variation exists within the speaker or if one or the other form may be idiosyncratic to individual speakers. No reflexive possessives of the type *sovoj knjiga* ‘my/your/his/her own book’ exist, which is untypical from a Slavic perspective (see Breu 2017: 42). Instead, like German, English, Italian and other languages reference is always made at the level of the person. In some languages of the world, distinct forms of possessive pronouns exist for attributive (*my, their*) and pronominal/predicative use (*mine, theirs*). In Molise Slavic, no such distinction can be found:

2.3-13 Je reka ka ono neče, moj sin. Si maša bi **njevog**, biša **njevog** [...].

be.PRS.3SG say.PFV-PTCP.SG.M REL DEM.DIST-NOM.SG.N NEG-want.IPRV.3SG my.NOM.SG.M son.NOM.SG.M REFL must.IPFV-IPRF.3SG PART **his.NOM.SG.M** be.IPRF.3SG **his.NOM.SG.M**

‘My son said that he does not want this. If it were **his**, it would be **his** [...].’

(Breu 2011: [2.1.1]-141)

The position of possessives is interesting. In contrast to demonstratives, the possessive pronouns can occur pre-nominally as well as post-nominally. All examples thus far were instances of possessives pronouns in pre-verbal position that were parallel

structures as possessive NPs with full-noun possessors. The next two examples illustrate the other possibility:

2.3-14 Biša **drugi kumbanj naš** [...].

be.IPRF.3SG **another-NOM.SG.M fellow.NOM.SG.M our.NOM.SG M**

'There was **another fellow of us** [...].'

(Breu 2011: [1.2.2]-161)

2.3-15 Biša **na kumbanj naš** [...].

be.IPRF.3SG **ART.INDF-NOM.SG.M fellow.NOM.SG.M our.NOM.SG M**

'It was **a fellow of us** [...].'

(Breu 2011 : [1.2.2]-236)

In this regard, the possessive pronouns behave like adjectives in Italian and Molise Slavic. It was noted long before that in some languages possessive pronouns can only behave like genitive-based possessors and occupy only the same position, while in other languages they rather behave syntactically (and in agreement) like adjectives. Lyons introduced the (controversial) classification of *AG languages* and *DG languages*. In *DG (determiner-genitive)* languages like English and Irish, the possessor (pronoun, proper noun or full NP) resides in the position of determiners, hence blocking this position for a co-occurring determiner/modifier. On the other hand, in *AG (adjectival-genitive)* languages like Italian and Greek, no such restriction exists because possessors occupy the same position as adjectives (cf. Lyons 1999: 24-25). Adding a determiner to the possessor is possible (*il mio libro, il libro di Vittoria*). In other words:

in the first type a possessive has the effect of inducing a definite interpretation in the noun phrase it modifies, and a definite article cannot also appear; in the second type a possessive does not have this effect, and the article must co-occur with it to get a definite interpretation (in languages that have an article). (Lyons 1999: 130)

The type to which a language belongs can be tested by finding examples in which demonstratives and adjectives co-occur. With several queries in the corpus (combinations of possessives and demonstratives in Molise Slavic and Italian), only few clear examples emerged:

2.3-16 [...] **kaka je ovi moj**, ma bit pur **ta tvoj**!

how be.PRS.3SG **DEM.PROX-NOM.SG.M my-NOM.SG.M** but be-INF also **DEM.MED-NOM.SG.M your-NOM.SG.M**

'How **mine** is, should be also **yours**!'

(Breu 2011 : [1.2.2]-381)

2.3-17 Ti čaš razumit ka **ona tvoja** je unika na sfit za teba.

you.NOM want-PRS.2SG understand.IPFV-INF REL **DEM.DIST-NOM.SG.F your-NOM.SG.F** be.PRS.3SG unique-NOM.SG.F in world-ACC.SG.M for you.ACC.SG.M

‘You will understand that **this one of yours** is unique to you in the world.’

(Saint-Exupéry 2009 : 70).

2.3-18 Ti znaš, znaš **oni puč naš** nonda?

you.NOM know.IPFV-PRS.2SG know.IPFV-PRS.2SG **DEM.DIST-ACC.SG.M fountain.ACC.SG.M our.ACC.SG.M** there

‘Do you know **that fountain of us** over there?’

(Breu 2011: [2.5.2]-46)

Based on these two examples, it is impossible to decide with absolute certainty whether Molise Slavic qualifies as a AG or DG language. This needs to be tested further with grammaticality judgement tests. On syntactic grounds, one could argue that the possessive takes the position of the noun in the NP here, since no full noun is present. Due to this movement, the demonstrative position is free and can be taken by the demonstrative itself. Example 2.3-18 delivers additional support that the possessive usually occurs in the demonstrative slot in the NP because this example indicates that both cannot occur sequently. The possessive also moves to the post-nominal position with other modifiers (example 2.3-14) and the indefinite article (example 2.3-15), while the demonstrative and *drugi* can occur together pre-nominally (example 2.3-20)

2.3-20 [...] **oni drugi vina** [...].

DEM.DIST-NOM.SG.M another-NOM.SG.M wine-NOM.SG.M

‘[...] **this other wine** [...].’

(Breu 2011: [2.3.1]-52)

Based on this, Molise Slavic qualifies as a likely candidate for being a determiner-genitive language and therefore does not fully model the Italian example in this grammatical category. In theory, whether the respective matrix NP is considered definite or not depends on the additional use of a determiner in AG language. In Italian, this is also confirmed by the fact that Italian allows for omission of the definite article when possessives are used on certain kinship terms (see Lyons 1999: 128-129). We assume that in context reference to some family members is unique and therefore definite. In all other cases, Italian speakers must overtly mark the definiteness on the NP in addition to the possessor, as in *il mio libro* (ART.DEF.SG.M my-SG.M book-SG.M)

‘my book’. Both categories are separate. In Molise Slavic, we assume that possessive pronouns and possessives expressed bare nouns carry definiteness as a meaning. In case of possession expressed with either the do-construction this depends on contextual factors or explicit marking.

2.4. Numerals and quantifiers

Numerals and quantifiers are sometimes related to definiteness, either due to their semantics or regarding their syntactic behaviour. Numerals are “linguistic sign[s] used for conceptualising and expressing a set of discrete entities or the relative value of an element of such a set” (Gvozdanovic 2015). The cardinal forms of numerals (*two*, *ten*) carry no specification regarding definiteness. In an English NP including a numeral, the definiteness of the phrase is achieved by other features in the phrase; for example, the definite article in the example (2.4-1a). Numerals without articles are most probably not definite, but rather specific or generic (2.4-1b):

2.4-1a He bought **the two famous books** by this author. [+spec] [+def]

2.4-1b **Two pairs of socks** are more practical than one. [-def] [-spec] [+generic]

2.4-1c Biša jena... jena boghate e jimaša **dva dica** [...].

be.IPRF-3SG ART.INDF-NOM.SG.M ART.INDF-NOM.SG.M rich-NOM.SG.M and have-IPRF.3SG **two.F child.ACC.PL**

‘There was a... a rich person and he had **two girls** [...].’

(Breu 2011: [3.6.13]-1)

In the Molise Slavic example 2.4-1c, the referents are specific by context rather than by semantic meaning of the numeral *dva*. Due to its numerical origin, the indefinite article (*jena*) cannot co-occur with the cardinal numbers. This does not necessarily mean that non-specific readings are prohibited with numerals. In the following example, it does not appear necessary that we must assume a specific interpretation of the referent. Which exact two kilometres are meant is simply not relevant.

2.4-2 Je nami reka ka hočaša dijač mila lir: **dva kilometra puta**, hočaša dijač mila lir.

be.PRS.3SG us.DAT say.PFV-PTCP.SG.M REL want.IPFV-IPRF.3SG ten thousand.F lir.GEN.PL **two.M kilometre-GEN.SG.M way-GEN.SG.M** want.IPFV-IPRF.3SG ten thousand.F lir.GEN.PL

‘He said to us that he wanted ten thousand lira: **for two kilometres street**, he wanted ten thousand lira.’

(Breu 2011: [3.5.2]-95)

In the case of ordinal numbers, one could expect that there is a tendency for the respective referent modified by the ordinal to be at least specific. However, at least in English and German there seems to be a strong tendency to exhibit a definite article (2.4-3a) while prohibiting the sole use of the ordinal numeral or restricting it to a few contexts (2.4-3b). Ordinals may occur with indefinite articles as well (2.4-3c), although I assume that there are some restrictions at work which I cannot investigate further. Ordinals are always referring to some type of ranked entities and single instances of the set could be interpreted as being specific due to its membership in the rank.

2.4-3a **The 100th costumer** received a price.

2.4-3b **Zweite Wahl** ist nicht immer schlecht.

2.4-3c **Ein erster** Tag/ **dieser erste** Tag/ **der erste** Tag

In the corpus, ordinal numbers are not sufficiently frequent to explore them in depth in this study. Molise Slavic speakers borrowed the Italian ordinal system to a large extent. In the first form, there is huge variety in the use. Both, different variants of the Slavic ordinal *prvi* are used along with the Italian ordinal number *primo*, although the latter is probably a result of code-switching. The ordinals can occur attributively (2.4-4, 2.4-6, 2.4-7, 2.4-8) as well as predicatively (2.4-5).

2.4-4 Moj muž, kaka biša orfan, vamaša sfe **prvi post**

my.NOM.SG.F man.NOM.SG.M how be.IPRF.3SG orphan.NOM.SG.M take.PFV-IPRF.3SG
always **first-ACC.SG.M seat.ACC.SG.M**

‘My husband, since he was an orphan, he always took **the first seat**.’

(Breu 2011: [1.1]-152)

2.4-5 Ma oni dita biša **pervi** [...]

but DEM.PROX-NOM.SG.M child-NOM.SG.M be.IPRF.3SG **first-NOM.SG.M**

‘But that child was **the first** [...].’

(Breu 2011: [1.1]-304)

2.4-6 **Priji čeljada** ka je ubija vuca!

first.NOM.SG.M person.NOM.SG.M REL be.PRS.3SG kill.PFV-PTCP.SG.M wolf.ACC.SG.M

‘**The first person** who killed a wolf.’

(Breu 2011: [3.2.2]-10)

2.4-7 [...] **primo giugno** [...].

first-NOM.SG.M june.NOM.SG.M

‘[...] **the first of june** [...].’

(Breu 2011: [2.1.4]-39)

- 2.4-8 Maša uhitit dol, **primo pian, terzo pian** [...].
 must-IPRF.2SG begin.PFV-INF below **first-NOM.SG.M floor-NOM.SG.M third-NOM.SG.M floor-NOM.SG.M**
 ‘You must start down below, **the first floor, the third floor** [...].’

(Breu 2011: [1.2.1]-274)

For ordinals higher than ‘one’, only instances of ‘second’ and ‘third’ occurred in the corpus. Those forms are all borrowed from Italian, but adjusted partly to the Molise Slavic phonology. The borrowing led to a reduction in meaning in the form *drugi*, which in contrast to Croatian only expresses the meaning ‘(an)other’ while the ordinal meaning is expressed by the Italian form. Interestingly, in some cases the Italian agreement of the form is overtaken when it co-occurs with a borrowed noun (2.4-8).

- 2.4-9 Je riva **sekond taksišt** [...].
 be.PRS.3SG arrive.PFV-PTCP.SG.M **second.NOM.SG.M cabdriver.NOM.SG.M**
 ‘A **second cabdriver** arrived [...].’

(Breu 2011: [3.5.2]-106)

- 2.4-10 **Tèrc dan** sma hitil vršaj.
third.ACC.SG.M day.ACC.SG.M be.PRS.1PL throw-out.IPFV-INF grain.ACC.SG.M
 ‘On **the third day**, we throw away the grain.’

(Breu 2011: [2.1.3]-16)

In BCS, the numerals behave mainly like adjectives and can be considered as belonging to this category. Since the ordinal forms are derived from ordinal numerals by adding the ending *-i* (plus some additional material), in theory they are long (“definite”) forms of adjectives (see next chapter). While this might be true regarding the historical development of the ordinal forms, this does not mean that ordinal forms in BCS are automatically marked for definiteness. They can occur together with the numeral ‘one’ (2.4-11), which has the effect of excluding definiteness (while specificity is preserved).

- 2.4-11 Među prvim damama i **jedan prvi gospodin** (Croatian)
 among first-INS.PL women-INS.PL.F and **one.NOM.SG.M first-NOM.SG.M mister-NOM.SG.M**
 ‘Among first women and **one first man**.’

(<http://net.hr/danas/svijet/medu-prvim-damama-i-jedan-prvi-gospodin/#1>)

- 2.4-12 Kako je jedan rob postao **prvi čovjek** koji je oplovio svijet (Croatian)
 how be.PRS.3SG one.NOM.SG.M slave.NOM.SG.M become.PFV-PTCP.SG.M **first-NOM.SG.M person-NOM.SG.M** who sail.PFV-PTCP.SG.M world.ACC.SG.M
 ‘How a slave became **the first person** sailed around the world.’

(<<https://www.cdm.me/zabava/zanimljivosti/kako-je-jedan-rob-postao-prvi-covjek-koji-je-oplovio-svijet/>>)

2.4-13 **Ovaj prvi policajac** mi je rođak (Croatian)

DEM.PROX-NOM.SG.M **first**-NOM.SG.M **policeman**.NOM.SG.M I.DAT be.PRS.3SG
cousin.NOM.SG.M

‘**This first policeman** is a cousin of mine.’

(<<https://www.vecernji.hr/vijesti/ovaj-prvi-policajac-mi-je-rodak-kojem-je-pola-familije-bilo-u-ratu-1007980/komentari>>)

In the Molise Slavic and Croatian examples, it is rather plausible to assume specificity to accompany the occurrence of ordinal forms, but not necessarily definiteness. If it carried definiteness inherently, we would expect that demonstratives and ordinals exclude each other (as in the Croatian example 2.4-13). Due to the small sample, this cannot be tested in the case of Molise Slavic. Despite sharing some features of adjectives, ordinals apparently only occur in front of the noun in attributive position.

Another category related to both numerals and definites is word forms subsumed under the heading of quantificational expressions or quantors (*all, most, many*). In general, they “share some of the behaviour and distribution of definites” (Lyons 1999: 148). According to Lyons, they often correspond to demonstratives regarding their distribution:

They may occupy Det[erminer] position as in English, in accordance with their apparent definite determiner status; or they may be in some more internal position, with something like adjectival status, but then with the definite article present (in languages that have one) to ensure the definiteness with which the quantifier must be associated. (Lyons 1999: 148).

Determiners and demonstratives exclude each other from co-occurrence due to their semantics, while they can both accompany quantifiers. Note that in all examples from English the quantifier takes the first position, while other positions are ungrammatical (all behaves differently from most because they are not belonging to the same grammatical category on syntactical grounds):

2.4-14a All those people I met in the last couple of months.

2.4-14b All the people...

2.4-14c All people...

2.4-14d The most people...

In this regard, Molise Slavic behaves just like English and Italian in allowing co-occurrence of the quantifier *tuna* ‘all’ in the demonstrative position and demonstratives (2.4-15). In the predicative position, the demonstrative is in the head

position of the NP (2.4-16, first instance), while *tuna* is in the first position of the phrase. This structure is also preserved when both lexemes occur with a full noun (2.4.16, second instance). Together with possessives, *tuna* takes the first position, too (2.4-17).

- 2.4-15 Je kazala **tuna ono** ka nosaša.
 be.PRS.3SG show.PFV-PTCP.SG.F **all DEM.DIST-ACC.SG.N** REL carry-PFV-IPRF.3SG
 'He showed **everything** what he brought.'
 (Breu 2011 : [3.6.11]-8)
- 2.4-16 Sa, biša po dubrave na kačatur ka bi čuja **tuna ovo, tuna ova kumèdja**.
 now be-IPRF-3SG around forest-GEN.SG.F Art.INDF-NOM.SG.M hunter.NOM.SG.MREL PTL
 hear-PTCP.SG.M **all DEM.PROX-ACC.SG.N all DEM.PROX-NOM.SG.F fuss-NOM.SG.F**
 'Now there was a hunter about in the woods, who had heard **everything, all that fuss**.'
 (Breu 2017: 219)
- 2.4-17 [...] je razdilija **tuna njevu parcijunu** [...].
 be.PRS.3SG divide.PFV-PTCP.SG.M **all his-ACC.SG.F part-ACC.SG.F**
 'He divided **all of his part**.'
 (Breu 2011: [3.6.13]-5)

For testing other quantifiers, the data sample is too small. Instances of (most) can be found but are too rare to make any claims about its use and position in the NP:

- 2.4-18 **Veća voti** biša na žena ka činaša za mašatura.
most time-ACC.PL be.IPRF.3SG ART.INDF-NOM.SG.F woman-NOM.SG.F REL make.IPFv-
 IPRF.3SG to marriage broker-NOM.SG.F
 '**Most times**, there was a woman who made the matchmaking.'
 (Breu 2017: [2.4.1]-3)

I mentioned similarities between quantifiers and definites on the one hand, but also between quantifiers and adjectives on the other. Usually adjectives are not discussed in the context of definiteness, because they constitute an independent grammatical category. As we have seen before, in the nominal phrase there is typically a slot for demonstratives as well as a separate slot for adjectives, both normally being independent. To sum up what has been said before, we see that adjectives do not interfere much with other modifiers licensing a definite interpretation of the nominal phrase. Regardless, BCS is often quoted for having two paradigms of adjectives comprising long and short forms. The short forms are claimed to express a definite reading. This claim is discussed in the next chapter and the nature of this distinction in Molise Slavic is tested.

2.5. Adjectives

Molise Slavic still makes use of a distinction between long and short forms of adjectives. This system is well known from Old Church Slavonic and is also preserved in BCS and Slovenian, although there is a tendency to lose the distinction in most of the modern standards, “the ‘definite’ form being now used far more than the ‘indefinite’, and no longer conveying definiteness [...]” (Lyons 1999: 82, see also Trenkic 2004: 1405). In general, the distinction can only be marked morphologically in the singular masculine, although it is sometimes claimed that this function can be expressed with tone in the other forms (see Trenkic 2004: 1405, see also Lyons 1999: 82). Considering the critique by Trenkic, we can also reanalyse the Molise Slavic system based on his analysis for the standard BCS languages. Trenkic quotes some other grammars in which the distinction is still described with the definite – indefinite dichotomy (cf. Trenkic 2003: 1403). In his analysis, he reaches the conclusion that for adjectives that still preserve the distinction it is rather specificity that is expressed by the long forms of adjectives in BCS) see Trenkic 2003: 1404). For practical reasons, the focus is on NOM.SG.M forms, since the difference is only really visible here. A tonal analysis of adjectives could not be conducted within the realm of this study. In the classical sense, the phrase in the Croatian example 2.5-1a is considered as being marked for definiteness, while the phrase in 2.5-1b is understood as indefinite in contrast. In Molise Slavic, instances of this dichotomy can be found. In 2.5-2a, reference is made to an entity not necessarily being definite (but specific), while in 2.5-2b the (nominalised) adjective is specified for definiteness.

- 2.5-1a Mladi student (Croatian)
 young.NOM.SG.M student.NOM.SG.M
 ‘The young student.’
- 2.5-1b Mlad student (Croatian)
 young.NOM.SG.M student.NOM.SG.M
 ‘A young student.’
- 2.5-2a Nije dobar žita [...]. (Molise Slavic)
 NEG-be.PRS.3SG good.NOM.SG.M wheat.NOM.SG.M
 ‘It was not good wheat.’

(Breu 2011 : [2.1.4]-51)

2.5-2b [...] je rispuniija presenduzi. (Molise Slavic)
 be.PRS.3SG respond.PFV-PTCP.3SG idle-NOM.SG.M
 'The idle answered: [...]'.
 (Saint-Exupéry 2009: 42)

However, if we take a closer look, it becomes clear that the adjective system is undergoing a process of restructuring. Nonetheless, in all examples something else can be noted. All instances are [+spec]. Additional support for this analysis could stem from the fact that long forms of adjectives may co-occur together with demonstratives. This means that definiteness must be additionally expressed when it needs to be expressed overtly. In the corpus, only one such example was found (2.5-3), unfortunately for the analysis not with a full noun. Since this phrase could be reanalysed as comprising a nominalised adjective, one could assume that demonstratives play a role in this deadjectivising process. Example 2.5-4 is another instance of this process.

2.5-3 Prije maš pit **oni stari** o maš pit **oni bili** [...] .
 first must.PRS.2SG drink.IPFV-INF DEM.PROX-NOM.SG.M old-NOM.SG.M or must.PRS.2SG
 drink.IPFV-INF **DEM.PROX-NOM.SG.M white-NOM.SG.M**
 'First, you must drink **that old one** or you must drink **that white one**.'
 (Breu 2011: [2.3.2]-74)

2.5-4 One bihu **one crne**, se nenadam do je race bihu! Bihu crne.
 DEM.DIST-NOM.PL be-IPRF-3PL **DEM.DIST-NOM.PL black-NOM.PL** now NEG.know-
 PRS-1SG of which race-GEN.SG.F be-IPRF-3PL be-IPRF-3PL black-NOM.PL
 'They were **these blacks**, I don't know now, which race they were! They were black.'
 (Breu 2017: 126)

Breu notices that there is tendency in Molise Slavic that long form adjectives are used more often in the attributive position, while the short form can be found more often in the predicative position of a clause (cf. Breu 2017: 34), although it seems that both forms are used in the predicative position (see Breu, Piccoli 2000). This is in line with the observation of Lyons (based on Javarek, Sudjić 1963), who notices the same for BCS:

For adjectives that do have the two forms, the distinction has partly grammaticalized, the definite[long] form being obligatorily after demonstratives and possessives and with vocatives, and the indefinite form in predicative use (Lyons 1999: 82).

In 2.5-5, the short form is used in the predicative position. However, the same form *dobar* was used in the previous example 2.5-2a in the attributive position. This is not

in line with Breu's analysis, since short forms are rather expected to occur in predicative position only, while the long form is more flexible. We find examples such as 2.5-6 in which the long form is used in the predicative position, but also long forms in the expected attributive position. In our corpus, we find more examples of the short form *dobar* (20) than for the long form *dobri* (5), but the scale is too low to make predictions based on these numbers.

2.5-5 Vagošta [...] žita nije **dobar**.

this year wheat.NOM.SG.M NEG-be.PRS.3SG **good-NOM.SG.M**

'This year [...], the wheat was not **good**.'

(Breu 2011: [2.1.4]-50)

2.5-6 Moj sin biša **mali** [...].

my.NOM.SG.M son.NOM.SG.M be.IPRF.3SG **small-NOM.SG.M**

'My sun was **small** [...].'

(Breu 2011: [3.2.6]-19)

2.5-7 [...] Alesandr [...] biša **na dobri** čeljada.

Alexander be.IPRF.3SG ART.INDF-NOM.SG.M **good-NOM.SG.M person.NOM.SG.M**

'Alexander [...] was **a good person**.'

(Breu 2011: [1.2.2]-150)

The last example 2.5-7 is interesting from another perspective. Definite forms may occur together with the indefinite article. We find both instances of indefinite articles with long as well as short forms quite frequently:

2.5-8 Na dan sa vidija utra na štip **nu lipu butilju zelenu** [...].

ART.INDF-ACC.SG.M day.ACC.SG.M be.PRS-1SG see-PTCP.SG.M in ART.INDF-ACC.SG.M
sideboard.ACC.SG.M ART.INDF-ACC.SG.F **beautiful--ACC.SG.F bottle--ACC.SG.F green-ACC.SG.F**

'One day, I saw in a sideboard **a beautiful green bottle** [...].'

(Breu 2017: 86)

2.5-9 Alora ja sa vaza **jena trator rus** [...]

then I.NOM be.PRS-1SG take.PFV.PTCP.SG.M **one-ACC.SG.M tractor.ACC.SG.M Russian**

'So I bought **a Russian tractor** [...].'

(Breu 2017 : 97)

2.5-10 [...] **na spida veliki**.

ART.INDF-NOM.SG.M hospital.Nom.SG.M **big-NOM.SG.M**

'[...] **a big hospital**.'

(Breu 2011 : [1.1]-44)

2.5-11 Oni biša **na ljudun velki**!

DEM.PROX-NOM.SG.M be.IPRF.3SG ART.INDF-NOM.SG.M **man.NOM.SG.M big-NOM.SG.M**

‘That was **a tall man**!’

(Breu 2011 : [2.2.4]-15)

When we turn our attention to the positioning of adjectives, we can note that they can occur in both positions, before and after the noun that they modify. They do not have to be clustered together pre-nominally with other modifiers (2.5-12).

2.5-12 Alor moj did, biša **moj did stari** ka sa zovaša Česar [...].

so my.NOM.SG.M grandfather.NOM.SG.M be.IPRF.3SG **my.NOM.SG.M grandfather.NOM.SG.M old-NOM.SG.M** REFL call.IPFV-IPRF.3SG Cesare

‘Well, my grandfather, it was **my old grandfather** whose name was Cesare [...].

(Breu 2011: [1.1]-354)

It seems that possessives (2.5-12) and long form indefinite articles (2.5-7) force the adjective to the post-nominal position. The short form of the indefinite article does not necessarily have this effect on the adjective. In 2.5-5 and 2.5-6, an adjective is located between the article and the noun, but in 2.5-8 and 2.5-9 the adjective is moved behind the noun. 2.5-6 is a special case because it has two adjectives modifying the noun: one occurring before the noun and the other one behind. ITALIAN? We have seen that the classical claim that BCS (and MS) have two paradigms for adjectives marked explicitly for definiteness cannot be strictly supported. The difference lies in specificity rather than definiteness. Furthermore, in Molise Slavic as well as its closest relatives, there seems to increasingly be a functional split between the two sets regarding position within the sentence. This effect was not as strong as expected in our material, but the tendency is there. In addition, it is important to note that the distinction becomes increasingly lost for a wide range of adjectives.

2.6. Case distinction

In her grammar of Croatian, Kunzmann-Müller (2002: 198-199) mentions three ways in which definiteness can (rather than *must*) be expressed in Croatian. The grammar aims at readers with a German background, hence explaining the importance of providing them with a feeling for the subtleties of a language that does not have articles. One of the three options that are listed there is case distinction. Depending on the lexeme, the

opposition between the accusative case and genitive case for direct objects can also be used. In example 2.6-1, the contrast in Croatian is indicated. The accusative case specifies the entity. In German, the definite article can impose the same effect (2.6-2).

2.6-1a Daj mi **krumpira** (Croatian)
 give.PFV.IMP-2SG I.DAT **potato-GEN.SG.M**
 'Give me **a potato!**'

2.6-1b Daj mi **krumpir** (Croatian)
 give.PFV.IMP-2SG I.DAT **potato-ACC.SG.M**
 'Give me **the potato!**'

(Kunzmann-Müller 2002: 199)

2.6-2a Ich esse **Brot**.

2.6-2b Ich esse **das Brot**.

It is difficult to find exact correspondences for such contrasts in a relatively small corpus. A similar finding can be found in Birkenmaier's analysis of Russian definiteness, also on nouns that can have a countable versus uncountable reading (see Birkenmaier 1979: 106). To be precise, his aim is not exactly to encounter all cases in Russian where definiteness is expressed, but rather how the functions expressed by an article in article languages are expressed in Russian. He researches functions that are expressed by articles in German and by other means in Russian. By triggering a partitive meaning derived from its case meaning, the genitive indicates that some part of a larger whole is demanded without necessarily indicating what precisely is needed. This is by no means to say that those nouns are automatically indefinite. Eating a part of the bread or the whole bread deals with the same entity, known to the speaker and hearer once they bring up this topic. Based on my search in the corpus and additional literature sources, it is unclear whether Molise Slavic makes use of this contrast at all. It is easy to find examples in the accusative case but not in the genitive with verbs that could also be expected to be found with partitive meanings. In the following examples, I examine instances of the verb 'to eat':

2.6-3 Ja ne jidam **kruh!**
 I.NOM NEG eat.IPFV-PRS.1SG **bread.ACC.SG.M**
 'I don't eat **bread.**'

(Saint-Exupéry 2009: 68)

2.6-4 E činima sir, jidama **rikotu**, jidama **ma-čok**, tuna te stvare, seki dan.

And make.IPFV-PRS.3SG cheese.ACC.SG.M eat.IPFV-PRS.3SG **ricotta.ACC.SG.M** eat.IPFV-PRS.3SG **cream cheese.ACC.SG.M** all DEM-MED-ACC.PL thing.ACC.PL every-NOM.SG.M day-NOM.SG.M

‘And we made cheese, we ate **ricotta**, we ate **cream cheese**, all those things, every day.’

(Breu 2011: [2.1.1]-86)

All referents are marked by accusative case. Genitive forms are found only with quantifying expressions:

2.6-5 E nami dajahu **dujčjend gram kruha**.

and we.DAT give.IPFV-IPRF.3PL **two hundred.M gram.ACC.SG.M bread-GEN.SG.M**

‘And they gave us **200 gram of bread**.’

(Breu 2011: [1.2.1]-211)

2.6-6 [...] su smočil **na mala kruha** [...].

be.PRS.3PL wet.PFV-PTCP.PL ART.INDF-ACC.SG.M **litte-ACC.SG.M bread-GEN.SG.M**

‘[...] they dipped **some bread** into water.’

(Breu 2011 : [3.6.6]-31)

In 2.6-6, a clear case of a partitive meaning is given. However, in this case the preposition requires the use of genitive article. One could interpret this entity as not being marked for definiteness or specificity, because it is mainly about eating some bread without further reference to the bread itself. The actual central meaning in all instances is clearly that of partitive, definiteness perhaps being a secondary level of meaning that is sometimes added by context. Partitive meaning is expressed in analogy to Italian with a preposition *do* (ital. *di*) or with other types of quantifying expressions indicating a partitive meaning (‘200 grams’). Moreover, the examples 2.6-3 and 2.6-4 do not necessarily provide a definite interpretation either. In 2.6-3, generic reference is made to ‘bread’. The referents in example 2.6-4 should be considered specific rather than definite. It seems that the case distinction lost its functional load here. Additionally, the actual gradual reduction of case endings and the presumable beginning of the loss of the case system (development to an analytical type, in classical words) complicates the identification. Therefore, we can conclude that case distinction is not really a strategy for the expression of definiteness. Thus far, the picture is heterogeneous. Some categories (demonstratives, personal pronouns) are inherently definite by their lexical meaning, while others are somehow related to definiteness at

a more abstract level. We have seen that contextual factors play a crucial role in the identification of an entity as being specific or definite. Therefore, in the second part of this study we turn our attention to pragmatic factors that trigger definiteness in Molise Slavic.

3. The article system of Molise Slavic

3.1. Simple indefinite: jena

In his survey of the article system in European languages, the language area of Western South Slavic languages (BCS, Slovene) constitutes a “transitional zone, this time between two areas with articles, the Western area and the Balkans” (Schroeder 2006: 576). Besides the non-existence of a definite article, there is a broad body of research dealing with the question whether Slavic languages and particularly South Slavic languages have an indefinite article based on the numeral ‘one’. We find a mention of Serbo-Croatian indefinite article in Meillet and Vaillant (1980: 100): “Le numératif *jèdan* ‘un’ soit employé couramment en fonction de l’article indéfini.” By contrast, opposing claims are also found often in the literature:

When considering the indefinite article, there is no Slavic language that has fully developed this grammatical category to this day. Nevertheless, all Slavic languages do have the indefinite determiners⁴ as a functional category. (Belaj, Matovac 2015: 2).

By contrast, Breu states that the indefinite article in Molise Slavic covers the same functional range as the Italian indefinite article (cf. Breu 2017: 38). As will become apparent in the following, we are mainly dealing with a terminological issue here. Molise Slavic has an indefinite article and makes extensive (because obligatory) use of it. It is not the aim of this study to explore the functions and role of the indefinite article in depth (a task that has already been undertaken to a significant extent, cf. Breu 2012, Kreisberg 2007). For our discussion of the question whether Molise Slavic has a definite article at its disposal, we need to characterise some of the main features of what would presumably be its counterpart. There are several ways in which to introduce and describe this fully-grammaticised word form. As in many other

⁴ The term *indefinite determiner* is used in the study by Belaj and Matovac without theoretical implications or reference to a particular syntactic theory. They only want to make clear that it is not exactly the same as an indefinite article in the sense of English *a*, German *ein*, etc. (Cf. Belaj, Matovac 2015: 2).

languages in Europe and beyond, the indefinite article is based on the numeral ‘one’. They constitute the unstressed forms of the numeral *jéna* ‘one’ (see Breu 2017: 37-38). The numeral forms are stressed and carry a pitch accent in the same sense as in BCS. This accent can be rising or falling depending on some prosodic changes, which are not discussed here (cf. Breu 2017: 37 for a more detailed explanation of the phonological differences between the numeral and the indefinite article). Without paying attention to the two types of pitch accent found in the numeral forms, I only want to briefly illustrate that the difference between the numeral and the indefinite article is mainly achieved by the contrast stressed/pitch accent and unstressed/neutral pitch. This is shown in two pitch analyses conducted with the tool *Praat*. All analyses conducted with *Praat* can be found in appendix 3. The pictures are given in the appendix due to higher resolution for the sake of readability. The example of the numeral in the *praat* picture (1) clearly shows a falling pitch, while the examples of the indefinite article more or less carry a constant pitch (*praat* picture 2).

Short (*na*) and long forms (*jéna*) of the indefinite article exist. It is important to note that the short form⁵ (in the nominative and accusative case) is extensively used in the Molise Slavic texts of our corpus. Just like their long counterparts, they do not carry pitch accent in any sense. Before we finally turn our attention to the claim of a definite article in Molise Slavic, we will briefly focus on the grammaticisation of the indefinite article and its main functions from a comparative perspective. Before turning our attention to the functions that it covers within the language system nowadays, some comments on how the article came into being are in order. Breu describes the development of an indefinite article in Molise Slavic as a type of *pattern replication* (in the sense of Matras 2008, often called *calque* in contact linguistics). The use of the numeral *jéna* ‘one’ is extended following the model of *uno* ‘one’ in Italian:

Die Struktur des Italienischen ermöglichte zwar die funktionale Erweiterung von *jéna* ‘eins’ zum Indefinitartikel, verhinderte aber die formale Herausbildung eines definiten Artikels aus den Demonstrativa, da im Italienischen diese beiden Wortarten stets verschiedene Formen aufweisen, z.b. *il* ‘der’ ≠ *quello* ‘jener’.” (Breu 2017: 95).

⁵ It is impossible in this study to investigate the functional differences or distribution between the short or the long forms.

There is a strong tendency in the languages of the world to use the numeral having a singular meaning as a basis for an indefinite article. This observation has gone so far that this grammaticisation path has sometimes even been considered a universal or “quasi-universal” (Weiss 2004, quoted by Belaj, Matovac 2015). In a sample of 534 languages in the WALS, 112 languages (21%) use the same word form for the numeral one and the indefinite article. In counting only languages that have a grammaticised indefinite article at all (238), the share of languages with numeral-based articles is even higher (47%). This makes clear that it is not a globally-valid universal, but rather a frequently-used strategy in Europe. It is important to note that not all forms that are classified as indefinite articles exhibit the same range of different function. In the following, I illustrate the functional development of the indefinite article to sub-classify which type the Molise Slavic article is in contrast to other languages in Europe. Schroeder (2006) describes four stages in the functional extension of a numeral into an extended indefinite article:

(numeral) > emergent indefinite article > pragmatic indefinite article > referential indefinite article > extended indefinite article

Before we discuss the single stages, a second, more fine-grained grammaticisation path (based on Givon, Heine and Heine Kuteva) is presented. This was formulated by Belaj, Matovac (2015: 3):

numeral > presentative marker > specific indefinite marker > non-specific marker > generalised article

When we compare the two grammaticisation paths in the sources quoted here, we note that Schroeder’s stages are more precise at the boundaries, adding a second stage between numerals and the presentative marker, namely in the form of the emergent indefinite article. By contrast, Belaj, Matovac’s (2015) list is more detailed in introducing more distinctions in the middle way of the path. Hence, combining the two leads to a more structured and detailed tool for analysis:

numeral > emergent indefinite article > pragmatic indefinite article (presentative marker > specific indefinite marker) > non-specific indefinite marker/ referential indefinite article > extended indefinite article > generalised article

In this scale, both terminologies are included to see which terms roughly equal each other conceptually. On the **numeral stage (3.1-1)**, no component of specificity or definiteness is included in the NP (3.1-1):

3.1-1a Svako jutro pojedem **jedno jaje**. (Croatian)
 every.ACC.SG.M morning.ACC.SG.M eat.PFV-PRS.1SG **one-ACC.SG.N egg-ACC.SG.N**
 'Every morning, I eat **one egg**.'

(Belaj, Matovac 2015: 4)

3.1-1b Mitaj još **nu divojku!** (Molise Slavic)
 invite.PF-IMP.2SG more **one-ACC.SG.F girl-ACC.SG.F**
 'Invite **one more girl!**'

(Breu 2012: 284)

When the numeral can be used to introduce "a new and major participant" into discourse, it serves as the function of a **presentative marker (3.1-2)** or pragmatic indefinite article (see Belaj, Matovac 2015: 4). Russian and Polish are claimed to be at this stage now, while Czech and BCS are already taking up increasingly more functions of the next level (see Belaj, Matovac 2015: 4). In other accounts, BCS is still considered to remain at this stage (see Reichenkron (1966: 345; Schroeder 2006: 576).

3.1-2a Bio jednom **jedan kralj**. Taj kralj je imao... (Croatian)
 Be-PTCP.SG.M once **one.NOM.SG. king.NOM.SG.F** DEM.MED-NOM.SG.M king-
 NOM.SG.M be.PRS.3SG have.IPFV-PTCP.SG.M
 'Once there was **a king**. This king had...'

(Belaj, Matovac 2015: 4)

3.1-2b Biša nu votu **na žena stara**. (Molise Slavic)
 be.IPRF.3SG ART.INDF-ACC.SG.F time-ACC.SG.F **ART.INDF-NOM.SG.F woman-**
NOM.SG.F old-NOM.SG.F
 'One time, there was **an old woman**.'

(Breu 2012: 281)

The numeral develops into a **specific indefinite marker (3.1-3)** when it can be used for the mention of specific participants who are considered major participants of the discourse. In their analysis, Heine and Kuteva (2006) consider Bulgarian and Macedonian as being at this stage. In contrast to their account, Belaj assumes that Croatian is already at this stage (see 3.1-3), while Bulgarian, Macedonian and Upper Sorbian are further along the grammaticisation path (see Belaj, Matovac 2015:13). At

this stage, the indefinite article can be used for all indefinite NPs without limitation due to pragmatic factors (see also Schroeder 2006: 557):

- 3.1-3 Čula sam to od **jedne prijateljice**. (Croatian)
 hear.PFV-PTCP.SG.F be.PRS.1SG DEM.MED-ACC.SG.N from **one-GEN.SG.F friend-GEN.SG.F**
 'I heard it from **a friend**.'

(Belaj, Matovac 2015: 4)

When a non-specific entity is introduced or mentioned together with the numeral form, it is likely to be a **non-specific indefinite marker (3.1-4)**. This is the case in Upper Sorbian, where the indefinite form can be used for non-specific reference, although there is often the restriction that this only applies to singular countable nouns. The indefinite article is used with all indefinite noun phrases as well as with non-referential noun phrases (see Schroeder 2006: 559). In examples 3.1-4b and 3.1-4c, the direct object NPs are not specific. This is a typical function of the indefinite article in Molise Slavic and hereby confirms its status as a true article.

- 3.1-4a Ja cem **jen mikser** měć. (Upper Sorbian)
 I.NOM want.PFV-PRS.1SG **ART.INDF-ACC.SG.M mixer.ACC.SG.M** have.IPFV-INF
 'I want to have **a mixer**.'

(Belaj, Matovac 2015: 8)

- 3.1-4b Ja jiskam **na mičicij**. (Molise Slavic)
 I.NOM search.IPFV-PRS.1SG **ART.INDF-ACC.SG.M friend-ACC.SG.M**
 'I am looking for **a friend**.'

(Breu 2012: 282)

- 3.1-4c Zov **na medik**! (Molise Slavic)
 Call.PFV-IMP.2SG **ART.INDF-ACC.SG.M doctor.ACC.SG.M**
 'Call **a doctor**!'

(Breu 2012: 284)

In conclusion, Molise Slavic reaches this stage in the grammaticisation but goes even further in some contexts (see example 3.1-6). In the last stage, the indefinite noun can be used with all types of nouns, in some languages even with plural (Spanish, 3.1-5 or mass nouns (Italian and Molise Slavic, 3.1-6).

- 3.1-5 **Unas mujeres** llegaron al edificio. (Spanish)
ART.INDF-NOM.PL.F woman-NOM.PL arrive-PRT.3PL at building-SG.M

‘**Some women** arrived at the building.’

(Belaj, Matovac 2015: 4)

The indefinite article in Molise Slavic is not grammaticalized to the stage of Spanish but covers all functions as its Italian counterpart. For instance, it is even possible to use the indefinite article (optionally) with mass nouns:

3.1-6 Nonda biša ø/ **na trava niska friška**.

(Molise Slavic)

there be.IPRF-3SG ART.INDF-NOM.SG.F grass-NOM.SG.F low-NOM.SG.F fresh-NOM.SG.F

‘There was **low soft grass**.’

(Breu 2012: 285)

It is important to remember that this scale describes the distribution of functions for the same word form. It is rare that the form fully gives up some of the earlier functions when progressing in the path (as in English “a”, not covering the numeral meaning). However, phonological changes often take place, as just described earlier. It is safe to say that Molise Slavic is at the stage of a non-specific indefinite marker, as becomes clear from examples like in. For other uses, e.g. in generic NPs, see Belaj, Matovac (2015: 5) for more theoretical deliberations.

It is unclear in which cases the short or long form of the indefinite article is preferred. In 3.1-7, both types are used alternating:

3.1-7 Nu votu **na zec** prohodaša jizbane **jene rike**.

ART.INDF-ACC.SG.F time-ACC.SG.F ART.INDF-NOM.SG.M rabbit.NOM.SG.M
pass.IPFV-IPRF-3SG on.side ART.INDF-GEN.SG.F river-GEN.SG.F

‘Once there was **a hare** walking along **a river**.’

(Breu 2017: 89)

In conclusion, we have seen that Molise Slavic has an indefinite article that is highly developed and further along the grammaticisation path in comparison to European languages, but only one step away from Croatian. Breu mentions a last context in which the indefinite article is used obligatorily: in rhematic reference, the use of the indefinite article is obligatory in Molise Slavic (cf. Breu 2012: 282):

3.1-8 Sfe skup je uliza **na ljud tusti**.

always together be.PRS.3SG enter.PFV-PTCP.SG.M ART.INDF-NOM.SG.M
man.NOM.SG.M fat-NOM.SG.M

‘Suddenly **a fat man** came in.’

(Breu 2012: 282)

This indicates that the ordering of the sentence in a theme and rheme (or topic and comment) affects the assignment of definiteness to a nominal phrase. In the first part of the study, we mainly considered modifiers that assign definiteness to a nominal phrase. In the following, we will examine how the actual structure of the sentence can directly entail a definite interpretation on a noun or nominal phrase. If the information structure determines the definiteness of a phrase, it is questionable whether the concept of a zero article as a contrasting partner of the indefinite article is truly necessary. We will explore this question in the remaining parts of this study. The hypothesis is that contextual factors may be more important in assigning definiteness to a noun.

3.2. Simple definite: zero article

As we have seen in the previous chapter, Molise Slavic is the Slavic language with the furthest grammaticised indefinite article. This fact is even more striking when we consider that it did not develop a definite article in the morphological sense. In Breu's view, this constitutes a violation of a universal in language change, namely that definite articles first appear before indefinite articles in a given language (cf. Breu 2017: 90). To heal these diagnostics, he introduces the idea that there is a definite article in the form of a zero morpheme. Since all contexts in which an indefinite article typically occurs are marked by this article, he hypothesises that nouns not marked with the indefinite article are automatically marked for [+definite]:

Unabhängig davon, wieweit die ursprüngliche Zahlbedeutung abgeschwächt wurde, steht der Indefinitartikel obligatorisch, wenn Definitheit ausgeschlossen ist. Das heißt umgekehrt, daß die beiden Substantive in s1 nur eine definite Lesart haben, es liegt ‚definitiver‘ Nullartikel vor. Im Endeffekt ergibt sich aus der Opposition von jena:ø ein vollständiges Artikelsystem.“ (Breu 2017: 95).

The concept of zero articles is refused here. As we have seen in the previous chapter, Croatian also makes use of a type of indefinite marker and still does not require the use of a definite article. I predict that from a comparative perspective with other Slavic languages, the concept is not necessary in accounting for definite interpretations of referents without having a definite article. Another problem stems from a terminological deliberation, because it is used in different contexts. In theories of

relatedness, it is used for instances of nouns without an article that still have some type of definite reading (cf. the discussion in Birkenmaier 1979: 25-26.). It is a type of article contrasting with the definite article on the one hand and the indefinite article on the other, as in the German example in 3.1-9:

3.1-9 Sie waren **Soldat** und wissen, was ein Befehl ist. (Birkenmaier 1979: 26).

The aim of the second part of this study is to research whether the concept of zero articles – as postulated by Breu for the variety under discussion – is necessary at all. I will test whether the interpretation of a NP or clause as definite or indefinite by the speaker can be explained by other factors, namely by analysing the role of topicality. Without exploring this aspect from a theoretical perspective, it is noteworthy to mention instances in the Molise Slavic data in which a form without article (hence, in Breu’s analysis, a definite NP) cannot be interpreted as being definite:

3.1-10 Ja mahu kupi **trator** nonda, one dana.

I.NOM must-IPRf-1SG buy.PFV-INF **trator.ACC.SG.M** then DEM.DIST-ACC.PL day-ACC.PL.M

‘At that time, I had to buy **a tractor**, those days.’

(Breu 2017: 97)

This sentence is uttered at the beginning of a piece of discourse. It might be possible to interpret the NP tractor as being [+spec] if we assume that the speaker has the specific tractor that he bought in mind (and that he mentions later). However, it rather appears that up to this point in the utterance tractor does not refer yet to any specific, identifiable entity. Nonetheless, the alleged definite form is used although we would expect marking with an indefinite article. By contrast, in example (3.1-11) the direct object of the sentence is located in the post-verbal position of the sentence and should be understood as [+def], because the entity denoted was already mentioned twice in the discourse, once even explicitly marked with a demonstrative:

3.1-11 E naka zec je zgubija **skumasu**.

and so.DIST rabbit.NOM.SG. be.PRS.3SG lose.PFV-SG.M **bet-ACC.SG.F**

‘And so the hare lost **the bet**.’

(Breu 2017: 94)

Furthermore, in generic contexts it is unclear whether a generic NP needs to be interpreted as being specified as [+def] only based on the morphosyntactic analysis of not being marked with an indefinite article.

3.1-12 Čeljade velke jesu činjene naka.

people-NOM.PL big-NOM.PL be.PRS-3PL make.PFV-PTCP.PASS.PL so.DIST

'Grown-ups are like that.

(Saint-Exupéry 2009: 19)

Generic contexts are not discussed here. However, how this function is expressed in Molise Slavic requires further research. If Breu's analysis is correct, generic meaning can be expressed with a definite noun like in French, optionally German, but not English. Alternatively, it is possible that a morphologically unmarked, bare noun can be used for a variety of different, unrelated functions; for example, generic use as in the English translation in 3.1-12, predicative sentences as in German (3.1-9), unspecific reference as in 3.1-10 as well as clearly-definite contexts (both referents in 3.1-11). If this assumption is true and the bare form of nouns can be used for such a broad variety of contexts, other factors in discourse must play a role in assigning a definite reading of a noun. Thus far, we have investigated how modifiers can be used for this purpose. Additionally, it is possible that features of the verbal domain – in particular aspect – also contribute to this topic. Given the limitations of this study, no exploration of the verbal system could have been undertaken. A third option is that more abstract levels of the language system enhance our understanding of what is to be considered definite or otherwise. In the next chapters, we turn our attention to the question of how information structure and its related component intonation contribute to the category of definiteness in Molise Slavic.

4. Topicality and zero articles: The interplay of information structure and definiteness

4.1. Theoretical deliberations: Referentiality and information structure

At this point, we are leaving the area of morphosyntactic descriptions and approaching the concept of definiteness from the wider perspective of discourse and communication. Language manifests itself in its use in discourse and the purpose of referring to actual events and entities in the outer world is one of the crucial features of this human activity. The following investigation is largely informed by research in pragmatics, particularly in the field of information structure (IS) research. It is necessary to clarify some of the notions and concepts since there is huge confusion about the use of certain terminology in the field. Starting with the overarching level, I understand “[...] discourse [as] an (interface) module that is concerned with information pertaining to individuals and events and the relations between them [...]” (Burkhardt 2005: 5). One linguistic level that is active between the actual linguistic form and the communicative needs is information structure:

The term information structure refers to the ways linguistically encoded information is presented relative to the speaker’s estimate of the temporary mental state of the receiver of the message. (Matić 2015:95).

It is less about the actual content of a message and more about how the items of the message are structured and presented to the hearer by considering communicative needs. In line with processing needs of the human mind and general requirements of communication as a cooperative activity, information structure determines how and in which order information is presented, as well as which parts of the information are emphasised or defocussed:

As a bridge between mental states and linguistic form, information structure serves as a means of signaling the mental states of the speakers with respect to the information they intend to convey, and thus provides comprehenders with information that could help guide the processing and interpretation of sentences. (Cowles 2012: 293).

The concept itself is seen mentally but with clear and direct consequences for the structuring of the linguistic code:

[...] it seems clear that models of language production conceive of information structure status as a reflection of the mental representations of the speaker, and what the speaker intends to communicate to an addressee. (Cowles 2012: 289).

The term itself was coined by Halliday (1967), although the notion behind this concept dates back to a long discussion since antiquity (cf. von Heusinger 1999). It is mainly about how information is presented, rather than what the information content is about. Chafe (1976) specified this under the term *information packaging*. In discourse, two (or more) interlocutors are sharing a set of information that is mutually known by both. This shared set of information is called *common ground* (CG) in IS research. In the process of communication, the common ground is constantly updated and the information structure guides the hearer by providing assistance on how to process the content.

This idea of information structure is rooted in a particular model of communication, the incremental model (Stalnaker, 1999). According to this model, communication consists of reducing the differences in the knowledge of the interlocutors by increasing the common ground between them, i.e., their stock of shared knowledge. (Matić 2015: 95).

This will become important in the following discussion of topicality, because it is fair to say that effective communication requires the cooperation of the interlocutors in terms of agreeing on a topic and adding information to it when discourse unfolds:

That is, the common ground consists not only of a set of propositions that is presumed to be mutually accepted [...], but also of a set of entities that have already been introduced into the common ground previously." (Krifka, Musan 2012: 3).

It is noteworthy that the main ideas of information structural concepts date back to Mathesius and the Prague Functionalist School, therefore being acknowledged in Slavic studies tradition ever since:

The decisive move from psychology to linguistics was undertaken by the linguists of the Prague School, especially Vilém Mathesius, who used the categories derived from psychology and philosophy to account for phenomena of word order variation and prosody." (Matić 2015: 97).

Before we turn our attention to the smaller units of information structure – namely topic and focus – one last clarification is in order: there is some discussion in the field concerning which levels are covered by the notion of information structure. On the one hand, it is unresolved if the concept only affects the sentence level, and therefore only propositional entities can also cover the component parts of such entities. (cf. Matić 2015: 96) On the other hand, there is some discussion on the question whether information structure should be restricted to this level or if it can also cover broader pieces of discourse. For our purpose, it is necessary to clearly distinguish the concepts of information structure and discourse structure. The first is limited to the structuring

within the sentence⁶, while the latter related to the larger unit of discourse. The former affects the sentence level:

Information Structure is a Sentence Internal partition of the information in an utterance according to its relation to the discourse context under dichotomies such as topic/comment, theme/rheme, given/new, focus/background, etc. Such categories are essentially referential in nature. (Steedman, Kruijff-Korbayová 2001: 1)

By contrast, discourse structure is the level that deals with the structuring and alignment of larger pieces of discourse:

Discourse Structure concerns the Inter-clausal relations of explanation, elaboration, exemplification, and illocutionary force that hold between successive utterances of a discourse or dialog, supporting inference about the domain and purposes of the discourse.” Steedman, Kruijff-Korbayová 2001: 2

Following this distinction, we can also assume the existence of *sentence topic* and *discourse topic* as two distinct categories, which will be further elaborated in the next chapter. For instance, in the example text *Zec aš rak* (see appendix 2), both the hare and the crayfish are discourse topics of this story. Nonetheless, only one of the two serves as a sentence topic in the single utterances. In our discussion, discourse structure is only relevant regarding CG management. We are assuming that only what is already in the common ground can serve as a topic of a sentence. Everything that was mentioned in the discourse or is in the immediate awareness of the interlocutors (in the context) can qualify for this (cf. Erteschik-Shir 2014: 27). There is general agreement that information structure and definiteness are somehow related in language, whereby some accounts even suggesting that definiteness or at least one type of definiteness is subsumed as a sub-category of information structure itself.

Interestingly, already older accounts in Russian linguistics consider the morphologically overt marking of (in)definiteness with articles as one of several possible strategies for the communicative sentence perspective (which more or less equals the nowadays concept of information structure). They relate the content of the utterance to the interlocutors and make claims about the information that the speaker and the hearer have about an entity mentioned. The definiteness – indefiniteness dichotomy is hereby similar to the idea of given – new, but not identical. It is

⁶ There is discussion in the literature, if IS only affects sentence level or also smaller (or larger) units of the sentence. See Matić 2015: 96 for further notes on this discussion.

language-specific which type of strategy (morphological marking, syntactic ordering, special constructions or intonation) is used for contrasting nouns for these features:

Bestimmtheit/ Unbestimmtheit ist also in den Sprachen, in denen diese Opposition durch den Artikel ausgedrückt wird, eine morphologische Kategorie innerhalb der syntaktischen Kategorie der kommunikativen Satzperspektive oder eine auf Substantive beschränkte kommunikative Satzperspektive“ (Birkenmaier 1979: 29).

It has also been noticed in Russian linguistics that the category of definiteness depends on the actual context of an utterance:

‘Bestimmt’ und ‘unbestimmt’ sind nicht immanente Eigenschaften eines Gegenstandes, sondern ‘sind durch die Stellungnahme des Sprechers und des Hörers in der Sprechsituation bedingt. Entscheidend bei dieser Kennzeichnung ist der Kenntnisstand des Gesprächspartners über den in Rede stehenden Gegenstand’“ (Moskalskaja 1971, 202. Zitiert nach Birkenmaier 1979: 31).

The relation between information structure and definiteness becomes clearer when we consider the role that both play in discourse. As previously mentioned, we are separating the two levels discourse structure and information structure. Nonetheless, this is an analytical assumption following our research purposes. In human communication, the two levels are necessarily intertwined. In the context of discourse representation theory (DRT), the notion of common ground was extended to discourse referents (Kamp 1981, Heim 1982). The concept of *referentiality* is understood in my study as the heading concept for definiteness. In our exploration, we are mainly dealing with referring expressions:

[...] a referring expression is a linguistic form that the speaker uses with the intention that it correspond to some discourse entity and bring that discourse entity to mind for the addressee. (Birner 2013: 111).

In our discussion, referring expressions in the form of deictics, definites, indefinites, anaphoric expressions and demonstratives hold relevance, although other types also exist. A (discourse) referent is the entity to which the referring expression is referring. By contrast, definiteness is defined from two different perspectives. In uniqueness accounts, only unique referents can be definite. Sometimes this view is preferred over identifiability accounts:

Following Heusinger (2002), I take definiteness to express uniqueness of items which are not necessarily identified, while specificity expresses referential dependency between items introduced in the discourse. (Tomić 2008: 450).

In identifiability accounts it is less about the uniqueness of the entity or referent, but about the information the speaker provides the hearer to process the reference to an

entity: “[...] certain noun phrases refer to entities that the speaker judges should be identifiable by the addressee.” (Payne 1997: 263). I do not want to proceed deeper into this discussion. In my understanding for this study, I consider the three categories of referentiality, specificity and definiteness as being closely related and defined regarding the role that they play in discourse. We distinguish three types of referentiality in our discussion. Objective referentiality is understood as a more general property that a referent must have: “An entity is objectively referential if it exists as a bounded, individuated entity in the message world” (Payne 1997: 264). By principle, this is in line with uniqueness accounts. As previously mentioned, every referent that the speaker mentions and which he knows in a way can be considered as being specific. Once the entity is brought to the awareness of the hearer by being introduced into the discourse, the referring expression is definite after the first mention. In other words, an entity that exists and to which we can refer is referential. If a speaker has awareness of this entity and can single it out from other entities or attribute it to a subset of entities, the referents of this entities should be considered specific. Once he introduces these entities to the set of referents active in the shared awareness of the discourse (common ground) and when the hearer is able to identify this entity, it is seen as definite. This view is strongly influenced by practical deliberation concerning communication. Note that specific referents are sufficient for the preservation of discourse. All of these notions do not deal with the question of whether the entity is really known to the same extent by speaker and hearer. A referent is already definite when the hearer has minimal information about him, even when it is only the knowledge of the existence of the referent (*principle of minimal information*).

4.2. Topicality in Molise Slavic

4.2.1. The linguistic concept of *topic*

Slavic languages (except Bulgarian and Macedonian) are prominent in the European context for not having an article system. However, every language must possess some means for expressing definiteness in the sense of a semanto-pragmatic category, albeit not always obligatory on every NP. Therefore, alternative strategies for the expression of definiteness are regularly mentioned. In particular, word order alternations and intonation are known to trigger this interpretation. In example (4.2-1 and 4.2-2), the sentence-initial position is claimed to impose a [+def] interpretation on the entity, although it is important to note that the authors clearly only mention a tendency to be at work (see Sussex, Cubberley 2006: 418). In the Croatian example (4.2-3), the two entities *crni* and *bijeli vragovi* ‘white and black devils’ are first introduced in rhematic, post-verbal position. In the second sentence, they are in the sentence-initial, pre-verbal position. It is claimed that the position marks them for definiteness. Since they are introduced in the common ground, they occur in the first position of the sentence.

- 4.2-1 [Dziewczynka]_{TOP} [dała mi bukiet kwiatów.]_{COM}
(Polish)

girl-NOM.SG.F give-PST.SG.F I-DAT bouquet-ACC.PL flower-GEN.PL
The girl gave me a bouquet of flowers

- 4.2-2 [Bukiet kwiatów]_{TOP} [dała mi dziewczynka.]_{COM} (Polish)

bouquet-ACC.PL flower-GEN.PL give-PST.SG.F I-DAT girl-NOM.SG.F
The bouquet of flowers was given to me by a girl.

(Sussex/ Cubberley 2006: 418)

- 4.2-3 [Na svijetu postoje **crni i bijeli vragovi**.]_{COM} (Croatian)
on earth-DAT.SG.M exist.IPFV-PRS.3SG **black-NOM.PL** and **white-NOM.P** devil-NOM.PL
‘On earth, there are **black and white devils**.’

[**Crni vragovi**]_{TOP} [su užasno opasni vragovi.]_{COM} (Croatian)
black-NOM.PL devil-NOM.PL be.PRS.3PL uncannily dangerous-NOM.PL devil-NOM.PL
‘**The black devils** are dangerous devils.’

[**Bijeli vragovi**]_{TOP} [nisu opasni [...].]_{COM}
white-NOM.PL devil-NOM.PL NEG-be.PRS.3PL dangerous-NOM.PL

'The white devils are not dangerous [...].'

(Kunzmann-Müller 2002: 99)

In some accounts, the function of the article in article languages is equalled to the role that word order and intonation can play in Slavic languages. In other words, word order and intonation can be used to express the article function in article languages. For this claim, see the following two examples. In one grammar of Croatian, word order and intonation are given as equivalent to the German article:

Das Kroatische als eine artikellose Sprache hat eine vergleichbare morphologische Kategorie nicht. Das gen. Merkmal wird demzufolge nicht obligatorisch gekennzeichnet. In der Organisation von Rede und Text ist aber die Kennzeichnung von Neuerwähnung vs. Nichtneuerwähnung relevant und es existieren Mittel der formalen Charakterisierung. (Kunzmann-Müller 2002: 99)

In his study of definiteness in Russian, Birkenmaier puts it even more explicitly:

In den Artikelsprachen haben die Artikel die Aufgabe, ein Substantiv als Träger bekannter oder unbekannter Informationen zu charakterisieren. In den Sprachen ohne Artikel wird die gleiche Funktion von Wortstellung und Intonation übernommen. Das bedeutet, daß in bestimmten Wortstellungsmustern bestimmte Positionen als thematisch markiert sind. Das gilt jedoch nur, wenn der Hauptakzent am Ende des Satzes liegt und keine Kontrastbetonung vorhanden ist. (Birkenmaier 1979: 67).

Both quotes have been included to illustrate the common view in literature that word order and intonation in Slavic can fulfil the exact function of the article in article languages. Both hint that the link between these two notions lies in the thematic status of the respective NP. This thematic status is one of the sub-domains of information structure, namely the division of a sentence into topic and comment (or theme and rhema in older accounts). To sum up the first point just made, I want to repeat that explicit claims are made regarding Slavic languages that the expression of definiteness in these languages that do not exhibit articles is largely determined by word order driven by information structure. The notion of topic is assumed to be responsible for this effect. Therefore, I understand topic as a mediator between the semanto-pragmatic need to express definiteness and the morphosyntactical means of word order and intonation (for a summary of this view, see Trenkic 2004: 1403). It is argued in these accounts that the informational structural notion of topic is responsible in one way or the other for causing a nominal phrase to be interpreted as definite (or indefinite if it does not function as topic in a sentence). The final goal of this chapter

is to test these claims on our Molise Slavic material. On the one hand, the role of topics and topicalization strategies is contrasted with Breu's assumption of a grammaticised zero article in this language. Analysing the impact of topics on the definite reading of a clause in Molise Slavic, we are trying to decide whether the concept of a zero article is necessary and if it captures the concept of definiteness in an adequate way. On the other hand, the study contributes to the information structure research on Slavic languages. It remains far from fully understood how and in what sense topicalization and definiteness are related in these languages and beyond. First, we need to debate the notion of topics and evaluate the tools and tests proposed for identifying topics in a given piece of discourse. Despite the fact that the aforementioned claim in case of Slavic languages is often made, the concept of topic itself is not questioned in many of these studies, although "[t]he notion of 'topic' has, unfortunately, been defined in a number of distinct and sometimes conflicting ways" (Birner 2013: 212). I take up the notion as identified by Krifka and Musan in their seminal work on information structure: "The topic constituent identifies the entity or set of entities under which the information expressed in the comment constituent should be stored in the common ground content" (Krifka, Musan 2016: 28). In other words, the topic is what a sentence is about and the comment provides further information about this topic. Other concepts also exist, but what I examine here is what is sometimes called *aboutness topic* or *sentence topic*: "The most common use of the word *topic* is with respect to the sentence topic, defined as what the sentence is about" (Gundel 1989; Lambrecht 1994). It is necessary in our study to distinguish sentence topic from discourse topic (cf. Birner 2013: 213). Topics of sentences can connect virtually by reference. Nonetheless, this connection is at the level of discourse topic. Sentence topics themselves are autonomic. For defining topics in a more general way, the definition by Reinhart (1981) is used here:

Sentence topics are a pragmatic phenomenon which is specifically linguistic [but related to non-linguistic aspects of human interaction]: Only sentences can have a sentence topic, and what the topic of a given sentence is is determined both by its context of utterance and by its linguistic structure (Reinhart 1981: 53).

When I use the term *topic* in this study, I am referring to *sentence topic* in the sense of Reinhart (1981), which is equivalent to the concept of *theme* in the sense of Prague

Functionalism. Later research has identified a number of different types of topic. In this sense, Reinhart's analysis focusses on the notion that is nowadays labelled *aboutness topic* in the literature. Contrastive topics and other phenomena are not considered here. More specifically, Lambrecht defines the topic as "the matter of current interest which a statement is about and with respect to which a proposition is to be interpreted as relevant" (Lambrecht 1994: 119). In terms of common ground, the topic must be an entity that is present where both interlocutors are aware of the entity that the sentence is about. The structuring into topic and comment therefore provides instructions to the hearer how to interpret the sentence and helps him to build up the dependency of knowledge that he has about the entities in discourse (see Reinhart 1981:80, also for the famous file-card model of topics which is not reproduced here). We assume that every sentence has a topic due to this communicative need. A sentence at the beginning of a discourse or more general every all-new focus expression in which the whole sentence comprises new information and in which the sentence is focused is also aligned for topic and comment (see Erteschick-Shir 2014: 25). In this sense, *new topics* do exist (Krifka, Musan 2012: 28). Nonetheless, there is some restriction regarding referentiality whereby entities of a sentence can function as topics. The main entity that serves as a topic is a NP. This is not a necessary but practical condition (cf. Reinhart 1981). Lambrecht emphasises the fact that under his notion of topic it is possible for every grammatical category in a sentence to be the (sentence) topic (cf. Birner 2013: 213). However, for the sake of brevity in this account only nominal entities are granted topichood. Nouns being topics must be able to be identifiable. Hence, they must be at least specific from the moment that they serve as a topic:

Only referential expressions serve as topics. Topics are prototypically referential DPs with a discursial antecedent. (Erteschick-Shir 2014: 24).

We assume that entities in the common ground (or in the actual surrounding of speaker and hearer) can be taken as a topic, while deictic references can also serve as a topic. It is also relevant to note that: "Languages mark topics in a variety of different ways. Topics can be marked by topicalization, by a (clitic) pronoun, morphologically, by topic drop or by intonation (including destressing)" (Erteschick-Shir 2014:24). In my

analysis, brackets around a NP or clause mark information structural constituents. Finally, the question arises how to identify topics in a sentence.

4.2.2. Topic tests

A few testing criteria for the identification of topics have been proposed in the literature. Some are oriented at the linguistic form itself, like linear order, grammatical terms or intonation (e.g. non-stressed/de-accentuated expressions as topic; cf. Reinhart 1981: 56 for an overview). Those accounts are insufficient in identifying topics because they make no reference to the role that they play in context. For the same reason, I refuse the topical hierarchies as outlined by Firbas (1975) or Givon (1976) because they contain parameters like (semantic) type of verb or definiteness. These are morphosyntactic or semantic criteria. Topic is a pragmatic category and should be defined by pragmatic features. Morphosyntactic and prosodic features related to information structure are determined by the topic – comment alignment instead of imposing it on a sentence.⁷ A third possibility to infer topics from structural examinations is also rejected here given that MS is not yet sufficiently researched in this regard. One could start by identifying topicalization strategies to identify topics. This is a circular approach, because it would be unclear what serves as a topicalization strategy as long as we have not made clear what a topic is. Nonetheless, based on typological work, some strategies (e.g. Left dislocation and other sorts of fronting, topicalization particles, etc.) are well known to mark or even emphasise topics in other languages. This is taken up in chapter 3.3.3. where an initial exploration is given that remains incomplete to date. Alternatively, a set of question tests were elaborated to check for topics. None of the tests can fully account for all types of topics nor identify them absolutely precisely (for examples and critique, see Birner 2013: 214).

For a number of reasons, then, the term ‘topic’ has not proven as useful as one might hope. For now, it may be safer to avoid the term altogether and search instead for what lies behind the intuition that certain constituents are given, or topical, and others are not. All of the above tests are related to the intuition that, for a constituent to be a topic, it must have a certain degree of pragmatic accessibility. (Birner 2013: 214).

⁷ In my understanding of language production models, I expect IS assignment to occur earlier in the language production process than grammatical encoding because information structure is a mediator between the (not language specific) conceptual level and the (language specific) encoding mechanisms. This view is informed by the Level model of language production (see Levelt 1989).

We do not have to go as far as Birner to give up on the concept fully, because we can combine several sub-features into one coherent testing tool. As Birner emphasised, taking pragmatic or contextual factors into account seems important. The concept of common ground provides such a test. To account for several features of topic, in the following a set of criteria is presented that is used in this study to identify topics on the material. These criteria are largely informed by the tools elaborated by Reinhart (1981), because she has offered one of the best testing tools to identify topics to date. The testing tools in Reinhart are more formally elaborate, but we reduce them to their main statements. In order to make up for this formal weakness I add two criteria not considered (albeit recognised) by Reinhart. Each criterion is described with an example from the literature or with a Molise Slavic example. The features (1) to (4) were applied to all sentences, while the additional feature (5) intonation was only applied to the second text and third test in the testing sample.

Criterion 1: Singularity of topic and focus within a sentence

Every sentence has one topic and one comment (cf. Reinhart 1981: 56). The notion of several topics within one sentence is refused here, because it is more practical to subsume them within one topic. A sentence can also have no topic. It can have no topic at the beginning of a discourse or when a discourse referent is introduced completely new (cf. Reinhart 1981: 56 and 67; Lyons 1999: 229; see text examples in the appendix). Other accounts allow several topics within one sentence (e.g. Erteschick-Shir 2014: 23). This assumption is refused here for the sake of simplicity. If and which entity is chosen as a topic depends on the context. Subordinate clauses have their own topic, which can be distinct from the topic of the main clause (4.2-5).

4.2-4a [Chlapec]_{TOP} [odnesl dřevo do kůlny]_{COM} (Czech)
 boy.NOM.SG.M carry-away.IPFV-PTCP.SG.M wood.ACC.SG.N to shed-GEN.SG.F
 ‘The boy took wood to the shed.’

4.2-4b [Dřevo]_{TOP} [odnesl do kůlny chlapec]_{COM}
 wood.ACC.SG.N carry-away.IPFV-PTCP.SG.M to shed-GEN.SG.F boy.NOM.SG.M
 ‘The wood was taken to the shed by a boy.’

(Birkenmaier 1979: 49-50)

- 4.2-5a [Ovi ašterojid]_{TOP} [je vidan lu 1909 do jenga aštrónoma turk]_{COM}. (Molise Slavic)
 DEM.PROX-NOM.SG.M asteroid.NOM.SG.M be.PRS.3SG see.PFV-PTCP.SG.M in 1909 by
 ART.INDF-GEN.SG.M astronomer-GEN.SG.M Turkish-GEN.SG.M
 ‘This asteroid was seen in 1909 by a Turkish astronomer.’
 (Saint-Exupéry 2009: 18)
- 4.2-5b [Aštrónom]_{TOP} [je čija jopa njevogu dimoštracijunulu.]_{COM}
 astronomer.NOM.SG.M be.PRS.3SG make.PFV-PTCP.SG.M again his-ACC.SG.F
 demonstration-ACC.SG.F
 ‘The astronomer held his presentation again.’
 (Saint-Exupéry 2009: 19)
- 4.2-5c [ovi {ka [mi]_{sub-TOP} [ga [proda]_{sub-FOC}]_{sub-COM}}]_{TOP} [je [umbra]_{FOC}]_{COM}.
 DEM.PROX-NOM.SG.M REL I.DAT he.ACC sell.PFV.PTCP.SG.M be.PRS.3SG
 die.PFV.PTCP.SG.M
 ‘the one who had sold it to me died.’
 (Breu 2017: 99)

Depending on the position of the NP, it must be considered marked for definiteness, although the content of the sentence remains the same: “[...] it is a crucial fact about sentence topics, that equivalent sentences may have different topics (even if they mention precisely the same referents” (Reinhart 1981: 58).

Criterion 2a: Principle of presumption of knowledge (Strawsons 1964: 96)/
 Active reference in CG

The important point is that the entity to which the NP refers is already in the common ground of both interlocutors or can be derived easily from context or world knowledge. If so, the NP is likely to be the topic. If not, it should be the comment of the sentence. Here, it is important that this does not fully equal the notion of given and new information: “[...] although in most cases the topics tend indeed to represent old information, this is neither a sufficient nor a necessary condition for topichood” (Reinhart 1981: 73). It is rather assumed that the topic presents knowledge that is already in the shared attention of the interlocutors. This was already outlined by Strawson. Nowadays, this means that the discourse referent that is already in the CG can serve as a topic. Note that this is not a sufficient condition, since not everything in CG is a topic. For a more fine-grained analysis about what it means to be activated in the shared discourse knowledge of interlocutors, see the features by Prince (1979), presented in Reinhart (1981: 61). In the first brief analysis in chapter 1.2., the first NP

was identified as a topic. The sentence is given here again in short. From context, we know that the Little Prince is travelling from planet to planet and meets a couple of different people. In sentence (4.2-6), reference is made to the business man, who was already introduced to the common ground and therefore presents given information.

4.2-6 [Komerčand]_{TOP} [je dvignija glavu]_{COM}: [...].

businessman.NOM.SG.M be.PRS.3SG lift.PFV-PTCP.SG.M head-ACC.SG.F

‘The businessman raised his head.’

(Saint-Exupéry 2009: 45)

Criterion 2b: The principle of relevance (Strawson 1964)

Closely related to the notion of shared knowledge in the CG is a second point mentioned by Strawson “that an expression will be understood as representing the topic if the assertion is understood as intending to expand our knowledge of this topic [...]” (Reinhart 1981: 59). This point might appear redundant but is just as crucial as criterion 2b. Several things are activated in the CG at the same time. However, only the entity that we want to talk about more, that we want to expand the knowledge of a hearer about this entity (topic) can serve as a topic (unless we have a topic change). Everything that was just said as a comment on something else is not a topic, but rather a focus/comment/new information. In sentence 4.2-5a, the speaker (or narrator) wants to provide more information about the asteroid and not about the astronomer, while in 4.2-5b it is the astronomer who the hearer (or reader) should learn more about.

Criterion 3: (What) about – test (Gundel 1989, Reinhart 1981: 64-65)

If we want to question what the NP that we identified as a topic with the first two assumptions is doing or how it is affected by the new information of the focus, we can simply ask “What about this entity?”.⁸ Connected to this aboutness assumption is the

⁸ Other tests do exist as well, but turned out to be impractical for our current analysis: A couple of tests were introduced how to identify topics in discourse. Three of the most well-known are presented in short.

As-for test (Gundel 1989)

- a) [Dorothy]_{TOP} [is bringing chicken salad.]_{COM}
- b) [As for Dorothy]_{TOP}, [she’s bringing chicken salad.]_{COM}

tendency that topics are the NP that is also subject in grammatical terms. This point is only considered briefly here, and not further elaborated (but see Reinhart 1981: 86). Hence, since it is the next planet to which he travels, we can fairly assume that *pjanet numar četar* is a topic. We could ask the question: What about the planet number four? A possible reply (in the comment) would specify that the 4th planet is the one of the businessman.

4.2-7 [Pjanet numar četar]_{TOP} [biša oni do jenga komerčanda.]_{COM}
planet.NOM.SG.M number-NOM.SG.F four be.IPRF-3SG DEM.DIST-NOM.SG.M of
ART.INDF-GEN.SG.M businessman-GEN.SG.M

'The planet number four was the one by the businessman.'

(Saint-Exupéry 2009: 45)

Criterion 4: Paraphrasing as a question (context-dependent) (cf. Reinhart 1981:56)

A closely-related approach is to paraphrase the proposition of the sentence into a question. By asking who or what is fulfilling or undergoing the action specified in the sentence, we might limit what can be considered as a topic. For example, if we want to ask in 4.4-4a who brought the wood to the shed, we would end up with the boy as the topic. We might ask in reverse what the boy did to determine the comment. Here we encounter a problem in our analysis: depending on the question (context-dependent), the boundary of topic and focus regarding the verb shifts (see also Reinhart 1981: 58 on this context dependency). Hence, we have to add one assumption. In our example, we have no further information about the situation in which this utterance was given. If we had information on how much the interlocutors already knew about wood that was carried to the shed, we could more easily identify the

What-about test (Gundel 1989)

c) What about Dorothy? > [She]_{TOP}['s bringing chicken salad.]_{COM}

Say-about test (Reinhart 1981)

d) [Roger said that]_{NAI} [Dorothy]_{TOP} [is bringing chicken salad.]_{COM}

Without access to speakers of the language, we cannot really apply the As-for test and the say-about test without checking if such constructions exist in the language. The What-about test is possible.

appropriate interpretation. In example 4.2-5 from Molise Slavic, the first mention of the person who died is the focus because he provided new information (the topic is rather virtual, because we already know that the speaker was bought a tractor by somebody). If we paraphrased the 2nd sentence, we could ask what about the guy who sold the tractor and receive the answer that he died as a focus. If we paraphrased the focus now as “Who died?”, we receive the “guy who sold the tractor” as the topic. In another discourse, when the speaker mentions that somebody died, the question for the new information (so the focus) would be “Who died?”. Here, we can see the dependency on context in the interpretation of focus and topic. The alternative view is that *oni* is a topic of a newly-introduced referent that it specified later (see Reinhart 1981: 77 on this point).

(Additional) criterion 5: Intonation

In contrast to Reinhart (1981), we also apply intonation as a testing tool. She acknowledges this possibility but does not use it (see Reinhart 1981: 85). We assume that focus is heavily stressed in contrast to topic, which is rather neutrally stressed. This point is elaborated in more detail in chapter 4.2.4.

These tools combined still may fail to identify the topic of a discourse, but it should provide us with a practical tool to account for the focus. We assume that most foci occur in the comment, except for the special situation of focused topics (see chapter 4.2.4 for more on intonation). The tools applied in this chapter for a consistent analysis of topic were applied to two pieces of discourse in Molise Slavic. The first text is an excerpt from the translation of the little prince. The advantage here is that we can easily compare the Molise Slavic version to the original text (French as an article language) and BCS, genetically the closest Slavic language. The second text is a story by a Molise Slavic speaker in Breu (2017) about the Russian tractor that he bought and some issues related to this, and a type of fairytale/fable about a crayfish and a hare who are betting on who reaches first the top of a hill. The second text was chosen especially for an additional intonation analysis because this recording is publicly available. The whole version of analysis is given in the appendix. Not every step is explained in detail, but the testing kit as presented here was applied for every

identification of topic and focus. In cases where the analysis is not absolutely unambiguous, notes are given. In the remaining chapters, selected sentences and pieces of discourse from those two texts are discussed in detail.

4.2.3. Sentence-initial position

It was mentioned earlier that in the case of Slavic languages (except Bulgarian and Macedonian) the sentence-initial position is often seen as triggering a definite interpretation of the entity that takes this position. We have seen examples of this effect in Polish (4.2-1 and 4.2-2), Croatian (4.2-4), Czech (4.2-4) and Molise Slavic (4.2-5). In the literature, we find claims such as: “Topicalization marks definiteness and focusing marks indefiniteness. [...] In Russian IS (via word order) therefore plays a critical role in determining definiteness” (Erteschik-Shir 2014: 38). Thus far, we have seen that in some examples the sentence-initial position coincides with the topic, which again can be related to a definite interpretation of its nominal phrase. We must discuss the point of position in further detail before we turn to Molise Slavic. In the examples mentioned, we had transitive sentences with one arguments of the verb. The pre-verbal NP was the topic in all examples, regardless whether they were the subject or object of the sentence (4.2-1 versus 4.2-2 and 4.2-4a vs 4.2-4b). Those pre-verbal NPs were considered as definite.

4.2-8 [Mal’čik]_{TOP} [prišel]_{COM} (Russian)

boy.NOM.SG.M arrive.PFV-PST.SG.M

‘The boy arrived.’

[Prišel]_{TOP} [mal’čik]_{COM}

arrive.PFV-PST.SG.M boy.NOM.SG.M

‘A boy arrived.’

(Birkenmaier 1979: 63)

Besides this, in sentences containing only one NP and a verb, inversions are also possible affecting the definite interpretation. When the verb takes the sentence-initial position, we have the special effect that the topic is verbal. This strategy is well known in the literature as *narrative inversion*, an inversion to VSO order (see Jasinskaja 2016:

7). Example (4.2-10) is one such example where the verb moves to the sentence-initial position. Existential constructions of the type *there is/ was* as in 4.2-11 are not considered to belong to this category.

4.2-9a [On]_{TOP} [recenziroval stat'ju.]_{COM} (Russian)

he.NOM review.IPFV-PST.SG.M article-ACC.SG.F

'He reviewed the/ an article.'

4.2-9b [On stat'ju]_{TOP}⁹ [recenziroval]_{COM}

he.NOM article-ACC.SG.F review.IPFV-PST.SG.M

'He reviewed the article.'

(Birkenmaier 1979: 63)

4.2-10 [je reka]_{TOP} [rak]_{COM} (Molise Slavic)

be.PRS.3SG say.PFV-PTCP.SG.M hare.NOM.SG.M

'The hare said.'

4.2-11 [bihi mali soldi]_{COM}

be.IPRF.3PL little-NOM.PL money-NOM.L

'There was few money.'

(see texts in the appendix)

Another effect that is probably related to the role of sentence-initial positioning is the ordering of objects in a sentence. The order of indirect and direct objects can also affect the definiteness of the entities that they denote. In the following example, the first object is considered definite.

4.2-12a [Mužčina]_{TOP} [podaril devuške cvety.]_{COM} (Russian)

man.NOM.SG..M give.PFV-PST.SG.M girl-DAT.SG.M blossom-ACC.PL.M

'The man gave the girl flowers.'

4.2-12b [Mužčina]_{TOP} [podaril cvety devuške.]_{COM}

man.NOM.SG..M give.PFV-PST.SG.M blossom-ACC.PL.M girl-DAT.SG.M

The man gave the flowers (to) a girl.

(Birkenmaier 1979: 65)

Interestingly, in languages displaying articles, the order of objects also seems to be the same. Definite objects occur before indefinite ones. Examples from German indicate that the order of direct and indirect objects is determined by its degree of

⁹ Depending on the theory, one could assume one or two topics here. For the sake of keeping the analysis simple, only one topic per sentence is assumed here. Hence, the object of the sentence moves into the topic.

determination, hence definiteness status. The higher the determination of an object, the more likely it will appear earlier. In cases in which the degree of definiteness is the same, the indirect object will usually occur before the direct object (4.2-13a)¹⁰. If the definiteness status is not the same, the definite object will occur in front of the indefinite object (4.2-13b). Note that examples in 4.2-13c are not ungrammatical, but stylistically marked. Hence, the neutral ordering follows the a tendency to first process the already-known information and then add the new (see Birkenmaier 1979: 65; examples 4.2-13a-c are my own).

(4.2-13a) Ich gebe dem Mann das Buch./ Ich gebe einem Mann ein Buch.

(4.2-13b) Ich gebe dem Mann ein Buch./ Ich gebe das Buch dem Mann.

(4.2-13c) Ich gebe einem Mann das Buch./ Ich gebe ein Buch dem Mann.

This principle will be discussed below. In general, the order of object constituents in the comment position should be investigated in future research.

For now, I have to limit this discussion to the topic position. Birkenmaier compares examples from Russian with their German translation equivalents. He makes the important discovery that only the initial position is specified for definiteness. By contrast, the post-verbal position is neutral. This might also relate to the role of agentivity which is not accounted for in this study (but see Brunetti 2009b on Spanish and Italian for a first attempt in this direction).

34.2-14a [Direktor]_{TOP} [izučæet proekt.]_{COM}

director.NOM.SG.M study.IPFV-PRS.3SG project-ACC.SG.M

‘The director studies a/ the project.’

4.2-14b [Proekt]_{TOP} [izučæet direktor.]_{COM}

project-ACC.SG.M study.IPFV-PRS.3SG director.NOM.SG.M

‘The project is studied by a/ the director.’

4.2-14c [Direktor proekt]_{TOP} [izučæet.]_{COM}

director.NOM.SG.M project-ACC.SG.M study.IPFV-PRS.3SG

‘The director studies the project.’

(Birkenmaier 1979: 48-49)

¹⁰ In contrast to German, but fully in line with other South Slavic languages the neutral order of objects in Molise Slavic is indirect (dative) object > direct (accusative) object as can be seen in the ordering of pronominal clitics (see Breu, Piccoli 2000: 405).

The first referent in every sentence is definite according to Birkenmaier's analysis, but the second referent can be both definite and indefinite, specific or unspecific. The first position in the sentence is reserved for definite referents. Subject and object NPs can both occur in the first position:

Das Auffallendste an der Übersetzung – von unserer Problematik aus gesehen – ist die Möglichkeit von zwei Artikelformen in bestimmten Fällen gegenüber dem alleinigen Vorkommen des bestimmten Artikels in anderen Fällen. Die Schwankungen kommen nur in der Endposition vor. In Erst-Position erscheint nur der bestimmte Artikel. (Birkenmaier 1979: 49)

In cases where both translations are accepted, a decision whether indefinite or definite reading is to be chosen can only be determined by context. Birkenmaier already relates the sentence position with the concept of topic. In his view, the first position in the sentence functions explicitly as a topic position:

Das erlaubt den Schluss, daß die Themaposition in Bezug auf Bestimmtheit markiert ist, während die Rhemaposition in dieser Hinsicht unmarkiert ist (Birkenmaier 1979: 49).

I do not want to discuss here how indefinite referents should be interpreted when they appear in the sentence-initial position in other Slavic languages. I will return to this issue below regarding the Molise Slavic data. In our small sample of texts analysed in terms of information structure, we find many instances that seem to confirm the hypothesis that just as in other Slavic languages there is a tendency for referents that should be interpreted as definite to appear in the sentence-initial positions.

4.2-15 [Čeljade velke]_{TOP} [jesu činjene naka.]_{COM}

People-NOM.PL big-NOM.PL be.PRS-3PL make.PFV-PTCP.PASS.PL so.DIST

‘Grown-ups are like that.

(Saint-Exupéry 2009: 19)

4.2-16 [Aštronom]_{TOP} [je čija jopa njevogu dimoštracijunulu 1920 s jenme višitam čuda elegand.]_{COM}

Astronomer.NOM.SG.M be.PRS.3SG make.PFV-PTCP.SG.M again his-ACC.SG.F demonstration-ACC.SG.F 1920 in ART.INDF-INS.SG.F clothing-INS.SG.F much elegant-INS.SG.F

‘The astronomer gave his demonstration again in 1920 dressed in a more elegant costum.’

(Saint-Exupéry 2009: 18)

Not only bare nouns can be found in the initial position, but also instances of demonstratives (4.2-17) and pronouns, even the clitic forms (4.2-18).

4.2-17 [Ovo]_{TOP} [ma ne moraša čit zbaučit čuda.]_{COM}
 DEM.PROX-NOM.SG.N I.ACC NEG can-IPRF.3SG make-INF surprise.PFV-INF much
 ‘But that did not really surprise me much.’

(Saint-Exupéry 2009: 18)

4.2-18 [Ga]_{TOP} [zova presembj : << ašterojid 325>>.]_{COM}
 He.ACC call.IPFV-PRS.3SG for example asteroid.NOM.SG.M 325
 ‘He calls it for example: asteroid 325.’

(Saint-Exupéry 2009: 18)

At this point, one has also to take into consideration the ordering with subordinate sentences.

4.2-19 [Kada jena aštronom skopriva]_{COM} [jena do’vihi]_{TOP}, [mu]_{TOP} [daja kana jiman jena numar.]_{COM}
 when ART.INDF-NOM.SG.M astronomer.NOM.SG.M
 ‘When an astronomer discovers one of them, he gives it a number instead of a name.’

(Saint-Exupéry 2009: 18)

Italian exhibits two different topic positions: topics occurring on the left side of a sentence are prototypically-shifted topics, while continued topics are attached to the right (see Brunetti 2009a, quoted by Erteschik-Shir 2014: 37). The Molise Slavic data does not support the hypothesis that its information structure simply replicates the information structure of Italian in this regard. In example (4.2-20a), a continued topic is located on the left side, whereas in example (4.2-20b) a shifted topic is found on the right side if we analyse the right constituent as topic (instead of considering it an instance of the narrative inversion).

4.2-20a [Zec]_{TOP} [je reka.]_{COM}
 rabbit.NOM.SG be.PRS.3SG say.PFV-PTCP.SG.M REL
 ‘The hare said [...]’

(Breu 2017: 91)

4.2-20b [je reka]_{COM} [rak]_{TOP}
 be.PRS.3SG say.PFV-PTCP.SG.M REL crayfish.NOM.SG.M
 ‘The crayfish said [...]’

(Breu 2017 : 90)

Example 4.2-20 could also be reinterpreted in the sense of a narrative inversion as described in the previous chapter. Thus far, only single examples have been given for instances of referents in the sentence-initial position that should be considered as

being definite, whereas we have not discussed instances of indefinite referents in this position. In example (4.2-21), we have an example:

- 4.2-21 [Nu votu]_{NAI} [na zec]_{TOP} [prohodaša jizbane jene rike.]_{COM}
 ART.INDF-ACC.SG.F time-ACC.SG.F ART.INDF-NOM.SG.M rabbit.NOM.SG.M pass.IPFV-
 IPRF-3SG on.side ART.INDF-ACC.SG.F river-GEN.SG.F
 ‘Once a hare was walking along a river.’
 (Breu 2017: 89)

Counter-examples like this are sometimes taken as an argument against the connection of definiteness and topicality. Trenkic analyses the claim that topic is necessarily definite on Croatian material. Moreover, in the Acquaviva part of Breu’s corpus that is made public, we find only one instance in which the indefinite article introduces the sentence (4.2-22). In the sentence following this part, direct reference to the father is made with a demonstrative.

- 4.2-22 [Na tata]_{TOP?} [jimaša dva sina.]_{COM}
 ART.INDF-NOM.SG.M father-NOM.SG.M have-IPRF-3SG two.M son-GEN.SG.M
 A father had two sons.
 (Breu 2017: 194)

We can conclude that it is possible for indefinite referents to appear in this position. However, single instances of definite or indefinite referents in topic position cannot confirm that there is a stronger tendency for definite categories to be related to topichood and therefore that they are found in what was identified as a topic position.

For this purpose, we need quantitative figures of occurrences in a corpus. It exceeds the scope of this study to provide a fully-detailed analysis of a large number of sentences controlled for their sentence-initial position. Therefore, only an indirect measure is given here. Thus far, we have assumed that indefiniteness must always be marked in the respective NP. Therefore, as Breu assumed, every unmarked, bare noun should be considered as definite. The main assumption for this short corpus exploration is that if the number of instances of indefinite referents in the sentence-initial position is low, this provides further confirmation for our claim. By contrast, if clearly-definite categories like demonstratives and pronouns appear more frequently, this is also a good hint for the claim. The weakness of this approach is that bare nouns that are actually the subject of this discussion are largely ignored. The following

numbers should therefore only be taken as an indirect measure for the sake of economy. For this purpose, the corpus as described in chapter 1.3. was refined for alphabetical, sentence-based analysis. With the open-source tool notepad ++, all sentences were brought into an own line semi-automatically. A sentence was definite as every unit of words ending with a full stop, question mark or exclamation mark. All other symbols like quotation marks were deleted with the replace function. With this approach, relative clauses and clauses connected by conjunctions are not taken into consideration. Nonetheless, this approach provides a quick overview and can be used as a first hint for further corpus-based research. Everything was brought into alphabetical order. This led to a list of sentences with about 4,523 lines (=sentences).

The resulting list provides us with a rough, simple tool for the exploration of the sentence-initial position. It turned out that about 66 lines (=sentences as defined above) start with a form of the indefinite article (numerals excluded). Upon deeper exploration of these 66 cases, it emerged that 35 actually are temporal adverbials/phrases (of the type *nu jistru* ‘one morning’, *na dan* ‘one day/ once’.). Overall, only 31 sentences (0.69% from all sentences) start with an indefinite noun phrase, since the indefinite article is fully grammaticised and claimed to be obligatory. Upon closer examination, further instances needed to be removed because they either comprised sentences without a verb or were instances of the numeral (4.2-23, 4.2-24), the preposition (*na štacijunu* ‘at the station’) or fixed expressions as *na mala* ‘a bit’. Considering this reduction, only 20 cases (0.44%) could be counted as truly being examples of the use of the indefinite article. The following examples exemplify this type of occurrence:

4.2-23 [Na žena]_{TOP} [jinveč ju bolaša nog(a).]_{COM}

ART.INDF-NOM.SG.F woman-NOM.SG.F instead she-ACC hurt.IPFV-IPRF.3SG leg-NOM.SG.F

‘Her legs hurted one woman.’

(Breu 2017: [3.5.2]-83)

4.2-24 [Nu avenduru]_{TOP} [ka neč’ ju zabit maj!]_{COM}

ART.INDF-ACC.SG.F adventure-ACC.SG.F REL NEG-want.PRS.1SG she-ACC forget.PFV-INF never

‘An adventure which I will never forget.’

(Breu 2011: [3.5.4]-14)

4.2-25 [Jena]_{TOP} [sa zovaša Romadori Alfredo [...]]._{COM}
 one-NOM.SG.M REF1 call.IPFV-IPRF.3SG Romadori Alfredo
 'One was called Romadori Alfredo.'

(Breu 2011: [1.2.2]-50)

Counter-examples like this are sometimes taken as an argument against the connection of definiteness and topicality. All examples in 4.2-23 to 4.2-25 (also counting the numeral) are specific in the context of the utterance. Therefore, I hypothesise that unspecific referents are not allowed in the topic position. Trenkic takes comparable instances in Croatian as a criterion that topic position is not related to definiteness (see examples in Trenkic 2004: 1403). For this study, a full in-depth account of all sentences would exceed the scope of this study. Therefore, not all instances of full-noun phrases in the sentence-initial position can be studied here. In order to compare only all those NPs starting with a demonstrative, clearly overtly marked definite NPs are counted. 383 sentences (8,47%) start with a NP with a demonstrative (ambiguous forms *ona*, one only counted when attributively attached to the noun to avoid erroneous counts of pronominal forms, all other instances of these two forms are counted as pronouns; since they are added up in the final result of this exploration this error is healed). Sentences starting with a pronoun were even more frequent: 9.79% of our test sentences started with a pronoun (forms of *ma* were not counted). This means that that about 18.26% of all sentences started with a category for which we are absolutely sure of its demonstrative nature. This already gives a hint that the sentence-initial position is reserved for NPs carrying given, old information, because it seems that definitely-marked phrases occur more often in this position.

In order to exclude the possibility that this is only due to general discrepancy in the distribution of indefinite articles and demonstratives, the same sub-corpus was explored by using the word count function in AntCon. The 4,523 sentences comprise 44,881 words (tokens). In the word list, it is not possible to discriminate instances of *ona* as pronoun or demonstrative. Therefore, this form is excluded completely in the following estimate. This means that demonstratives are slightly under-valued/-ranked. With this limitation, 1,619 instances are counted as demonstratives (3.61%). Personal pronouns occur at a rate of 6.92% in the corpus. Therefore, taken together, demonstratives and pronouns can be found in around 10.53% of all tokens in the

corpus, while in the sentence-initial position they occur even twice as frequently. A similar problem exists in counting instances of the indefinite article, given that *na* can function as a preposition (presumably frequent) and as a short form of the NP. Since both occur in front of nouns and an assessment of all instances would be too time-consuming, rough estimates are given. The minimum number of instances (632, 1.41%) counts only 265 instances of 658 of *na* as cases of the article (because that is actually the same distribution as in the sentence-initial position in comparison to instances of *nu*, which occurs about 0.75% as frequently as *na*). As a maximum, all instances of *na* are counted, which is empirically wrong, but serves as a check for maximum (1,025 instances, 2.28%). In any case, it becomes clear that overall the demonstratives and pronouns occur about (maximally) 4-5 times as frequently as the indefinite article (regardless of which count). However, in the sentence-initial position, demonstratives occur more than 26 times more often than the indefinite article. As additional support for this admittedly raw calculation that only provides a first (incomplete) evaluation, occurrences in inversion cases are counted. In cases where the two existential or presentational verb forms (*biša* 'there was', *bihu* 'there were') are in the sentence-initial position, indefinite NPs (with an indefinite article) occur in 29% of all instances, while clearly-definite cases as defined before only about 4.34% of the sentences. As mentioned before, I assume that this type of sentence does not have a topic.¹¹ Therefore, they are instances of sentences which are completely part of the comment. In this position, indefinites are far more frequent than in the topic position. This short quantitative examination supported the findings. A more fine-grained analysis of all instances and particularly the comment position must be conducted in future research.

It was discussed before how topics relate to definiteness. There is a pragmatic relation considering what topics mark:

Given that topics almost invariably represent given information, it is to be expected that topic noun phrases will frequently be definite (Lyons 1999: 232).

¹¹ Lyons explains this with the newness of the referent: "Certain presentational verbs, which can be used to express the introduction of a new referent into the discourse, particularly favour a non-topic interpretation of their subjects [...] Particularly in the case of sentences beginning a discourse or section of discourse, there need not be a point of departure expressed." (Lyons 1999: 229).

However, it is also clear that specific (and probably also generic) indefinite NPs can also be used as a topic (cf. Reinhart 1981: 66), with the only valid criterion in this direction being pragmatic referentiality. The question arises what underlying principles lead to a definite interpretation of nominal entities in the sentence-initial position. In example (4.2-6) for instance, the referent *komerčand* ‘businessman’ was clearly the topic in the sentence “[Komerčand]_{TOP} [je dvignija glavu]_{COM}.” Based on context and position, it is fair to say that the businessman is definite here. The question arises whether this is achieved mainly by the topic position or simply as a side-effect of differing reasons. The sentence-initial position is related to topics for a cognitive reason. Except for the case that a new referent is introduced sentence-initially, it is prototypically the position where definite referents are to be expected in Molise Slavic. This follows the proposed *given – new Principle* in the literature:

Given information tends to appear closer to the beginning of a sentence, while new information tends to appear closer to the end of a sentence (see Halliday 1967 [...]). (Birner 2013: 210).

In some accounts of information structure, it was even assumed that the dichotomy of new versus given information alone is responsible for the assignment of definiteness status. This ordering also derives from the role that topics play in discourse:

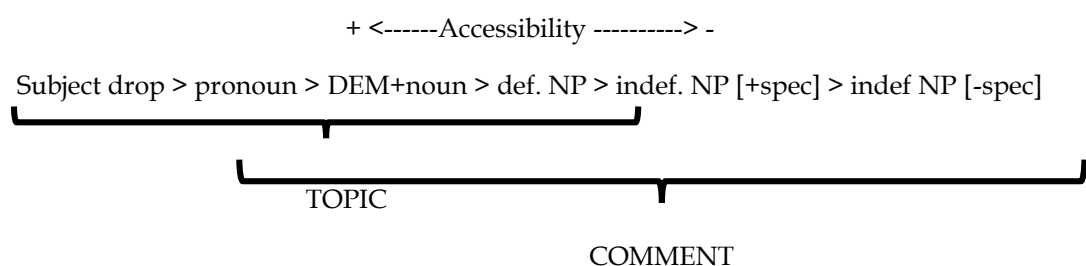
The topic expression foregrounds something already in the consciousness of the participants in order to make it the point of departure for some new information” (Lyons 1999: 228).

Trenkic comes to the conclusion that definiteness is not related to information structure “in a straightforward way” (Trenkic 2004: 1404). She identifies two possible ways in which given and new can be understood. If it is understood as a category of the consciousness, “given information is ‘activated’ information, believed by the speaker to be currently ‘lit up’ in the consciousness of the hearer” (Trenkic 2004: 1404). New information in this category can be either accessible or inaccessible to the hearer. In her account, accessibility equals familiarity of older accounts. The alternative view is to see the dichotomy as a category of knowledge. Given is what is identifiable to the hearer, because it is either activated or accessible to him. New information is always inaccessible then at the point of first mention. In this view, the first category relates to the concept of topichood, and the latter to the category of definiteness. This is illustrated in table (5), which is taken from Trenkic (2004: 1404):

Category of consciousness:	GIVEN (=theme)	NEW (=rheme)	
Referent status:	activated	accessible	inaccessible
Category of knowledge:	GIVEN (=definite)		NEW (=indefinite)

Table 5: Types of givenness categories (based on Trenkic 2004: 1404)

According to this solution, a referent can only be the theme (or topic) when it is activated (in the common ground in our view). It seems that there is a hierarchy regarding how this activation maps on grammatical form in the sentence. Assuming two distinguishable but related levels provides us with a tool to describe which types of formal marking are to be expected in positions specified for topic. The higher that something is activated or accessible, the more likely it will occur in the topic. To be precise terminology-wise, this is a question of information status and degrees of explicitness rather than information structure. Taking the category of knowledge into consideration helps us to define what types of nouns we can expect to occur in the topic position. A referent that is fully accessible and activated can be dropped completely in an utterance. Pronouns can only be used when the reference is unambiguous. Based on my investigation of Molise Slavic, I present a first attempt of an accessibility hierarchy which tries to predict how referents are morphosyntactically marked depending on their accessibility in regard of topic. The higher accessible a referent is, the more likely it will be marked with one of the constructions to the right. Topic and comment is hereby also dependent on accessibility.



The hierarchy captures the fact that only definite and in few contexts specific referents can occur in topic position. In comment almost every construction can occur.

Trenkic's argumentation is fully in line with the data presented here. It explains why definites may occur in either the topic or comment position, while indefinites are much more likely expected in the comment of an utterance. Therefore, definiteness in the broad sense and topic – comment structure are related but not identical. Birkenmaier comes to the same conclusion in his analysis:

Die nominale Determination ist also als ein Nebenergebnis der aktuellen Satzgliederung zu verstehen, das sich unter ganz bestimmten Bedingungen – und nach Einzelsprachen verschieden – einstellt“ (Birkenmaier 1979: 50).

It turned out in this study that comment or rheme must be investigated in following research. This could shed more light on the question of how definites are marked in a sentence. Thus far, we have no testing tool to account for topics that are not in the propose topic position. A last possibility remains concerning how definites could be made identifiable by means of information structure. If definites outside of the topic position tend to be marked by either intonation or morphosyntactic topicalization markers, this could solve the question of how definites might be marked in a sentence. We do not assume that every form carrying a definite meaning is actually marked by one of these strategies. It is well known that Slavic languages may sometimes be vague when it comes to explicitly marking definiteness, whereby in some cases ambiguities remain without necessarily affecting communication in a negative way. This last claim is tested rather briefly in the next two chapters.

4.2.4. The role of intonation

In some languages, a topic is marked by intonation and therefore this option must also be taken into consideration. There is a tendency whereby constituents related to newness and focus are “encoded with a wider pitch range, a longer duration, and a higher intensity, compared to givenness and categories like topic” (Chen 2012: 251). At present, no systematic account of intonation in Molise Slavic exists (see Breu 2017: 20). This complicates the task because it is unclear how to deal with prosodic effects in Molise Slavic. Therefore, the rough distinction claimed by Chen (2012) is assumed here and only pitch is considered as a measure. It was necessary to develop a tool to describe the main features of intonation in Molise Slavic. The best framework to date is the family of ToBI annotations developed in the area of autosegmental-metrical

phonology (for a general overview of ToBI annotation, see Grice 2006; for its application on BCS, see Godjevac 2005). Despite the fact that Italian largely influenced Molise Slavic in the domain of intonation, I follow the framework developed in Godevaj (2005) for the application of the ToBI transcription rules for Serbo-Croatian. This has the advantage that it accounts for lexical pitch accent, which is still productively used in MS. I mainly follow the conventions used as in the classical ToBI analyses with the modifications added by Godjevac. All examples are analysed in *Praat*, whereby the pitch ranges were adjusted to the gender of the speaker. For the sake of readability, the lexical pitch accent is given on the first tier and the phrase level intonation is annotated on the second tier. Length is no longer a feature of pitch accent in MS but rather a question of vowel quantity. Therefore, only two pitch accents are distinguished in my analysis: the falling tone typically has its highest pitch on the stressed syllable followed by a decrease in pitch (H*+L), while the rising tone has a reverse pitch contour (L*+H). Additionally, the highest pitch in an intonational unit is marked on the second tier (HiF0). Possible downsteps are marked by an exclamation mark (!). Upsteps – which seem to occur occasionally – are marked by #, which is not used in the classical ToBI transcriptions. Global downsteps that occur at the phrasal level are indicated by the symbol Ø, when the whole pitch of the utterance decreases after a focus (a modification borrowed by Godjevac 2005/2010). This effect can be found in BCS and clearly also exists in MS. At the phrasal level, the boundary tone of words (either %H or %L) and the boundary tone of the intonational phrase (H% or L%) are marked. Additional tiers are provided for the break indices, glosses into English and full translation of the text. Although the intonation underwent a clear impact from Italian, I stick to the transcription framework for BCS due to its simplicity. When describing the intonation of Molise Slavic, one encounters the problem that three different sources trigger the shape of intonation (not to mention side-effects due to phonetic articulation, e.g. pitch raising or lowering by voiceless obstruents). At the smallest unit, the lexical pitch accent shapes the form of the word itself when carrying the stress. Second, the actual type of sentence (declarative, imperative, wh-question, intonation question) affects the phrase level intonation as a whole. The third presumed force is information structure. Due to this interplay of three different

processes, it is difficult to identify actual effects of information structure. Therefore, I limit my discussion to only one type of sentence, namely the declarative statements. In addition, lexical pitch accents are marked on the second tier as they are expected based on their word form and not how they are actually realised in the given piece of speech. For our analysis, a fine comparison of the phrasal-level intonation is the only analysis that can be conducted based on the material used here. In praat picture 3 in the appendix, the typical intonation of a declarative sentence is illustrated (compare also picture 9). Despite some special features, we can clearly see the gradual pitch decrease from %H within the sentence, terminating in L%. The constitute *vaza* ‘took’ is emphasised with a heavier pitch. This indicates that it is the focus of the sentence, because in languages that mark focus with an intonation it is usually the most prominent pitch realisation in an utterance (cf. Grice 2006): “[...] prosody is one of the key cues to focus structure in many spoken languages” (Cowles 2012: 293). To test the proposal that focus is marked with heavy pitch in MS, it is the best to test the answer to a *swh*-question. In the fairytale *Little red riding hood*, the grand-daughter asks the wolf (who is hiding in her grandmother’s bed) why she has such big ears. In the response in the version recorded in Acquaviva Collecroce, the adverbial *bolje* ‘better’ in the response is clearly emphasised by stress (see praat picture 4 in the appendix). Other examples of heavily stressed foci can be seen in praat picture 10, 11, 12 and 15. A more detailed study on prosody and intonation is necessary, not only regarding information structure but also in relation to the tonal system and the influence of Italian intonation, which can be clearly heard. For now, we assume focus as being the most prominent pitch in the sentence. Based on this, we can investigate whether intonation also affects topics. Since it is not the purpose of the study, I only give a few illustrations of the intonational effects on topics, excluding all other effects. As in other languages, it is also possible to focus a topic. The praat picture 5 gives an example of this. Note that the palatal approximant somehow distorts the pitch signal. This was also observed in other examples. In this example, a secondary focus might also exist. The same process of bringing a focus into sentence-initial position – in other words, focussing the topic – can be found in Russian. Quoting examples from Bel’skij (1956), Birkenmaier gives the following example to account for this. 4.2-27 a and b present the

intonational neutral possibility. The topic structure is achieved by word order and determines the interpretation as (in)definite. Since no other constituents or intonational cues are present in the intonationally-neutral utterance, the comment equals the focus.

4.2-26a [Mal'čik]_{TOP} [[prišel]_{FOC}]_{COM}

4.2-26b [Prišel]_{TOP} [[mal'čik]_{FOC}]_{COM}

If we take intonation into account, we have four instead of two possibilities, each with a different reading that they trigger. In the constituent order of 4.2-26, we get the same reading as in 4.2-28a when the second constituent is stressed. A change in information structure will occur when the first word is stressed (4.2-27b) By intonation, mal'čik receives the focus.

4.2-27a [Mal'čik]_{TOP} [prišel]_{COM}

4.2-27b [[Mal'čik]_{FOC}]_{TOP} [prišel]_{COM}

The same happens when word order is reversed. Nonetheless, intonation can overtake the function of assigning focus status to a constituent in the sentence (see Birkenmaier 1979: 56).

4.2-28a Prišel mal'čik > mal'čik [-/+spec], [-def]

4.2-28b Prišel mal'čik > mal'čik [+spec], [+def]

One interesting test case regarding information structure is word order alternations and the accompanying prosodic effects compared to the regular intonation as just described in the two examples. One possible word order change discussed in the literature is the beforementioned narrative inversedion (see chapter 4.2.3 and praat picture 6). This word order is restricted to broad focus thetic sentences, putting the verb into the topic indicating the topicality of the situation time (see Jasinskaja 2016: 7 with reference to Junghanns and Zybatow 2009: 697-698). Based on what was said, about 30 sentences (mostly equalling intonational phrases) were analysed, whereby 15 are given with full annotation in the appendix. The following table summarises the intonational effects that were found. In conclusion, it emerged that topic does not seem to have a large effect on intonation in Molise Slavic. The information structural

category affecting pitch range is focus. All material occurring post-focally is reduced in pitch. This effect can also be found on right dislocated topics (praat picture 13 and 14 in the appendix). When focus coincides with topic, this constituent is heavily stressed due to the focus intonation. It seems that shifted topics are accompanied by a slight variation in intonation. In the data presented here, it seems that shifted topics are marked by a *bridge* intonation LH* (plus H*+L in graphic 7 in the appendix) due to the pitch accent) (or *hat contour* ‘Hutkontur’ in German terminology). A bridge intonation is common in Slavic languages for shifted topics (see Junghanns, Zybatow 2009: 693). This becomes visual even more in listings of contrastive pairs (e.g. in a headline). In praat picture 8 in the appendix, the headline of the fairy tale about the hare and the crayfish is analysed. Here, the bridge is clearly visible and indicates the contrasting of two topics, because in this utterance there is no indication which of the two referents will serve as the topic in the following. Besides this bridge intonation and focused topics, no other pattern was found in the data under investigation. In conclusion, intonation does not seem to play a huge role regarding topic. This is in line with Birkenmaier’s brief analysis on Russian. Topic and comment are largely determined by word order and focus, and perhaps other categories are intonational in contrast. It has become apparent in this study that besides sentence-initial position, some other strategies might be applied for marking topics in the sentence. Accordingly, these strategies are briefly summarised in the following chapter.

4.2.5. Topicalization strategies

Besides word order and intonation, some languages make use of special constructions to mark topics. Japanese and Chinese deploy specific topic markers, while English and French make use of cleft-constructions, leading to the fronting of a NP and therefore assigning it topicalized status. Likely candidates for topicalization strategies can be found in Molise Slavic. Molise Slavic makes use of left dislocation as well as right dislocation. Left dislocation is often seen as a topicalization strategy that “[...]is used to change the current topic” (Reinhart 1981: 64). The following utterance was indeed made in contrast to another topic:

4.2-29 Tvo(ja) mat, ta [...]!

your.NOM.SG.F mother.NOM.SG.F DEM.MED-NOM.SG.F

‘Your mother, she [...]!’

(Breu 2011 : [1.3.8]-31)

Right dislocation also exists. Just like in Italian, RD may occur with a clitic in the main clause and without. Compare the Italian examples in 4.2-30a and 4.2-31a with the Molise Slavic examples 4.2-30b and 4.2-31b:

4.2-30a Il capo **li** odia, **i broccoli**.

ART.DEF.SG.M boss-SG.M they hate.PRS.3SG ART.DEF.PL broccoli

‘The boss hates broccoli’

(Brunetti 2009a: 759)

4.2-30b [Sa]_{COM} [ga]_{TOP} [[tija pita nu stvaru extra]_{FOC}]_{COM}, [ovga ndžinjira]_{TOP}.

be.PRS.1SG he.ACC want.IPFV-PTCP.SG.M ART.INDF-ACC.SG.M thing-ACC.SG.F extra
DEM.PROX-ACC.SG.M engineer-ACC.SG.M

‘I wanted to ask him a specific question, this engineer.’

4.2-31a Le verdure proprio non le vuole, **il capo**.

ART.DEF.PL.F vegetable.PL.F really NEG he.ACC want.PRS.3SG ART.DEF.SG.M boss-SG.M
the vegetables really not themcl he-wants the boss

‘He really doesn’t want vegetables, the boss’

(Brunetti 2009a: 759)

4.2-31b (ø) [Jimaša [čuda robi]_{FOC}]_{COM}, [ovi tata]_{TOP}.

Have-IPRF-3SG much good-GEN.SG.M DEM.PROX-NOM.SG.M father-NOM.SG.M

‘This father had many goods.’

(Breu 2017: 194)

These are first examples of these two topicalization strategies in Molise Slavic. In the literature, the claim is made that there is a restriction to LD that they do not comply with specific and generic indefinite topics (see Reinhart 1981: 64). This is essentially clear in cases where the topic NP is doubled with a clitic due to the definiteness of the article. Given the small number of occurrences, this claim cannot be tested here. Clitic doubling as a last strategy must be considered as necessarily imposing definiteness because of the pronominal component: “[w]eak (unstressed) pronouns are therefore by definition topics [...]” (Erteschik-Shir 2014: 24). This indicates that the clitic is responsible for licensing the topichood. Breu makes clear in his analysis that clitic doubling is not obligatory in Molise Slavic. Hence, this raises the question of what determines the doubling of an object with a clitic. If we consider CD in Bulgarian, topic

appears as one possible explanation for the doubling. A corpus analysis is not sufficient for determining this underlying factors. Experimental research is necessary in order to determine in more detail how topicalization strategies and definiteness are related and in which contexts they occur. Another syntactic process that is often seen as topicalising is passivisation (see e.g. Reinhart 1981: 64). However, due to the limitation of not including the verbal system in this study, I cannot account for this here. Based on our findings, there is a strong correlation between the definite status of a referent and the use of topicalization strategies. However, with the limited data set it is impossible to determine when these strategies are applied.

5. Conclusion

Molise Slavic has a fully grammaticalized indefinite article. This was confirmed in the analysis in chapter 3.1. However, in comparison to other Slavic languages, especially Croatian, it turned out that there is rather a gradual difference in functions which are covered by the indefinite article. In contrast, it was not quite clear in previous research how referents are marked for being definite. Therefore, Breu claims the existence of a definite article in form of a zero article. This view is challenged by this study. Due to the fact that other Slavic languages also do have some sort of indefiniteness marker, this study was guided by the assumption that other factors may help explain when a referent is marked for definiteness. Hypothesis 1 stated that definiteness of a referent can always be explained by either the presence of modifiers in the NP or by outcomes of information structural processes. This hypothesis was neither fully confirmed nor refuted with the present investigation. This study made a first attempt in seeking more detailed explanations for these processes. Future research must be carried out to test some aspects which were mentioned in this investigation. In the first part of the study, modifiers which can impose a definite interpretation were investigated. The hypothesis 2 was confirmed partly. Some of the modifiers under discussion express definiteness along their core meaning. Demonstratives (2.1), personal pronouns (2.2) and to some extent numerals (2.4) carry this meaning. In particular, the numeral all shares some syntactic features with demonstratives. Possessives (2.3) may be related to definiteness but this process is dependent on other contextual factors. A first

attempt was carried out for classifying Molise Slavic as a determiner-genitive languages despite the lack of a grammaticalized definite article. Other categories which were presumed to trigger definiteness turned out to be not relevant in this regard. In contrast to the distinction known from earlier stages of (South) Slavic languages, the long forms of adjectives (2.5) are not marked for being definite. The case distinction between genitive and accusative case (2.6) also is not triggering any definite interpretation. This is probably due to the development of a third construction for the expression of possession (do-construction). Modifiers alone cannot explain every instance of definite referents. Alternative explanations must be found for nominal phrases without modifiers (bare nouns). I attempted to explain this by the interplay of information structure and definiteness (chapter 4). It was hypothesized that topicality may be responsible for the definite status of a referent (hypothesis 3). This claim was tested in regard of topic position (hypothesis 3; chapter 4.2.3), intonation (hypothesis 3b; chapter 4.2.4) and other topicalization strategies (hypothesis 3c; chapter 4.2.5). This hypothesis was partly confirmed but cannot explain all instances of definite referents either. I proved that the sentence-initial position is reserved for topics. There is a strong tendency that this position is occupied by referents which can be considered definite. This was tested qualitatively with selected examples. In addition, a short quantitative analysis delivered support for this claim. Indefinite referents almost never occur in this position. Therefore, topics and indefinite referents excluded each other in most cases. However, it turned out that the relation of topic and definiteness is not direct, but rather a side-effect of its functions. Taking up ideas from Trenkic, this was explained with assuming two different categories of givenness to which topic and definiteness related. For further investigation, a first annotation tool based on the ToBI family of intonational analyses was developed for Molise Slavic. This tool can be used for further investigation of intonation in Molise Slavic, a task which still has to be undertaken. Topic is not marked intonationally, but focus is clearly achieved by pitch variation. Hypothesis 3b needed to be refuted. However, this part of the study led to a first investigation of Molise Slavic intonation. A first, short account of topicalization strategies was given. All of them turned out to be restricted to definite referents as well. This again

confirmed that there is connection between definites and topics. The hypothesis 3 was right in the sense that a correlation was found, but needs further investigation and experimental control. Intonation is not related topicalization but to information structure on other levels. This study also provided a first investigation of information structure in Molise Slavic. So far, information structural analyses of underresearched languages are still rare. This study contributes to this field as well. To conclude the results, Breu's claim was neither confirmed nor refuted. So far, the partial confirmation of hypothesis 1 is not incompatible with his claim. My approach rather focussed on pragmatic factors in the utterances leading to the definite interpretation of a referent whereas Breu's view, in contrast, deals with the morphosyntactic marking of definite referents when no modifier is used. However, this study gives a first hint that information structure may account for most instances. Future research should focus on the occurrence of bare nouns and indefinite NPs in comment position. As we have seen in this study, the comment structure is not restricted to either definite or indefinite referents. More fine-grained analyses should investigate under which circumstances definite NPs occur in the comment position. It became obvious due to this study that other areas also need to be investigated further in order to find a final solution to the question of definiteness in Molise Slavic: Topicalization strategies were only mentioned in this study but need to be researched further in regard of their function. Since they do not occur frequently in the corpus, they should be tested in elicitations with the speakers. Future research should also focus on generic expressions as well as unspecific indefinite referents. From a theoretical point of view, the relation and delimitation of the topic – comment structure and the given – new dichotomy needs further refinement. The proposed distinction by Trenkic into a category of consciousness and a category of knowledge must be tested empirically. Psycholinguistic experimentation could be used for testing this distinction as well as the question when information structure assignment takes place in the process of language production. This could help determine which role information structure play in assigning definiteness to a referent. This study relied on corpus data of Molise Slavic. Further research on definiteness and information structure in Molise Slavic must apply (field) experiments to further investigate the topic under discussion.

Appendices

Appendix 1: Tables of forms

Table 1: Molise Slavic Orthography with possible phonetic realizations (based on Breu 2017: 21)

a [a], [ʌ], [a]	e [e], [e]	i [i], [ɪ]	nj [ɲ]	u [u], [ʊ]
b [b]	è [ɛ]	j [j]	o [o], [ɔ]	v [v], [w]
c [ts]	ě [ə]	k [k]	ò [ɔ]	z [z]
č [tʃ]	f [f]	kj [c]	p [p]	ž [ʒ]
d [d]	g [g]	l [l]	r [r]	
dj [dj], [dʲ], [ɟ]	gh [ɣ]	lj [ʎ]	s [s]	
dz [dz]	h [x]	m [m]	š [ʃ]	
Dž [dʒ]	hj [ç]	n [n], [ɳ]	t [t]	

Table 2: Corpus frequencies of demonstrative forms in absolute numbers (corpus extracted from Breu 2011)

	DEM.PROX	n	DEM.MED	n	DEM.DIST	n
	[ovi]	805	[ta]	246	[oni]	1330
NOM.SG.M/ ACC.SG.M [-ANIMATE]	ovi ov vi	260 2 111	ta	44	oni	448
GEN.SG.M/ ACC.SG.M [+ANIMATE]	ovoga ovga voga vga	3 5 8 6	toga	4	onoga onga noga nga	12 6 9 9
DAT.SG.M	ovomu ovmu	4 0	tomu	0	onomu onmu	10 3
INS.SG.M	ovime ovme	1 2	time	0	onime onme	4 4

NOM.SG.F	ova va	111 10	ta		ona	180
GEN.SG.F/ NOM.PL	ove ve	131 28	toje te	0 93	one	302
ACC.SG.F	ovoju ovu	2 19	toju tu	1 4	onoju onu	6 36
INS.SG.F/ INS.SG.N	ovom	3	tom	0	onom onome	8 1
NOM.SG.N/ GEN.SG.N/ ACC.SG.N	ovo vo	86 7	to	90	ono	282
GEN.PL	ovihi ovhi	0 2	toihi tihi	0 6	onihi onhi	5 2
DAT.PL/ INS.PL	ovimi ovmi	1 3	Timi	4	Onimi onmi	6 7

Table 3: Full paradigm of personal pronouns (based on Breu 2017: 39.41)

NOM	GEN		DAT		ACC		INS
	Long	Clitic	Long	Clitic	Long	clitic	
'ja	mé:na _ə	mena	'men	mi	'mena _ə	ma	'menom
'ti	té:ba _ə	teba	'teb	ti	'teba _ə	ta	'tebom
'o:n	ɲé:ga _ə	ga	'ɲemu _ə	mu	'ɲega _ə	ga	'ɲime
'o:n	'je	je	'ɲo:ju _ə	ju	'ɲo:ju _ə	ju	'ɲo:m
'mi	'nasa _ə	nasa	'nami	nami	'nasa _ə	nasa	'nami
'vi	'vasa _ə	vasa	'vami	vami	'vasa _ə	vasa	'vami
'o:ne _ə	'ɲixi	xi	'ɲimi	ɲimi	'ɲixi	xi	'ɲimi

Table 4: Full paradigm of possessive pronouns (based on Breu 2017: 42-43)

	SG.M	SG.F	PL
NOM	moj	moja, mo	moje
GEN	mojoga, moga	moje	mojixi, moxi
DAT	mojomu, momu	moju	mojimi, momi
ACC	NOM:GEN	moju	moje
INS	mojime, mome	mojom	mojimi, momi
NOM	tvoj	tvoja, tvo	tvoje
GEN	tvojoga, tvoga	tvoje	tvojixi, tvoxi
DAT	tvojomu, tvomu	tvoju	tvojimi, tvomi
ACC	NOM:GEN	tvoju	tvoje
INS	tvojime, tvome	tvojom	tvojimi, tvomi
NOM	negov ~ nevog ~ nev	negova ~ nevoga ~ nevga	negove ~ nevoge ~ nevge
GEN	negova ~ nevoga ~ nevga	negove ~ nevoge ~ nevge	negovixi ~ nevogixi ~ nevxi
DAT	negovomu ~ nevogomu ~ nevmu	negovu ~ nevogu ~ nevgu	negovimi ~ nevogimi ¹² ~ nevгими
ACC	NOM:ACC	negovu ~ nevogu ¹³ ~ nevgu	negove ~ nevoge ¹⁴ ~ nevge
INS	negovime ~ nevogime ~ nevme	negovom ~ nevogom ~ nevgom	negovimi ~ nevogimi ~ nevгими
NOM	naʃ	naʃa	naʃe
GEN	naʃoga, naʒga	naʃe	naʃixi, naʃxi
DAT	naʃomu, naʃmu	naʃu	naʃimi, naʃmi
ACC	NOM:GEN	naʃu	naʃe
INS	naʃime, naʃme	naʃom	naʃimi, naʃmi
NOM	vaʃ	vaʃa	vaʃe
GEN	vaʃoga, vaʒga	vaʃe	vaʃixi, vaʃxi
DAT	vaʃomu, vaʃmu	vaʃu	vaʃimi, vaʃmi

¹² Apparently, the second form was misspelled in Breu 2017: 43.

¹³ See footnote 1

¹⁴ See footnote 1

ACC	NOM:GEN	vaŝu	vaŝe
INS	vaŝime, vaŝme	vaŝom	vaŝimi, vaŝmi
NOM	nihov ~ niŝog	nihova ~ niŝoga	nihove ~ niŝoge
GEN	nihova ~ niŝoga	nihove ~ niŝoge	nihovixi ~ niŝogixi
DAT	nihovomu ~ niŝogomu	nihovu ~ niŝogu	nihovimi ~ niŝogimi
ACC	NOM:ACC	nihovu ~ niŝogu	nihove ~ niŝoge
INS	nihovime ~ niŝogime	nihovom ~ niŝogom	nihovimi ~ niŝogimi

Appendix 2: Topic analyses

Note: Topic, comment and focus are marked with brackets in the following analysis. Utterances are marked in grey in the following analyses when their intonation was analyzed. Utterances of the referents within the narratives which are some sort of indirect speech are not marked for information structure. Glosses are only given in the examples which are discussed in the study itself. Englishes translations are added to every utterance. Topics and comments of subordinate clauses are specially indexed with the abbreviations SC-TOP (topic of the subordinate clause) and SC-COM (comment of the subordinate clause. Information in the sentence which is unimportant for the analysis is marked as NAI (not-at-issue) material.

Analysis 1: Excerpt from The Little Prince (Saint-Exupéry 2009: 18-19)

1. [Ø] _{TOP} [Naka sa bi znaja jenu second stvaru čuda jimbortand!] _{COM}	1. I had thus learned a second fact of great importance:
2. [Ova] _{TOP} [je ka [pjanet jiskla] _{SC-TOP} [je bi doša sama mala veča velki do jene hiže] _{SC-COM} !] _{COM}	2. this was that the planet the little prince came from was scarcely any larger than a house!
3. [Ovo] _{TOP} [ma ne moraša čit zbaučit čuda.] _{COM}	3. But that did not really surprise me much.
4. [Ø] _{TOP} [Znadahu dobra ka, [vana do pjaneti velke, kana Kaša, Džove, Marte, Venere,] _{SC-COM} [kojimi] _{SC-TOP} [su dal jena jiman] _{SC-COM} , [jesu] _{SC-COM} [stotine do drugihi] _{SC-TOP} , ka [jesu] _{SC-COM} [koju] _{SC-TOP} [votu naka male ka je difičil] _{SC-COM} [hi] _{SC-TOP} [vit s teleskopjam.] _{SC-COM}] _{COM}	4. I knew very well that in addition to the great planets--such as the Earth, Jupiter, Mars, Venus--to which we have given names, there are also hundreds of others, some of which are so small that one has a hard time seeing them through the telescope.

5. [Kada jena aštronom skopriva] _{COM} [jena do'vihi] _{TOP} , [mu] _{TOP} [daja kana jiman jena numar.] _{COM}	5. When an astronomer discovers one of these he does not give it a name, but only a number.
6. [Ga] _{TOP} [zova presembj: << ašterojid 325>>]. _{COM}	6. He might call it, for example, "Asteroid 325."
7. [Ø] _{TOP} [Jimam razloga serije za verjat ka [pjanet do di gredaša mali kraljič] _{SC-TOP} [je ašterojid B 612.] _{SC-COM}] _{COM}	7. I have serious reason to believe that the planet from which the little prince came is the asteroid known as B-612.
8. [Ovi ašterojid] _{TOP} [je bija vidan sama jenu votu di teleskopij, lu 1909 do jenga aštronomya turk.] _{COM}	8. This asteroid has only once been seen through the telescope. That was by a Turkish astronomer, in 1909.
9. [On] _{TOP} [je bi čija nonda jenu veliku dimoštracijunu] _{COM} do'no ka [Ø] _{TOP} [je bi skoprija, na jena Kongres Jindernacional do Aštronomyje.] _{COM}	9. On making his discovery, the astronomer had presented it to the International Astronomical Congress, in a great demonstration.
10. [Ma nikor nije bi] _{COM} [ga] _{TOP} [verja za kaka] _{COM} [Ø] _{TOP} [biša bučan.] _{COM}	10. But he was in Turkish costume, and so nobody would believe what he said.
11. [Čeljade velke] _{TOP} [jesu činjene naka.] _{COM}	11. Grown-ups are like that . . .
12. [Je bila na fortuna] _{COM} [za reputacijunu do ašterojida B 612] _{TOP} [ka] _{NAI} [jena ditator turk] _{SC-TOP} [je kumana zgora njevhi čeljadi, s kaštigom do smrtve, za [Ø] _{TOP} [sa buč kana čeljade europe.]. _{SC-COM}] _{COM}	12. Fortunately, however, for the reputation of Asteroid B-612, a Turkish dictator made a law that his subjects, under pain of death, should change to European costume.
13. [Aštronom] _{TOP} [je čija jopa njevogu dimoštracijunulu 1920 s jenme vištitam čuda elegand.]. _{COM}	13. So in 1920 the astronomer gave his demonstration all over again, dressed with impressive style and elegance.
14. [A ovu votu] _{NAI} [tuna sfit je] _{COM} [ga] _{TOP} [verja] _{COM} .	14. And this time everybody accepted his report.

Analysis 2: The Russian tractor (Breu 2017: 97-99)

S1	[Ja] _{TOP} [mahu [kupi trator] _{FOC} nonda, one dana] _{COM} . At that time, I had to buy a tractor, those days.
S2	[Bihu mala soldi, alora bihu [ove tratora russi] _{FOC} ke...] _{COM} There wasn't much money, but there were these Russian tractors, which...
S3	[Vabene] _{NAI} , [ø] _{TOP} [koštahu veča mala] _{COM} . Well, [they] cost less.
S4	[Alora] _{NAI} [ja] _{TOP} [sa [vaza] _{FOC} jena trator rus] _{COM} . So I took (=bought) a Russian tractor,
S5	[e] _{NAI} [ø] _{TOP} [sa hodiya torko lipa] _{COM} . [Oda, nu votu] _{NAI} , [su dol pur] _{TOP} [dva rusa] _{COM} : [<i>jena mekanik</i>] and it went very well. Once even two Russians came here: a mechanic
S6	<i>aš jena ndžinjir</i> _{TOP} , [ø] _{TOP} [su dol di sa ja] _{COM} [oda, nu votu] _{NAI} , [ø] _{TOP} [vit kaka gredahu ove magine] _{COM} . and an engineer, they came once to me here, to see, how these machines ran.
S7	[E] _{NAI} [ø] _{TOP} [sma kijekarijal na mala skupa] _{COM} . And we spoke a bit with each other.
S8	[Pèrò] _{NAI} [ja] _{TOP} [pa, nu votu, što sa čija?] _{COM} But once, what did I then (go and) do?
S9	[Sa [ga] _{TOP} [tija pita nu stvaru extra] _{FOC}] _{COM} , [ovga ndžinjira.] _{TOP} I wanted to ask him a specific question, this engineer.
S10	[Kaka stahu] _{COM} [kondadina Larusja] _{TOP} , What it was like for the farmers in Russia,
S11	[kaka] _{COM} [ø] _{TOP} [živahu nonde] _{COM} , [one vrima nonda] _{NAI} . how they lived there, in those times.
S12	[ø] _{TOP} [Je reka] _{COM} : „ <i>Sì ,stoju lipa, one jimaju kamba də ping pong, jokaju a ping pong, stoju torko lipa!</i> ” He said: “Yes, they’re fine, they have tennis courts, they play tennis, they’re getting on really well!”
S13	[Pèrò] _{NAI} [ø] _{TOP} [sa reka] _{COM} : « <i>Vi ka... sa čuja reč ka ona njiva ka rabi... rusa,</i> But I said: “Look... I’ve heard it said that that field that... the Russians work on...
S14	<i>koju rabu operaja za njifog kunat, čini robu.</i> that the farm labourers work on for themselves, yields something.
S15	<i>Ona ke rabi cond delo stat, ne čini.</i> That one (=those) he (=the Russian) works on behalf of the state, does not yield”.
S16	[Oni] _{TOP} [što je čija?] _{COM} [ø] _{TOP} [Je vaza, je sa usta, je si ga pošā] _{COM} . That one, what did he do? He got down to it, stood up, walked off.

S17	[Ė] _{NAI} , [[oni] _{FOC} {ka [mi] _{SC-TOP} [proda mutor] _{SC-COM} [men] _{sub-TOP} }] _{TOP} , Well, the guy, who sold me the tractor,
S18	[ovi {ka [mi] _{SC-TOP} [ga [proda] _{sub-FOC}] _{SC-COM} }] _{TOP} [je [umbra] _{FOC}] _{COM} . the one who had sold it to me died.

Appendix 4: The hare and the crayfish (Breu 2017: 89-94)

S1	[Zec eš rak] _{TOP} The hare and the crayfish.
S2	[Nu votu] _{NAI} [[na zec] _{FOC}] _{TOP} [[prohodaša] _{FOC} jizbane jene rike] _{COM} . Once there was a hare walking along a river.
S3	[ø] _{TOP} [Je frunda [jenga raka] _{FOC}] _{COM} . He met a crayfish.
S4	[ø] _{TOP} [Je mu reka] _{COM} : He said to him.
(S5)	«Kumba rak, di maš pokj, ka si jiskodija do vode?» “Godfather crayfish, where are you going (to), that needs you to come out of the water?”
S6	[[Je reka] _{FOC}] _{TOP} [rak] _{COM} : {ALTERNATIVE: [[Je reka] _{FOC}] _{COM} [rak] _{TOP} } The crayfish said:
(S7)	«Mam po nama-gor, zgora onga brda». “I will go up there, on the hill over there”.
S8	[[Zec] _{FOC}] _{TOP} [je sa vrga smijat] _{COM} , The hare began to laugh,
S9	[ø] _{TOP} [je reka] _{COM} : «A kada maš rivat ti nama-gor?» he said: “And when will you reach the top?”
(S10)	Ti sa hoče na dan!» You’ll need a day!”
S11	[Rak] _{TOP} [je reka] _{COM} : «Nomo sa smijat, The crayfish said: “Don’t laugh,
(S12)	ka ja jesa kapač rivat nama-gor prije do teba!» because I’m quite capable, of getting up there before you!”
S13	[Zec] _{TOP} [je reka] _{COM} : „Ma nomo ma či smijat!” The hare said: “Don’t make me laugh!”
S14	[[Rak] _{FOC}] _{TOP} [je reka] _{COM} : «Sa čaš vit? The crayfish said: “Do you want to see now?
S15	[[Homa] _{FOC} či nu skumasu] _{COM} ! Let’s bet on it!
(S16)	Andz, ču ta či bijat pur teba prije! I’ll even let you go first!
(S17)	Bija-sa e čaš vit ka ja rivoam prije do teba nama-gor!» Run, and you’ll see that I’ll get to that top before you!”
(S18)	[Zec] _{TOP} [je reka] _{COM} : The hare said:
S19	«[Alor homa vit] _{NAI} , [[homa čit] _{FOC}] _{COM} [ovu skumasu] _{TOP} !» “Well, let’s see, let’s bet on it!”

S20	[Zec] _{TOP} [je sija za parti veća bolje] _{COM} The hare crouched down, to get a better start,
S21	[e] _{NAI} [rak ka biša naza njega] _{TOP} and the crayfish, who was behind him,
S22	[je mu uhitija rep (s) škarami] _{COM} grabbed his tail with his claw,
S23	[e naka] _{NAI} , [kada] _{NAI} [zec] _{TOP} [je sa... je poča teč], and so, when the hare... began to run,
S24	[rak] _{TOP} [gredaša s njime] _{COM} . the crayfish went with him.
S25	[Zec] _{TOP} [teča, teča, riviva zgora brda, The hare runs and runs, gets to the top of the hill,
S26	sa brnjiva naza he turns around,
S27	za vit si rak gredaša zgoru brdam o si bi osta još dol jizbane rike.] _{COM} in order (to be able) to see, if the crayfish was running up the hill or had still remained beside the river.
S28	[Kaka] _{NAI} [on] _{TOP} [je sa zbrnija] _{COM} , As he turned back around,
S29	[rak] _{TOP} [je sa skinija do njega repata] _{COM} , the crayfish dropped down from his tail,
S30	[e] _{NAI} [ø] _{TOP} [je sa lundana na mala], [a] _{NAI} [ø] _{TOP} [je mu reka] _{COM} : and stepped back a bit, and said to him:
(S31)	«Kumba zec, sa si riva? “Godfather hare, have you arrived now?
(S32)	Ja, je no lipo mala ka ta čekam oda!» Me, it’s been ages since I’ve been waiting here for you!”
S33	[Zec] _{TOP} [je osta] _{COM} , The hare remained (astonished),
S34	[ø] _{TOP} [je reka] _{COM} : «Ma kada si riva?» he said: “But when did you get here?”
(S35)	«È, jesu, benja... “Well, it is, perhaps...
(S36)	je kvaš no lipo mala ka sa riva!» it is a fair while, since I got here!”
S37	[E] _{NAI} [[naka] _{FOC}] _{TOP} [zec je zgubija skumasu] _{COM} . {ALTERNATIVE: [E] _{NAI} [[naka] _{FOC}] _{COM} [zec] _{TOP} [je zgubija skumasu] _{COM} . And so the hare lost the bet.

Appendix 3: Intonation analyses

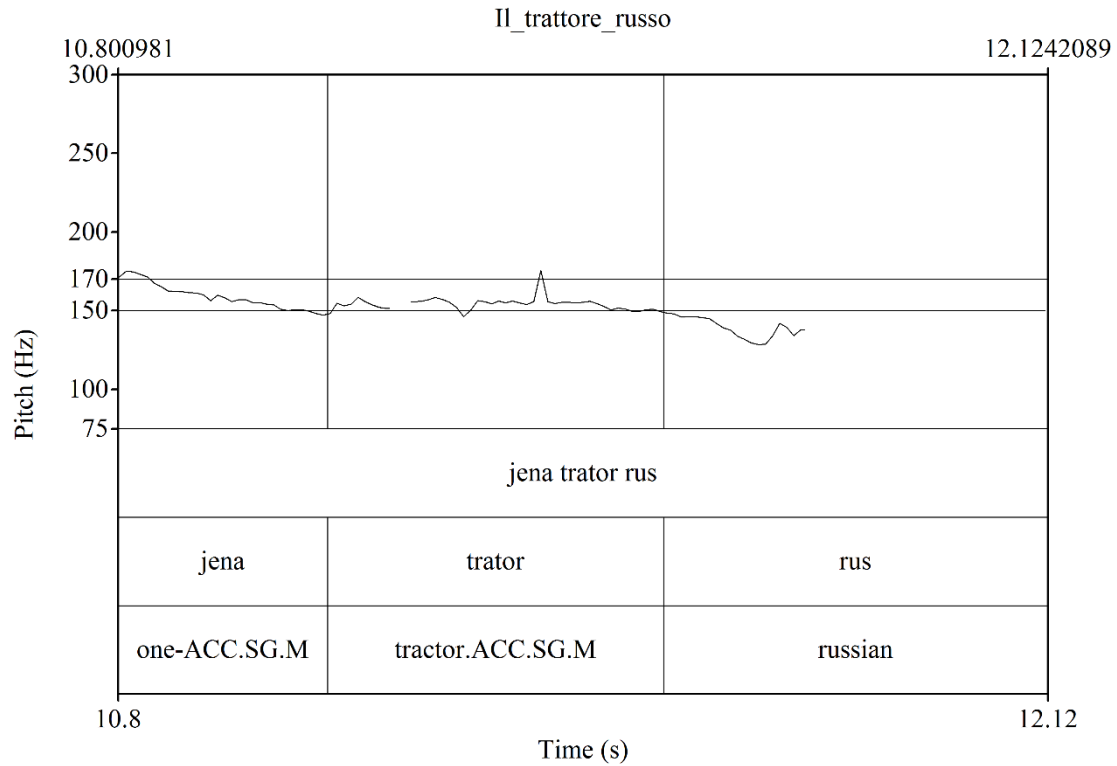
List of examples annotated with Praat and MS-ToBI

Note: Praat pictures appearing in the text are marked in gray in the following table. Glosses are provided in the praat analysis but not in the following table. Topic and focus are marked in the table.

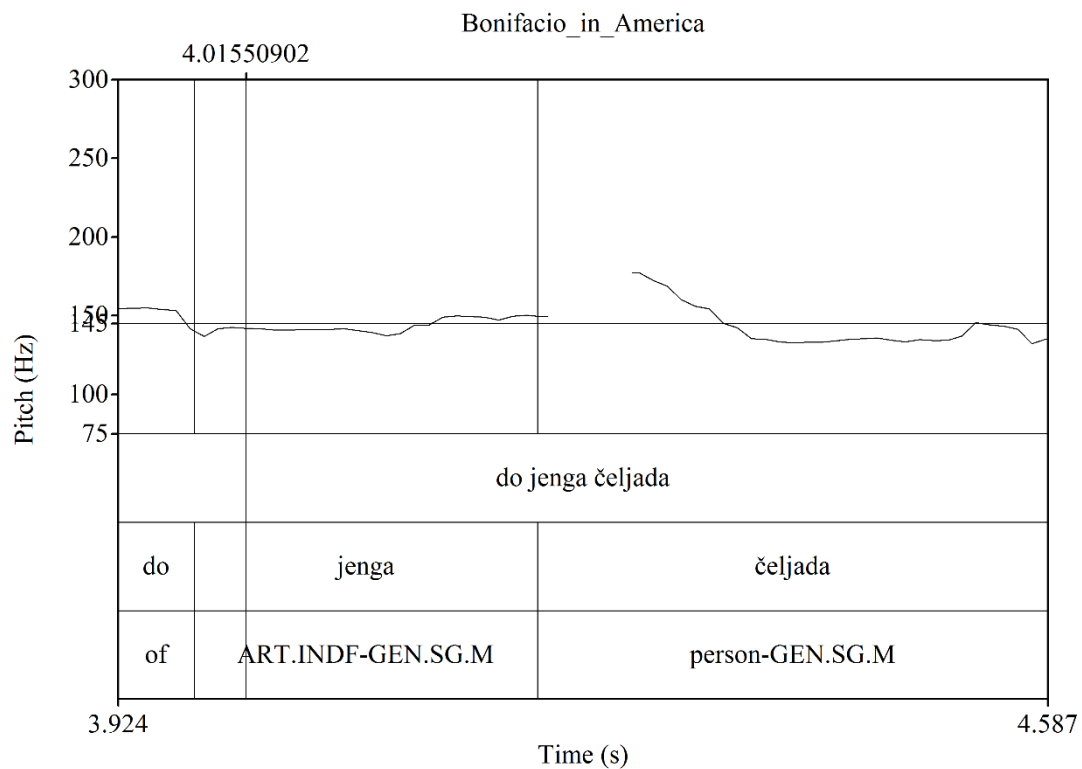
Number in text	Sentence	Type	Topic	Focus
1	jena trator rus	Numeral	-	-
2	do jenga čeljada	Indefinite article		
3	allora [ja] _{TOP} sa [vaza] _{FOC} jena trator rus	Declarative	Pronominal topic, slight bridge contour	Heavily stressed
4	za ta slušat [bolje] _{FOC} , moja divojka	Answer to wh-question	Subject-less sentence	Heavily stressed
5	[[zec] _{FOC}] _{TOP} [je sa vrga smijat] _{comment}	Declarative	TOP = FOC	stressed
6	[[je reka] _{FOC}] _{TOP} rak	Declarative	Verbal topic (Narrative inversion)	FOC = TOP
7	[[zec] _{FOC}] _{TOP} je reka	Declarative	TOP = FOC (bridge contour)	
8	[zec] _{TOP1} aš [rak] _{TOP2}	headline	Contrastive topic, bridge contour on first topic	Not present
9	[[jena tata] _{TOP} jimaša dva sina.] _{FOC}	Declarative, beginning of discourse	Not intonationally marked	All-new-focus with gradual pitch reduction
10	[nu votu [na zec] _{TOP} prohodaša jizbane jene rike.] _{FOC}	Declarative, beginning of discourse	TOP within all new focus	Heavily stressed, all-new-focus
11	[[ja] _{TOP} mahu [kupi trator] _{FOC} nonda one dana.]	Declarative	Pronominal topic, not intonationally marked	Heavily stressed
12	[ø] je frunda [jenga rak.] _{FOC}	Declarative	Subject drop (=zero topic)	Slightly stressed
13	[ø] jimaša [čuda robi] _{FOC} [ovi tata] _{TOP}	Declarative	Right dislocation with subject drop, in post-focal position	Heavily stressed
14	sa [ga] _{TOP} [tija pita nu stvaru extra] _{FOC} , [ovga ndzinjira] _{TOP}	Declarative	Right dislocation with clitic, clitic not stressed, dislocated object in postfocal position	Heavily stressed
15	[E naka] _{FOC} [zec] _{TOP} zgubija skumasu.	Declarative, end of discourse	No intonational effect	Focus heavily stressed

Praat pictures of pitch contours

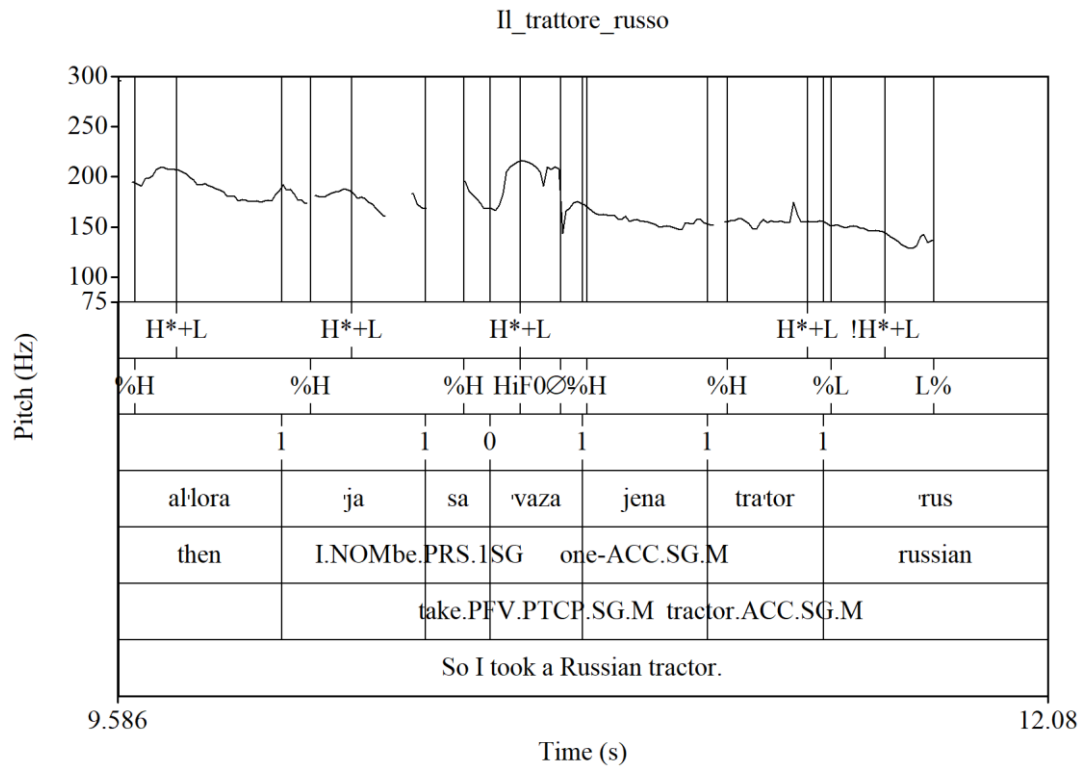
Example 1: Numeral *jena*



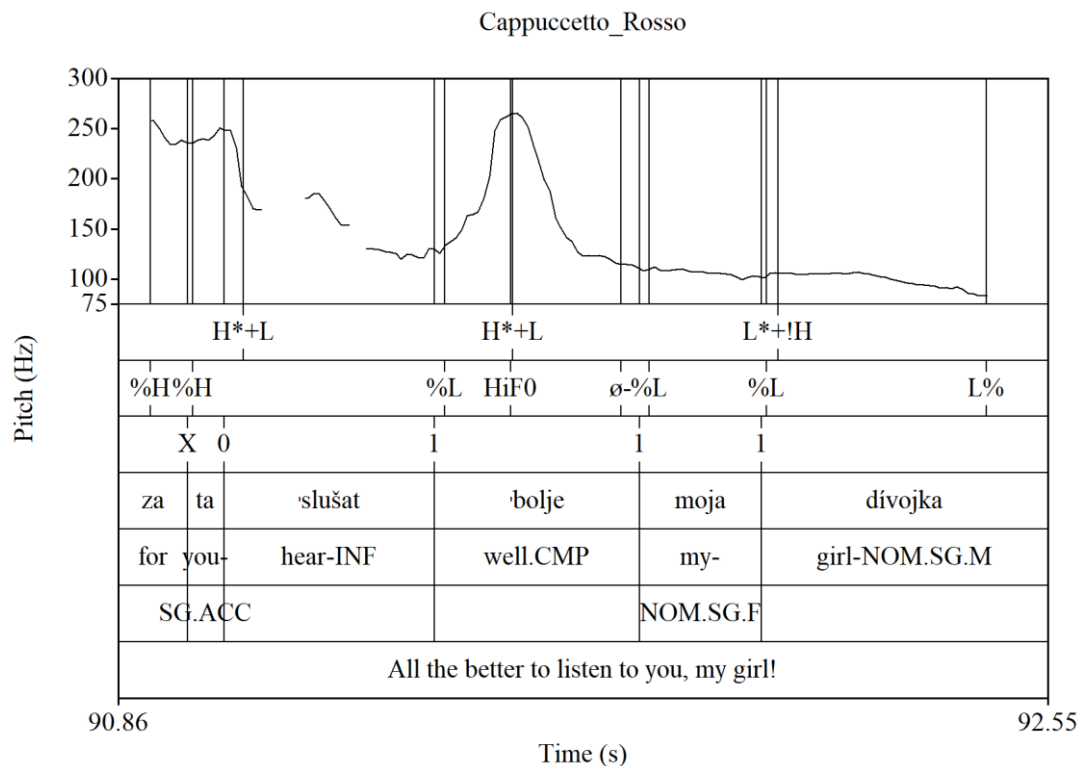
Example 2: Indefinite article



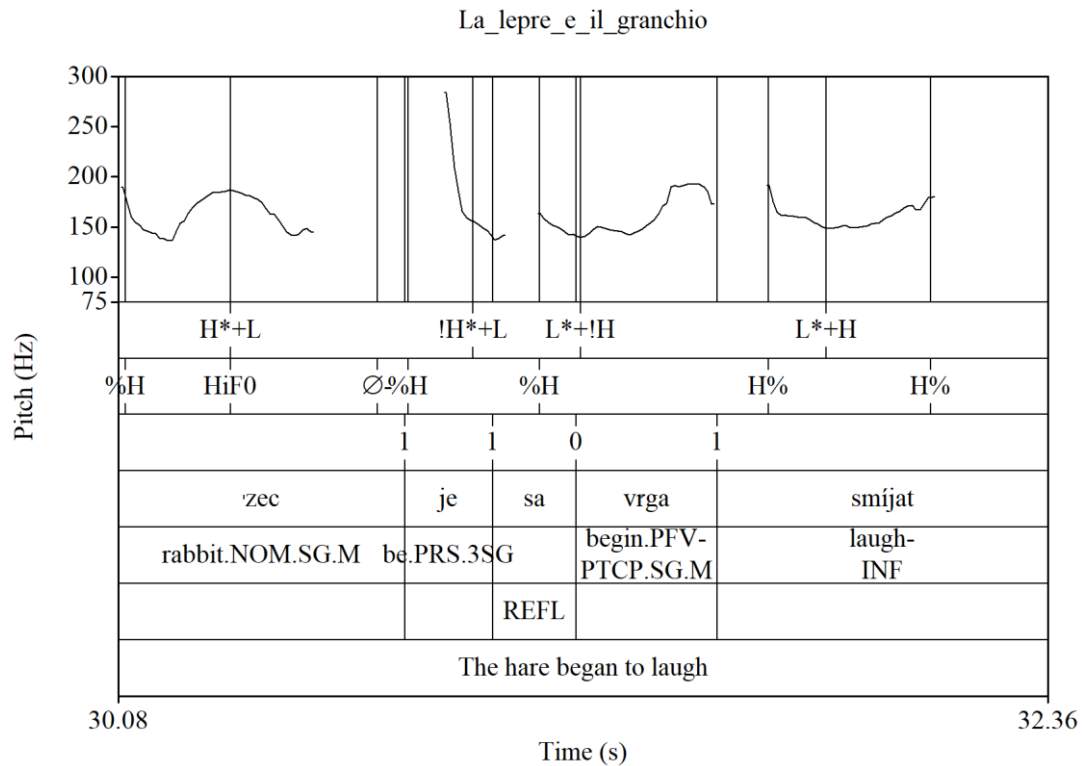
Example 3: Declarative sentence with pronominal topic



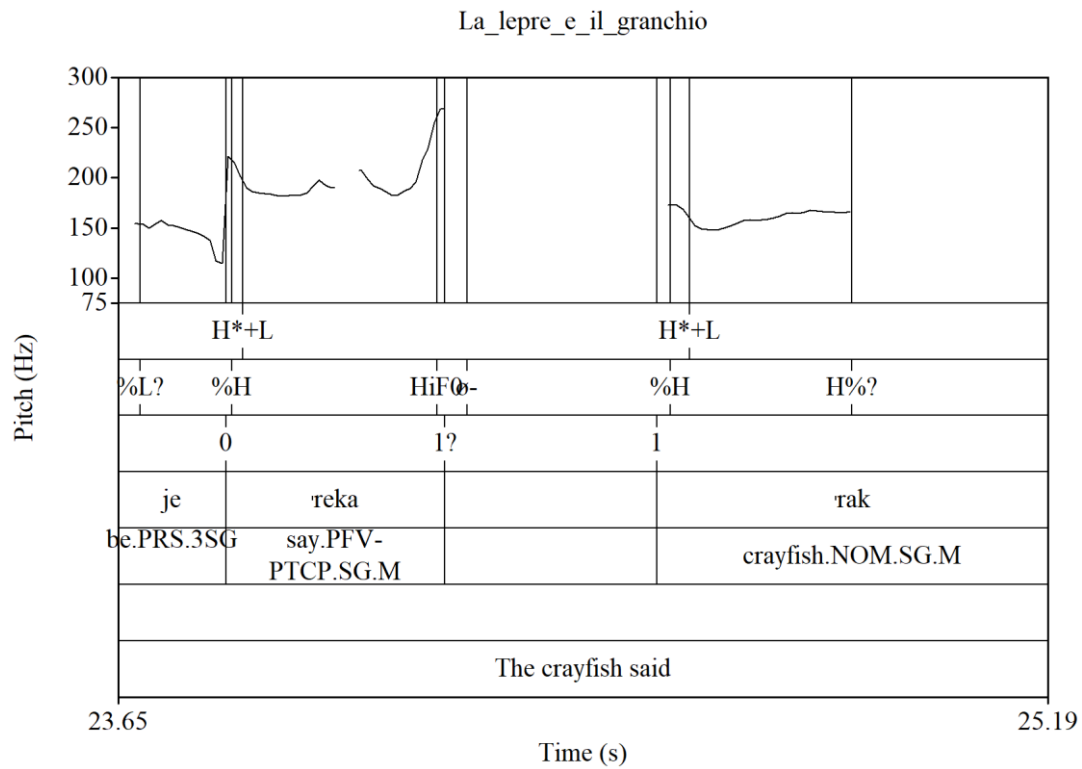
Example 4: Declarative sentence with subject drop



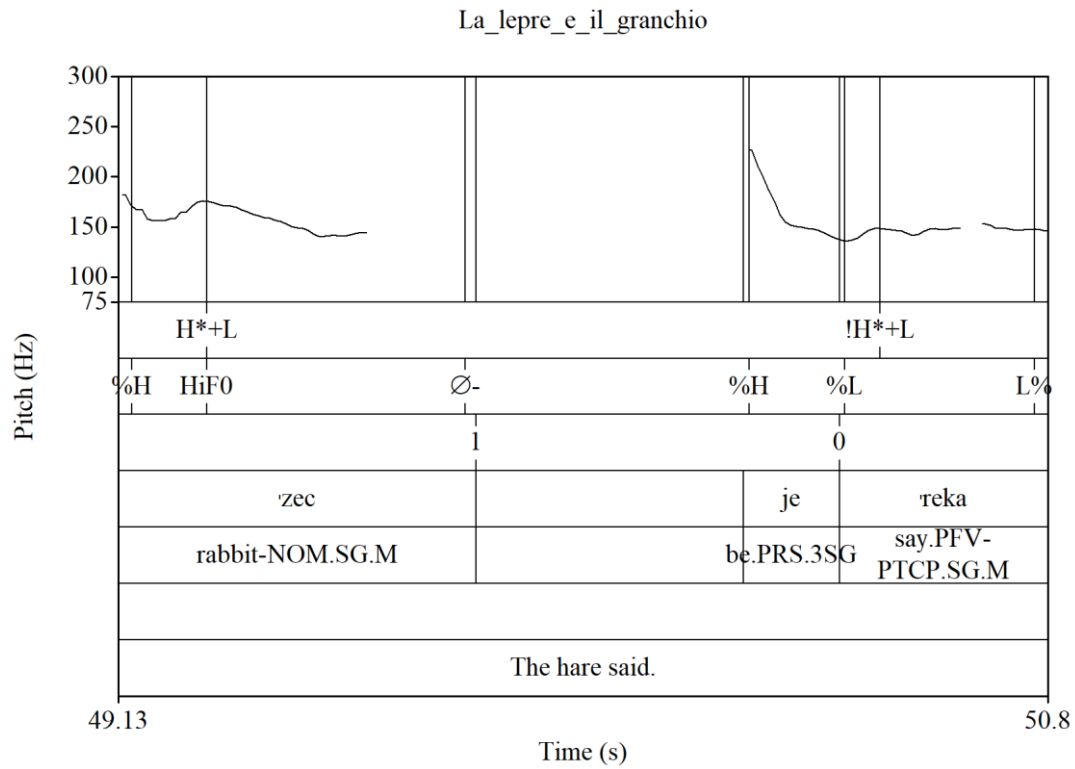
Example 5: Declarative sentence with focussed topic



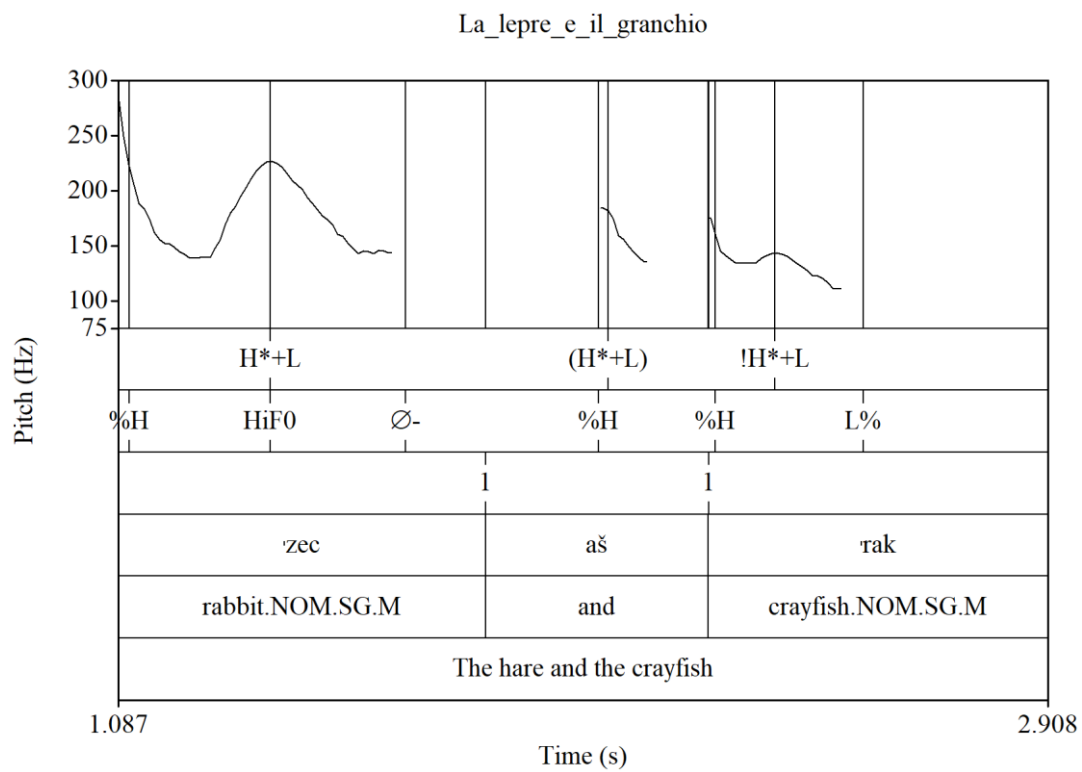
Example 6: Declarative sentence with verbal topic (Narrative inversion)



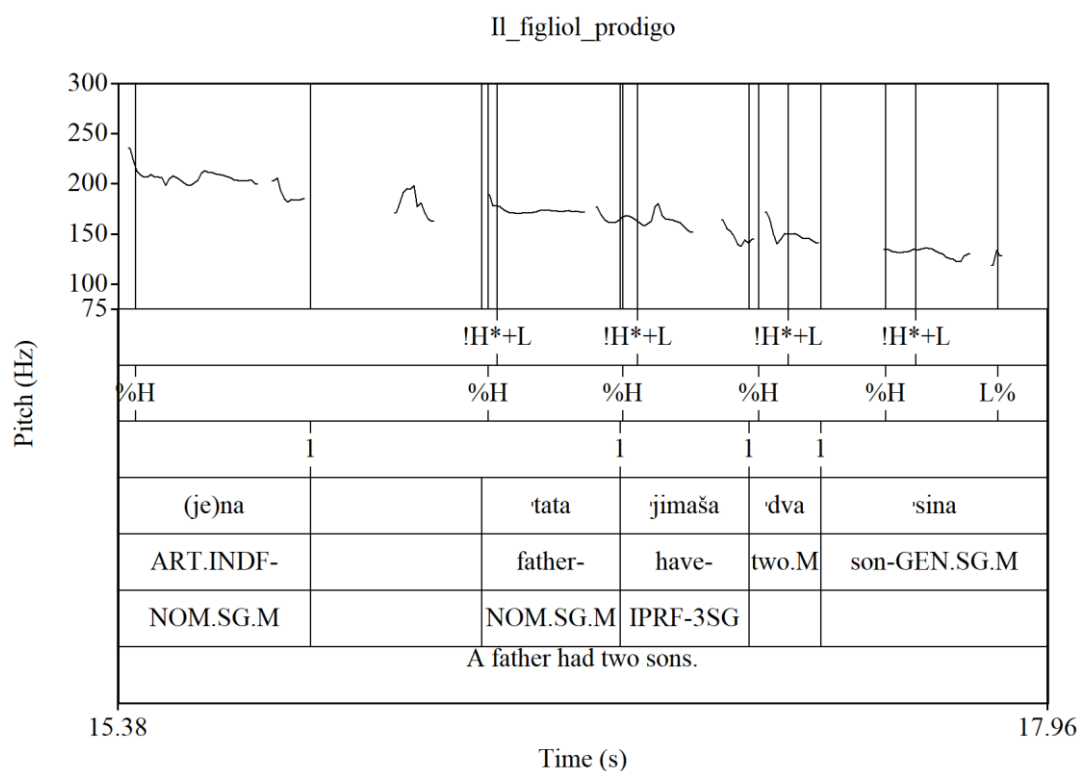
Example 7: Declarative sentence with focussed topic (bridge)



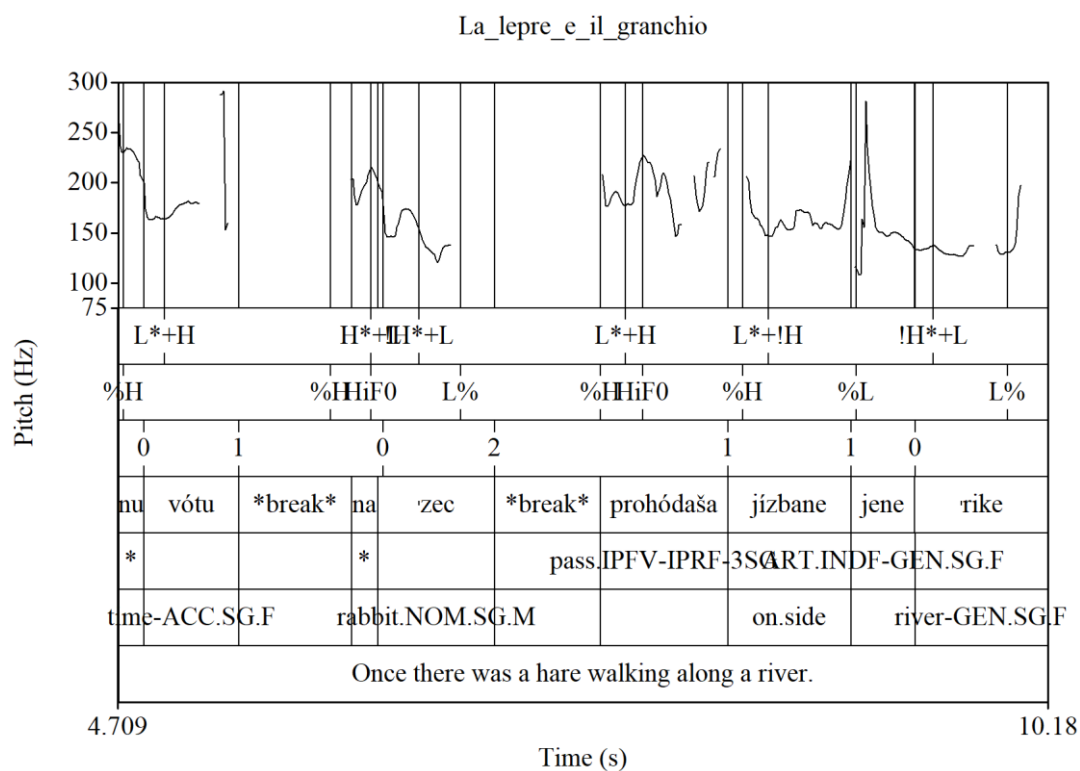
Example 8: Contrastive topic in headline



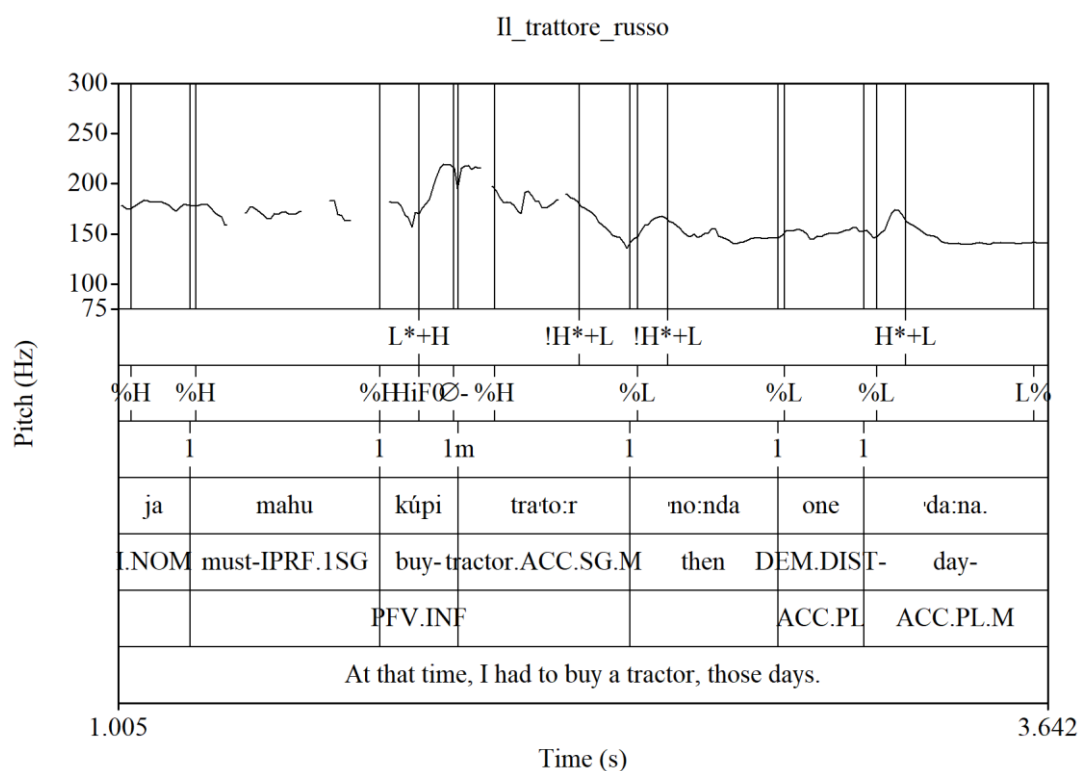
Example 9: Declarative sentence (beginning of discourse)



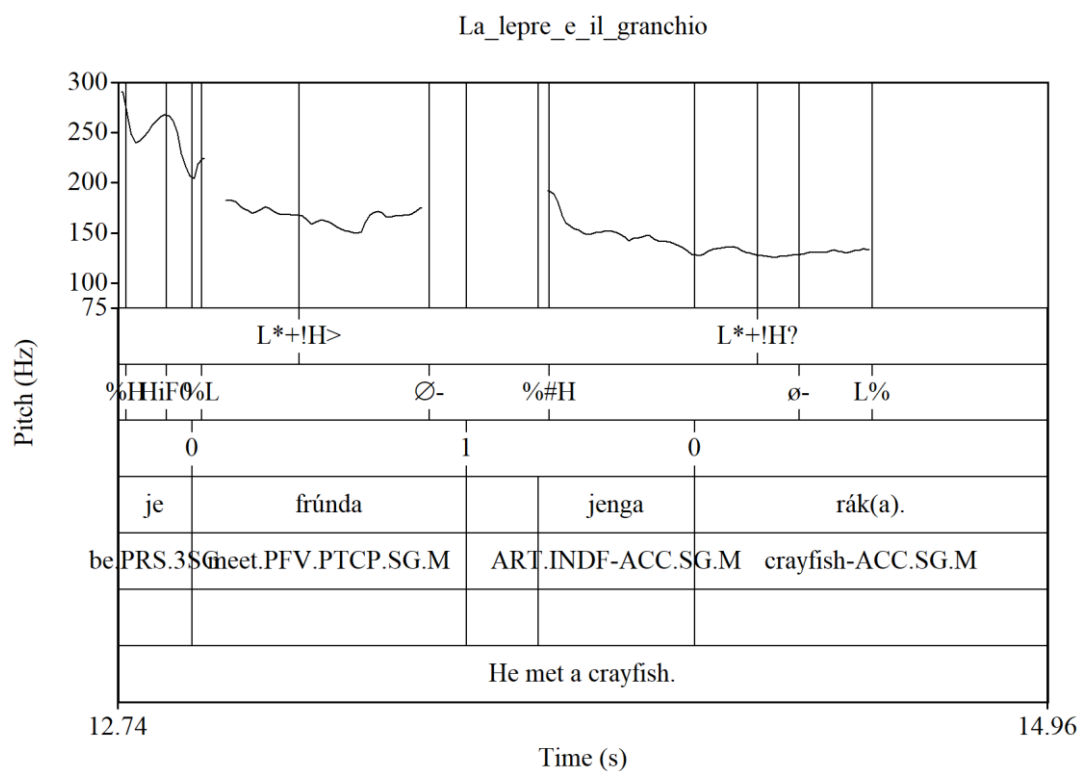
Example 10: Declarative sentence (beginning of discourse)



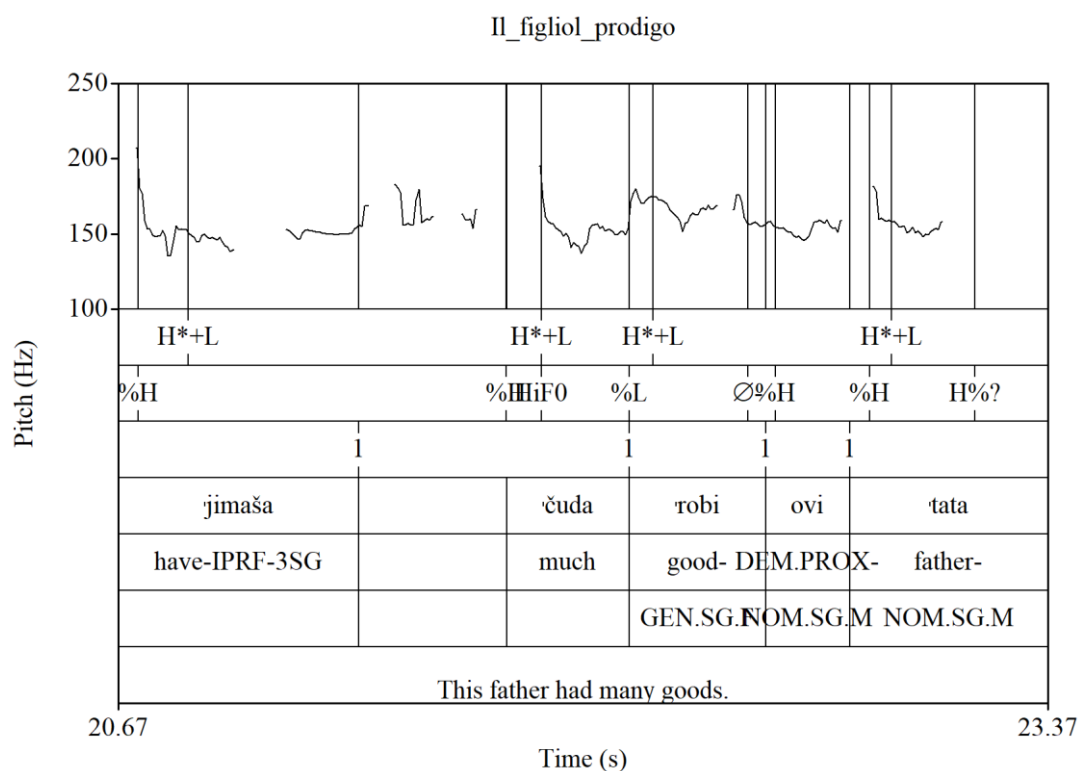
Example 11: Declarative sentence with pronominal topic



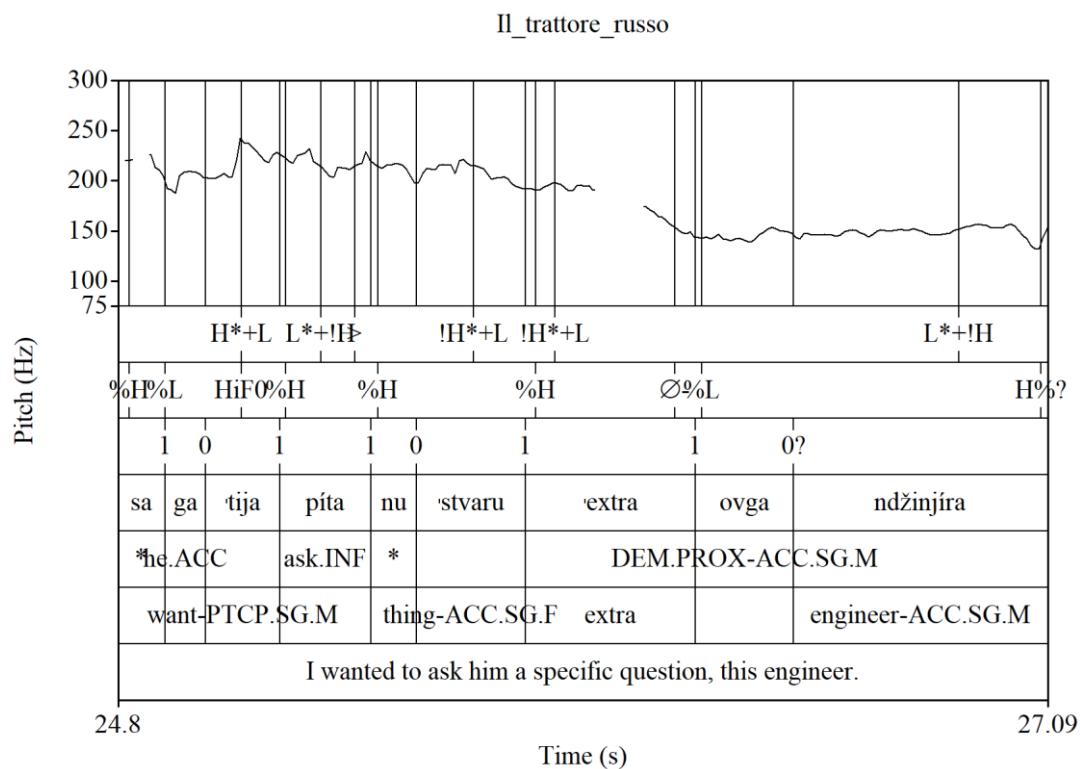
Example 12: Declarative sentence with subject drop



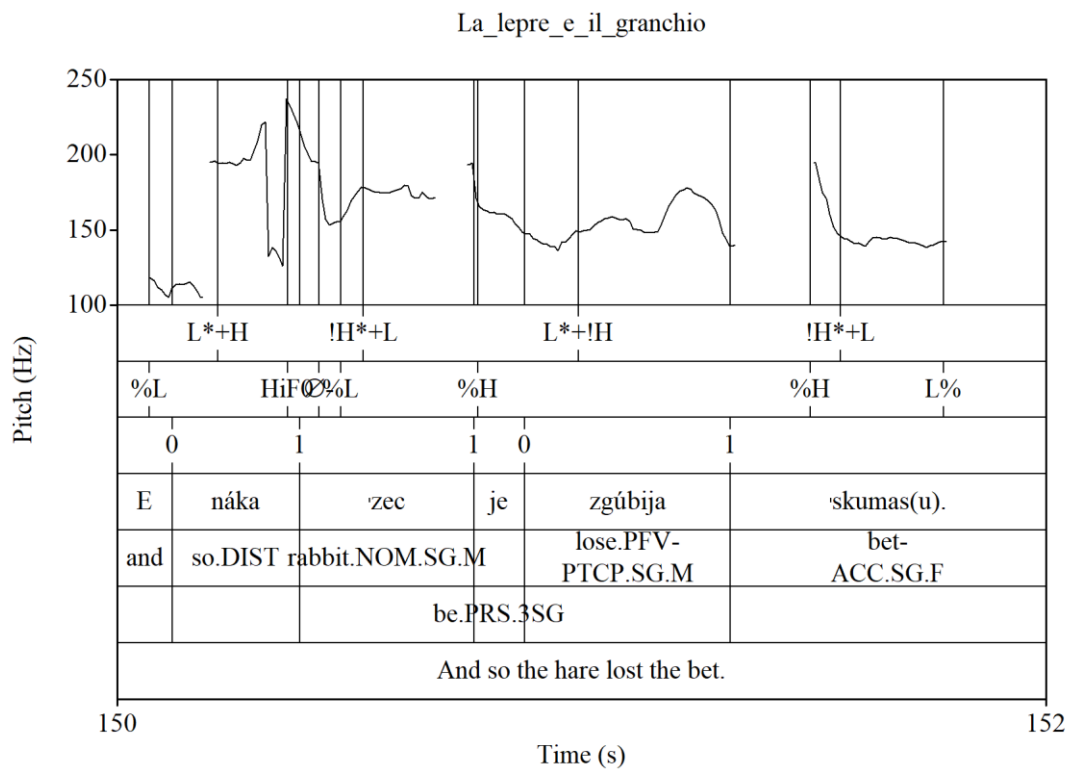
Example 13: Right dislocation



Example 14: Right dislocation with clitic



Example 15: Declarative sentence (end of discourse)



Literature

- Alexander, Ronelle. 2006. *Bosnian, Croatian, Serbian, a grammar: With sociolinguistic commentary*. Madison (WI).
- Belaj, Branimir, and Darko Matovac. 2015. On the article-like use of the indefinite determiners *jedan* and *neki* in Croatian and other Slavic languages. *Suvremena lingvistika* 79 (1). 1-20.
- Bel'skij, A. B. 1956. Intonacija kak sredstvo determinirovanija i predicirovanija v ruskom literaturnom jazyke. In: *Issledovanija po sintaksisu russkogo literaturnogo jazyka*. Ed. V. P. Suhotin. Moskva. 188-199.
- Birkenmaier, Willy. 1979. *Artikelfunktionen in einer artikellosen Sprache: Studien zur nominalen Determination im Russischen*. München.
- Birner, Betty J. 2013. *Introduction to Pragmatics*. Chichester (West Sussex).
- Breu, Walter. 1990. Sprache und Sprachverhalten in den slavischen Dörfern des Molise (Süditalien). In: *Slavistische Linguistik 1989*. Ed. Walter Breu. München: 35-65
- Breu, Walter. 2005. Il sistema degli articoli nello slavo molisano (SLM). In: *L'influsso dell'italiano sulla grammatica delle lingue minoritarie.: Problemi morfologici e sintattici. Atti del Convegno Internazionale - Costanza, 8 - 11 ottobre 2003*. Ed. Walter Breu. Rende. 111-139.
- Breu, Walter. 2012. The grammaticalization of an indefinite article in Slavic micro-languages. In: *Grammatical replication and borrowability in language contact*. Ed. Björn Wiemer, Bernhard Wälchli, Björn Hansen. Berlin. 275-322.
- Breu, Walter. 2017. *Slavische Mikrosprachen im absoluten Sprachkontakt: Teil I: Moliseslavische Texte aus Acquaviva Collecroce, Montemitro und San Felice del Molise*. Wiesbaden.
- Breu, Walter, and Giovanni Piccoli. 2000. *Dizionario croato molisano di Acquaviva Collecroce: Dizionario plurilingue della lingua slava della minoranza di provenienza dalmata di Acquaviva Collecroce in Provincia di Campobasso. Dizionario, registri, grammatica, testi*. Campobasso.
- Breu, Walter, and Giovanni Piccoli. 2011. *Südslavisch unter romanischem Dach: Die Moliseslaven in Geschichte und Gegenwart im Spiegel ihrer Sprache. Teil I. Texte gesprochener Sprache aus Acquaviva Collecroce*. München.
- Breu, Walter, and Giovanni Piccoli. 2012. *Südslavisch unter romanischem Dach: Die Moliseslaven in Geschichte und Gegenwart im Spiegel ihrer Sprache. Teil II. Texte gesprochener Sprache aus Montemitro und San Felice del Molise*. München.
- Burkhardt, Petra. 2005. *The syntax-discourse interface: Representing and interpreting dependency*. Amsterdam/ Philadelphia.
- Brunetti, Lisa. 2009a. On links and tails in Italian. In: *Lingua* 119(5). 756-781.
- Brunetti, Lisa. 2009b. On the semantic and contextual factors that determine topic selection in Italian and Spanish. *The linguistic review* 26. 261-289.

- Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In: *Subject and Topic*. Ed. Charles N. Li. New York. 25-55.
- Chen, Ajou. 2012. The prosodic investigation of information structure. In: *The expression of information structure*. Ed. Manfred Krifka and Renate Musan. Berlin. 249-286.
- Cowles, Heidi Wind. 2012. The psychology of information structure. In: *The expression of information structure*. Ed. Manfred Krifka and Renate Musan. Berlin. 287-317.
- Diessel, Holger. 2013. Distance Contrasts in Demonstratives. In: *World atlas of language structures online*. Ed. Matthew Dryer, Martin Haspelmath. Leipzig. <<http://wals.info/chapter/41>> (accessed on 2017-11-09).
- Dryer, Matthew. 2013. Indefinite articles. In: *World atlas of language structures online*. Ed. Matthew Dryer, Martin Haspelmath. Leipzig. <<http://wals.info/chapter/38>> (accessed on 2017-11-27).
- Erteschik-Shir, Nomi. 2014. Information structure and (in)definiteness. In: *Crosslinguistic studies on noun phrase structure and reference*. Ed. Nomi Erteschik-Shir, Patricia Cabredo Hofherr, and Anne Zribi-Hertz. Leiden. 23-50.
- Firbas, Jan. 1975. On the thematic and non-thematic ection of the sentence. *Style and text*. 314-334.
- Givón, Talmy. 1976. Topic, pronoun, and grammatical agreement. In: *Subject and Topic*. Ed. Charles N. Li. New York. 149-188.
- Givón, Talmy. 1981. On the development of the numeral 'one' as an indefinite marker. *Folia Linguistica Historica* 2 (1). 35-53.
- Godjevac, Svetlana. 2005. Transcribing Serbo-Croatian intonation. In: *Prosodic typology: The phonology of intonation and phrasing*. Ed. Sun-Ah Jun. Oxford: 146-171.
- Grice, Martine. 2006. Intonation. In: *Encyclopedia of Language and Linguistics*. Ed. Keith Brown. Oxford. 778-788.
- Gundel, Jeanette. 1989. *The role of topic and comment in linguistic theory*. Garland.
- Gvozdanovic, Jadranka. 2015. Numerale. In: *Wörterbücher zur Sprach- und Kommunikationswissenschaft (WSK) Online* (2015). Berlin. <https://www.degruyter.com/view/WSK/wsk_id_wsk_artikel_artikel_19396> (accessed on 2017-11-01).
- Halliday, Michael. 1967. Notes on transitivity and theme in English. *Journal of Linguistics* 3 (1-2). 177-274.
- Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. Amherst.
- Heine, Bernd. 1997. *Cognitive foundations of grammar*. New York.
- Heine, Bernd, and Tania Kuteva. 2006. *The changing languages of Europe*. Oxford.
- Jasinskaja, Katja. 2016. Information structure in Slavic. In: *The Oxford handbook of information structure*. Ed. Caroline Féry and Shinichiro Ishihara. New York. 709-732.

Javarek, Vera and Miroslava Sudjić. 1963. *Serbo-Croat*. London.

Junghanns, Uwe, and Gerhild Zybatow. 2009. Grammatik und Informationsstruktur. In: *Die slavischen Sprachen. Ein internationales Handbuch zu ihrer Struktur, ihrer geschichte und ihrer Erforschung* (=Handbücher zur Sprach- und Kommunikationswissenschaft 32.1). Ed. Sebastian Kempgen, Peter Kosta, Tilam Berger, and Karl Gutschmidt. Berlin/ New York. 648-707.

Kamp, Hans. 1981. A theory of truth and semantic representation. In: *Formal methods in the study of language*. Ed. Jeroen Groenendijk, Theo Janssen, and Martin Stokhof. Amsterdam. 277-322.

Kreisberg, Alina. 2007. Ancora a proposito dell'articolo slavo-molisano: Un tentativo di raffronto con altre lingue slave. In: *Minoranze linguistiche: Prospettive, strumenti, territori*. Ed. Carlo Consani, Paola Desideri. Rome. 205-217.

Krifka, Manfred, and Renate Musan. 2012. Information structure: overview and linguistic issues. In: *The expression of information structure*. Ed. Manfred Krifka and Renate Musan. Berlin. 1-43.

Kunzmann-Müller, Barbara. 2002. *Grammatikhandbuch des Kroatischen: unter Einschluss des Serbischen*. Frankfurt am Main.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge.

Levelt, Willem. 1989. *Speaking. From intention to articulation*. Cambridge.

Lyons, Christopher. 1999. *Definiteness*. Cambridge.

Matić, Dejan. 2015. Information structure in linguistics. In: *International encyclopedia of the social and behavioral sciences* 12. Ed. Neil Smelser, and Paul Baltes. Amsterdam.

Matras, Yaron. 2008. The borrowability of structural categories. In: *Grammatical Borrowing in Cross-linguistic Perspective*. Ed. Yaron Matras and Jeanette Sakel. Berlin. 31-73.

Meillet, Antoine, and André Vaillant. 1980. *Grammaire de la langue serbo-croate*. Paris.

Moskalskaja, Olga. 1971. *Grammatik der deutschen Gegenwartssprache*. 1971.

Payne, Thomas. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge.

Prince, Ellen. 1979. On the given/ new distinction. In: *Papers from the fifteenth Regional Meeting, Chicago Linguistic Society, Chicago*. Ed. Paul Clyne, William Hanks, and Carol Hofbauer. Chicago. 267-278.

Reichenkron, Günter. 1934. Serbokroatisches aus Süditalien. *Zeitschrift für Slavische Philologie* 11 (3/4). 325-339.

Reinhart, Tanya. 1981. Pragmatics and Linguistics: An analysis of sentence topics. *Philosophica* 27. 53-94.

- Rešetar, Milan. 1911. *Die serbokroatischen Kolonien Südditaliens: Mit 22 Abbildungen im Text*. Wien.
- Reumuth, Wolfgang, and Otto Winkelmann. 2001. *Praktische Grammatik der italienischen Sprache: Neubearbeitung*. Wilhelmsfeld.
- Roberts, Ian. 2016. Object clitics. In: *The Oxford guide to the Romance languages*. Ed. Adam Ledgeway and Martin Maiden. Oxford. 786–801.
- Saint-Exupéry, Antoine de. 2002. *The Little Prince*. Translated by Katherine Woods. London.
- Saint-Exupéry, Antoine de. 2009. *Mali Kraljič. Tradučan na-našu jiz frančezoga do Walter Breu aš Nicola Gliosca*. Neckarsteinach.
- Schick, Ivanka P. 1999. Doubling clitics and information structure in Modern Bulgarian. In: *Langue and parole in synchronic and diachronic perspective: selected proceedings of the XXXIst Annual Meeting of the Societas Linguistica Europaea, St. Andrews 1998*. Ed. Christopher Beedham. Amsterdam.
- Schroeder, Christoph. 2006. Articles and article systems in some areas of Europe. In: *Pragmatic organization of discourse in the languages of Europe*. Ed. Giuliano Bernini and Marcia L. Schwartz. Berlin. 545–611.
- Stalnaker, Robert. 1999. *Context and content*. Oxford.
- Steedman, Mark, and Ivana Kruijff-Korbayová. 2001. Two dimensions of information structure in relation to discourse structure and discourse semantics. In: *Information structure, discourse structure and discourse semantics. Papers of the 13th European summer school in logic, language and information*. Ed. Ivana Kruijff-Korbayová, and Mark Steedman. Helsinki. 1-10.
- Strawson, Peter Frederick. 1964. Identifying reference and truth values. *Theoria* 30. 377-394.
- Sussex, Roland, and Paul Cubberley. 2006. *The Slavic languages*. Cambridge.
- Tomić, Olga. 2008. Variation in clitic-doubling in South Slavic. In: *Microvariation in syntactic doubling*. Ed. Sjef Barbiers, Olaf Koenenman, Marika Lekakou and Margreet Ham. Leiden. 443-468.
- Trenkic, Danijela. 2004. Definiteness in Serbian/Croatian/Bosnian and some implications for the general structure of the nominal phrase. *Lingua* 114 (11). 1401–1427.
- von Heusinger, Klaus. 1999. *Intonation and information structure. The representation of focus in phonology and semantics*. Konstanz.
- von Heusinger, Klaus. 2002. Specificity and definiteness in sentence and discourse structure. *Journal of Semantics* 19. 245-274.
- Weiss, Daniel. 2004. The rise of an indefinite article: The case of Macedonian *eden*. In: *What makes grammaticalization. A look from its fringes and its components*. Ed. Walter Bisang, Nikolaus Himmelmann, and Björn Wiemer. Berlin. 139-168.

Opera Slavica Coloniensia

Herausgegeben von Daniel Bunčić und Jörg Schulte

Slavisches Institut der Universität zu Köln

ISSN 2748-6060

<https://slavistik.phil-fak.uni-koeln.de/osc>

https://kups.ub.uni-koeln.de/view/series/Opera_Slavica_Coloniensia/

- | | | |
|---------|------|---|
| Band 1 | 2015 | Anja Dillmann
„Wyśle ci maila“: Zur Belebtheitskategorie im Polnischen |
| Band 2 | 2015 | Marharyta Schödder
Der Akzent von Germanismen im Russischen |
| Band 3 | 2015 | Simone Maffezzoni
Das Rom der Dekabristen |
| Band 4 | 2016 | Iana Elger
Zur Stellung von Präfixen im russischen Wortbildungssystem |
| Band 5 | 2017 | Maxim Istomin
Von der Muse geküsst: Michelina in Brodskijs Werken und Leben |
| Band 6 | 2017 | Anastasia Romanuk
Die Entwicklung der nominalen Distanzanrede im Russischen:
Rückgang der Verwendung des Vor- und Vatersnamens
und Vordringen des Vornamens? |
| Band 7 | 2013 | Anastasia Smirnova
Der Genitiv der Negation im Russischen |
| Band 8 | 2016 | Tanja Keller
Wege der russischen Cicero-Rezeption |
| Band 9 | 2017 | Sabina Stacenko
„Contes des fées“ und „Contes du feu [...]“
Das Genre des Märchens und A. S. Puškins „Povesti pokojnogo Ivana
Petroviča Belkina“ |
| Band 10 | 2018 | Michael Beckers
Titus Petronius und die slavischen Dichter |
| Band 11 | 2019 | Giada Hitthaler
P. A. Stolypin and the renaissance of rhetoric |
| Band 12 | 2019 | Petar Pejović
Kyrillisches und lateinisches Alphabet
in serbischsprachigen Linguistic Landscapes |
| Band 13 | 2019 | Tamara Tanasijević
M. Ju. Lermontov and the science of passions |
| Band 14 | 2020 | Jelena Apostolović
Srpska satirična uspavanka u kontekstu evropske književnosti |
| Band 15 | 2020 | Michael Beckers
Die verlorenen Melodien slavischer romantischer Lyrik |

- | | | |
|---------|------|---|
| Band 16 | 2020 | Isolde Ruhdorfer
Akzent beim Schreiben?
Untersuchung zum deutschen Schriftakzent in der russischen Kyrillica |
| Band 17 | 2020 | Željana Vukanac
The history of the rondeau between East and West |
| Band 18 | 2021 | Ivana Dimitrijević
Von der Chovevet Tsion zur Chalutsa
Eine kulturgeschichtliche Studie (1897–1931) |
| Band 19 | 2021 | Joana Puci
Die ersten serbischen und griechischen Zeitschriften
Kulturelle und sprachliche Zusammenhänge der Aufklärung |
| Band 20 | 2021 | Antonia Shevchenko
Die vorwärtsgerichtete Diskursfunktion der
Genitiv-Akkusativ-Opposition bei Verneinung im Russischen |
| Band 21 | 2014 | Julia Bensch
Korrosion des Russischen bei MigrantInnen in Deutschland |
| Band 22 | 2016 | Lilit Kazaryan
Sentimentalismus auf Russisch und Französisch:
Leben und Werk der Fürstin Zinaida Volkonskaja |
| Band 23 | 2017 | Paul Compensis
Definiteness in Molise Slavic |