



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis
Focus and Topic in Meadow Mari

verfasst von / submitted by
Johannes Hirvonen

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Arts (MA)

Wien, 2023 / Vienna 2023

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on the
student record sheet:

UA 066 867

Studienrichtung lt. Studienblatt /
degree programme as it appears on the stu-
dent record sheet:

Allgemeine Linguistik: Grammatiktheorie und
kognitive Sprachwissenschaft (Master)

Betreut von / Supervisor:

Univ.-Prof. Dr. Daniel Büring

Contents

1	Introduction	5
2	Meadow Mari, a Uralic language at the Volga bend	6
2.1	Language ecology	6
2.2	Some remarks on Mari word order	7
2.3	Previous research	9
3	Information structure	10
3.1	General remarks	10
3.2	Topics	13
3.2.1	General considerations	13
3.2.2	Types of topics	14
3.2.3	Diagnosing topics	18
3.3	Focus	20
3.3.1	General considerations	20
3.3.2	Types of focus	22
3.3.3	Diagnosing focus	23
3.4	Contrastive topics	27
4	Strategies of information structure marking	29
4.1	Strategies of focus marking cross-linguistically	29
4.2	Strategies of topic marking cross-linguistically	33
4.3	Information structure marking in the Uralic languages	34
4.4	Summary	37
5	Method	38
6	Focus in Mari	42
6.1	Information focus	43
6.1.1	Constituent focus	44
6.1.2	Full clausal focus	46
6.1.3	Presentational focus	47
6.1.4	Correction focus	48

6.1.5	Verbal focus and verum focus	50
6.1.6	Focus in polar questions	51
6.2	Identificational focus	54
6.3	Focus negation	55
6.4	Summary	58
7	Topics in Mari	59
7.1	The left periphery	59
7.1.1	Non-subjects as topics	60
7.1.2	Antecedents for null subjects	64
7.1.3	Adverbials in the left periphery	66
7.1.4	Topic marking phrases	68
7.1.5	Left dislocation	71
7.1.6	Summary	72
7.2	Postverbal elements	73
7.2.1	Afterthoughts and completions	73
7.2.2	Familiar topics	74
7.2.3	Altruistic postverbal elements	75
7.2.4	Heavy XP shift	78
7.2.5	Genre-specific usage	80
7.2.6	Summary	81
8	The 3SG Possessive Suffix žE as Marker of Contrast and Topic Shift	84
8.1	Introduction	84
8.2	Previous approaches	88
8.3	Contrastive topics	90
8.4	žE as marker of topic shift	94
8.4.1	Interim summary – properties of žE	106
8.5	Other functions of žE and possible contact influence	109
8.6	Summary	116
9	Focus and Topic in Meadow Mari – Summary	116

Abbreviations

ACC = accusative	INS = instrumental
ADD = additive clitic	LAT = lative
ADJ = adjectivizing suffix	MSC = masculine
AUX = auxiliary verb	NEC = necessive
CAUS = causative	NEG = negation verb or constituent negation
CL = Tzotzil 2nd position clitic	NEG.EXIST = negative existential predicate
CMPR = comparative	NMLZ = nominalizer
CNG = connegative verb form	NOM = nominative
COMP = comparative case	PFX = verbal prefix
CVB = converb	PL = plural
CVB.ANT = converb of anterior action ('after doing X')	PREP = prepositional case
CVB.NEG = negative converb ('without doing X')	PST = past tense
CVB.POST = converb of posterior action ('before doing X')	PST1 = first past tense
DAT = dative	PST2 = second past tense
DET = determiner	PTCP.ACT = active participle
DIM = diminutive	PTCP.PASS = passive participle
EMPH = emphatic clitic	PTCL = particle
ENC = Tzotzil intonational phrase enclitic	PX = possessive suffix
EXIST = existential copula	Q = polar question marker
FEM = feminine	REFL = reflexive
GEN = genitive	SG = singular
ILL = illative	SUP = superlative
IMP = imperative	TO = Russian discourse particle <i>to</i>
INDEF = indefinite prefix	TOP = topic marker
INE = inessive	VOC = vocative
INF = infinitive	ŽE = Mari 3SG possessive suffix or Russian discourse particle <i>že</i>
INF.FUT = future infinitive	1–3 = 1.–3. person

Persönliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig verfasst habe. Ich habe alle Quellen, die ich für die Verfassung der Arbeit benützt habe, nach den Regeln guter wissenschaftlicher Praxis zitiert. Ich habe diese Arbeit bisher nicht als Prüfungsarbeit vorlegt.

Johannes Hirvonen, Wien, November 2023

1 Introduction

Meadow Mari is a Uralic language spoken in western Russia. As a scrambling SOV language, word order variation is widespread and typical. Although it has been recognized in old and recent descriptions of Meadow Mari clause structure that information structure is an important factor governing this variation in word order, Meadow Mari information structure marking has not been systematically described so far. The following quotes from recent publications stress the actuality of this issue:

“Deviating from SOV order is possible and obviously conditioned by information structure, but these phenomena have not been researched in detail.” (Saarinen 2022: 463)

“Information structure has received relatively little attention in the Mari literature, and is a significant topic for further investigation.” (Nelson, Vedernikova, and Bradley to appear: 7)

“The information structure of Finno-Ugric languages, apart from Hungarian and Finnish, is still very much under-studied.” (Surányi et al. to appear: 21)

The aim of this thesis is to fill this gap and provide a first descriptive overview about the strategies of information structure marking in Meadow Mari. At the center of this investigation are types of morphosyntactic marking in Meadow Mari which interact with the information structural notions of focus, topic, and contrastive topic. In particular, I describe different strategies of expressing focus, provide an overview about the effects of (non-)topicality on word order, and about different functions of the postverbal field in Mari. Finally, I discuss data on the discursive functions of the 3SG possessive suffix *-žE* and argue that it can be employed as marker of shift topics, a function of the suffix which was previously undescribed. The thesis is intended as starting point for further, more detailed investigations of Meadow Mari information structure.

This thesis is structured as follows: In Section 2, I introduce the Meadow Mari language. Section 3 deals with information structure in general terms, and there I

define the information structural categories relevant for this thesis – topics, focus, and contrastive topic. In Section 4 I discuss some tendencies of how these categories are marked in different languages of the world, including the Uralic languages, as a backdrop for the discussion about Mari. Section 5 shortly describes the method of data collection. The main part of the thesis starts with Section 6 where I discuss the marking of various types of foci. Section 7 continues the main part with a description of topics in Mari. I discuss the role of the left periphery as position for topics, and conversely, the reasons for why constituents may appear after the verb, including non-topicality. The results in this section about topicality set the context for the final Section 8 where I show that the suffix *žE* can, in addition to its contrastive function, mark shift topics in Mari, a function previously undescribed. In Section 9 I summarize the main findings of the thesis.

A few notes on language examples: All examples are in Standard Meadow Mari and English unless indicated otherwise. All language examples have been transliterated into a latin based orthography if the written standard is non-latin based (e.g. Mari and Russian). In the Standard Meadow Mari examples, the letter *a* stands for the central vowel which is usually transcribed as *â* in Finno-Ugric transcription. In Hill Mari examples, the palatal harmonic variant of the central vowel is transcribed as *ä*, the velar variant as *ə*. The transcription of the dialectal Mari examples is taken directly from the source of the examples. Focussed constituents and contrastive topics are marked with SMALL CAPS in the language example and the corresponding constituents in the translation. Underlined words mark constituents in a clause before they become sentence topics. Other highlighting is done by **boldfacing**.

2 Meadow Mari, a Uralic language at the Volga bend

2.1 Language ecology

Mari (in older sources Cheremis) is a pluricentric Uralic language that is spoken in European Russia. Most of its speakers live in the titular Russian republic Mari El but sizable speaker communities exist in the Republic of Bashkortostan to the east of Mari El as well. The two standardized literary varieties are the smaller

Hill Mari in the western part of the republic and Meadow Mari in the rest. In 2020, 15 352 people self-identified as speakers of Hill Mari and 258 722 people as speakers of Meadow Mari (*Perepis'* 2020–2021) which makes Meadow Mari the dominant variety. For recent overviews about Mari see Bradley and Luutonen (2023) and Saarinen (2022). In what follows, the term “Mari” will be used for Meadow Mari, the variety under investigation.

Despite the comparatively large number of speakers and good legal status as official language of the republic (*gosudarstvenny yazyk*), Meadow Mari is under threat since it lacks institutional support, e.g. as language used in education (Zamyatin 2016: 267–271; Bartholomä and Schötschel 2012), and the number of speakers has been declining in the last decades (Saarinen 2008). Hill Mari is considered “severely endangered”, and Meadow Mari is considered “definitely endangered” (Moseley 2010). Today, virtually all speakers of Mari have at least some knowledge of Russian, and without a doubt, the most important contact language for Mari today is Russian. Over the centuries, Mari furthermore has been in intensive contact with Turkic languages as well, most notably Chuvash and its predecessor Bolgar Turkic, and the Kipchak Turkic languages Tatar and Bashkir, leaving traces in the lexicon (e.g. Räsänen 1920, 1923; Saarinen 1997) and morphosyntax of Mari (e.g. Bereczki 1984; Hesselbäck 2005).

2.2 Some remarks on Mari word order

Mari is a relatively rigid head-final language: nouns are preceded by adjectives, compounds are head-final, the language employs postpositions, lexical verbs precede finite modal verbs, and clausal complements precede the main clause. The neutral word order is SOV with the dative object preceding the accusative object; time and place adverbials may be placed in front of subjects or topics, with time adverbials preceding place adverbials (see example 1; Chkhaidze 1941: 29–31; Riese et al. 2019: 419; Pregla 2023).

(1) (Chkhaidze 1941: 30)

TIME-ADV – PLACE-ADV – SUBJ – IO – DO – CIRCUMSTANCE-ADV – V

(2) (Chkhaidze 1941: 30)

Tače kolxoz pasu-što vič rveze traktor dene mlandə-m
today kolkhoz field-INE five lad tractor with land-ACC
udarno kural-ət.
enthusiastically plough-3PL

‘Today on the kolkhoz field, five lads enthusiastically plough the land with a tractor.’

The constituents of the above sentence may be scrambled rather freely, and this has an effect on the interpretation of the clause. For example, “with a tractor” may be ‘stressed’, according to Chkhaidze (1941) if it is placed directly before the verb:

(3) (Chkhaidze 1941: 31)

*Tače kolxoz pasu-što vič rveze mlandə-m **traktor dene***
today kolkhoz field-INE five lad land-ACC tractor with
udarno kural-ət.
enthusiastically plough-3PL

‘Today on the kolkhoz field, five lads enthusiastically plough the land **with a tractor**.’

Mari word order is thus “free” in the sense that a number of different word orders are allowed to occur under certain circumstances. Vilks (2022: 950) characterizes Mari as “quite flexible” SOV language, and most older work on Mari syntax stresses the fact that deviations from the unmarked word order are used to stress or emphasize certain constituents (e.g. Chkhaidze 1941: 29–31 Pengitov 1961: 122, Yegorova 1981: 275–276), although it is often not quite clear what exactly the authors mean by emphasis. The frequently used term *logicheskoye udareniye* ‘logical stress’ (e.g. Pengitov 1961: 117) in older Russian scholarly work likely refers to constituents that fulfill a focus-like function (Asztalos 2020: 18). The main factors influencing Mari word order are therefore to be sought in information structure (cf. Saarinen 2022: 463). The thesis is about the investigation of these factors.

In colloquial varieties of Mari, the word order can, however, deviate from

SOV, apparently without information structural reasons. Examples like ex. (4) have been adduced as evidence to argue that Mari is slowly shifting from a head-final to a head-initial syntactic orientation because complements follow their heads instead of preceding them in those examples (Anduganov 2001). Example (5) is especially illustrative; both the VP *lajkəm šəndaš* and the PP *fotosüret jəmake* follow the verb *mondaš*. However, constituents in a postverbal position are not unusual in some contexts in the literary variety, either (see Section 7.2).

(4) (Anduganov 2001: 74)

Čəla korno kond-a tide pajrem-aš.

all road lead-3SG this party-ILL

‘All roads lead to this party.’

(5) CSMM, VKontakte

[...] *da tə godəm-ak mond-aš ogəl lajk-əm šənd-aš fotosüret*

and this during-EMPH forget-INF NEG like-ACC put-INF picture

jəmake [...]

to.under

‘[...] and at the same time one should not forget to like the picture [...]’

Summarizing, we find that the unmarked word order in a Mari clause is SOV, but deviations from that order are possible in colloquial registers.

2.3 Previous research

In comparison to smaller Uralic languages, Mari is relatively well researched. Apart from several descriptive grammars and grammatical descriptions (Alhoniemi 1988, 1993, 2010; Bradley and Luutonen 2023; Chkhaidze 1941; Kangasmaa-Minn 1998; Riese et al. 2019; Saarinen 2022; Timofeyeva 1961; Tuzharov 1987), there are newer studies about specific issues in the realm of morphosyntax such as: the syntax of the noun phrase (Simonenko and Leont’ev 2012), the ordering of noun morphology (den Dikken 2018; Guseva and Weisser 2018; Hammer and Bradley 2022; Luutonen 1997; McFadden 2004: Section 7.4), the syntax of postpositional

phrases (Burukina 2023b), verb clusters (Georgieva, Salzmann, and Weisser 2021), possession and the semantics of possessive suffixes (Georgieva 2022a,b; A. Simonenko 2014), existential possession (A. Simonenko 2020), reflexives (Volkova 2014, 2017a), non-finite clauses (Pomozi 1997; Volkova 2017b), clausal complementation (Burukina 2023a), null subjects (Bradley and Hirvonen 2022), dative subjects (Burukina 2022), negation and negative concord (Hirvonen and Bradley to appear; Mitchell 2006; Saarinen 2015), various sentential construction models (Vasikova 1982), the expression of topic and comment (Kazyro 2006) and works on general syntax (Nelson, Vedernikova, and Bradley to appear; Pengitov 1961; Vilkuina 1998; Yegorova 1981). The most detailed description of focus and postverbal elements in Mari is found in Pregla (2023). Thus, despite being well-researched regarding these topics, information structure in Mari has not attracted a lot of attention so far. With this thesis I hope to lay the ground for fruitful future research on Mari information structure.

3 Information structure

This section deals with information structure in general terms. After a general introduction in Section 3.1, Topics are discussed in Section 3.2, Focus in Section 3.3, and Contrastive Topics in Section 3.4.

3.1 General remarks

A property of human language enabling successful communication is the constant optimization of communicative means in a discourse. The actual information contained in an utterance may be structured, or, in the words of (Chafe 1976), be *packaged* in various ways to conform to the temporary communicative needs of interlocutors. Information packaging is concerned primarily with the question *how* information is transmitted, not with *what* is transmitted. To package information, the speaker assesses on the basis of the temporary state of the addressee's mind how the addressee is able to process what is being said. This state of the interlocutors' knowledge can be modelled as Common Ground. The Common Ground contains a set of mutually accepted propositions and entities introduced

into the Common Ground, and additionally, the Common Ground contains information about the communicative intents of the interlocutors. The Common Ground changes continuously in the course of conversation, and therefore information packaging is adapted to the utterance (Krifka 2008: 243–246). This is the basis for an established understanding of *Information Structure*: “[n]formation] S[tructure] is concerned with those features of language that affect the local C[ommon] G[round]” (Krifka 2008: 243).

Information structure connects linguistic domains such as phonology, morphology, and syntax with general cognitive faculties which are responsible for “fixation of belief by way of information update, pragmatic reasoning, and general inference processes” (Zimmermann and Féry 2009: 1). While information structure is most likely a phenomenon universal to language, its linguistic coding is subject to cross-linguistic variation. Languages make use of a large variety of strategies for the coding of different information structural categories, including prosodic strategies, morphological strategies, or syntactic strategies. This means that information structural notions are linguistically significant since they are reflected in actual linguistic coding (Zimmermann and Féry 2009).

As an example for information structure at work, consider a speaker uttering the sentence in example (6a). The first clause introduces the fact that the speaker has a cat, into the Common Ground. In the second clause, the speaker can refer to said cat by using *my cat*, since they can assume that the cat was present in the common ground after it was introduced in the first clause. The second clause then adds further information, showing that the common ground is constantly changing. Information has to be packaged in an adequate manner relative to the common ground content, and violations of this principle can result in odd or infelicitous discourse. Consider the rather odd example (6b), where the order of the sentences has been reversed. Here, the phrase *my cat* in the first clause already introduces the cat into the common ground, but the second clause introduces the same information again (Krifka and Musan 2012: 1–2).

(6) (Krifka and Musan 2012: 1–2)

- a. I have a cat, and I had to bring my cat to the vet.
- b. #I had to bring my cat to the vet, and I have a cat.

Information packaging is, however, not only used to present information in an optimal manner relative to the common ground content. Krifka (2008: 244) and Krifka and Musan (2012) point out that some aspects of message optimization do in fact affect the *content* of a message itself and cannot be considered pure packaging, for example by leading to truth-conditional differences. Classic examples for this are the interaction of modals and focus-sensitive particles with the placement of sentential stress in English. Sentential stress can be used for simple information packaging such as in answers to *wh*-questions, but its placement can also lead to truth-conditional differences.

The case with modals is illustrated by a joke in Halliday (1967: 38). At the escalators in the London underground, there is a sign which says “Dogs must be carried”. A man reading the sign got very worried because he had no dog in the underground with him. This resulted from a reading of the sign which wrongly stressed *dogs* even though the order on the sign is felicitous in the situation of the man reading it, contrary to the sequence of sentences in example (6b). Consider the difference between the two readings in example (7). In the context of the sign and our world knowledge only the prosodic realization in example (7a) makes sense. This reading is something along the lines of “If dogs are taken onto the escalator, they have to be carried”, whereas example (7b) would mean “In order to use the escalator, a dog must be carried into it”. This is not true since the escalator may naturally be used without a dog. In this case, sentential stress can impact the truth conditions of the order and cannot be considered information packaging which is concerned with *how* information is transmitted. Information structure can therefore not only impact how information is managed in the common ground, but impact the actual content and truth conditions of an utterance.

(7) (Halliday 1967: 38)

- a. Dogs must be **CARRIED**.
- b. #**DOGS** must be carried.

Following this very general introduction into the matter I will introduce the three information structural categories which are under discussion in the main part about Mari information structure: Topics, Foci, and Contrastive Topics.

3.2 Topics

3.2.1 General considerations

Information is not added to the Common Ground in an unstructured way. Rather, speakers attempt to organize the information and give hints about how information is to be classified. This is the notion of Topic:

“Quite often in making an utterance, a speaker in some way brings our attention to an entity that is relevant at that point in the discussion, in order to tell us something about it. The relevant entity may be an individual or it may be a situation or event. In any such case, we say that the entity to which our attention is drawn is the Topic of the utterance.” (Roberts 2019: 381).

Drawing attention to certain parts of the discourse is necessary for structured discourse as the content of the Common Ground is possibly a large amount of information. For this reason Reinhart (1981) argues that the discourse context is not an unordered set of propositions but structured into subsets of propositions, which are stored under defining entries; for example, propositions about individuals. In example (8a), the new information provided by the sentence is stored as information *about* Aristotle Onassis, and in example (8b) it is stored as information *about* Jacqueline Kennedy.

(8) (Krifka 2008: 265)

- a. Aristotle Onassis married Jacqueline Kennedy.
- b. Jacqueline Kennedy married Aristotle Onassis.

The difference in word order does not have a semantic, truth conditional effect, but effects the context in which the sentence is felicitous, and the “verification strategies” in a given discourse (Reinhart 1981: 60; Büring 2016: 81). *Topic* is therefore a pragmatic notion relative to discourse.

The most influential approach to topics, Reinhart (1981), defines topics as the part that a given sentence is “about”, e.g. an individual, as illustrated above. Such topics are usually called Aboutness Topics. A common diagnostic for the

property of the sentence being “about” one of its parts is the “aboutness test”, as employed for English in the following examples. The main purpose of the test is to see whether adding *about* to a given constituent preserves the sense of the original sentence. If so, the sentence can be understood to actually be “about” the constituent, or that the constituent is “under discussion”, which makes it a topic. The “as for test” in English is similar but implicates a contrast between *John* and some other entity (Roberts 2019: 390–392) or to change the current topic (Reinhart 1981: 64).

(9) (Erteshik-Shir 2007: 19–20)

- a. Tell me about John. – He’s very nice.
- b. As for John, he’s very nice.

The notion of Topic is rather contested, however, and there are only few properties different authors agree on (Roberts 2019). *Aboutness* is a rather vague notion as well, and even the *about*-tests can lead to different results in a single language, or test for different types in different languages.

3.2.2 Types of topics

Reinhart (1981: 64) among others noted that different topic marking devices can be employed for different purposes; in English, left dislocation (e.g. via the *as for* test) is used to change the topic of conversation and introduce a new one, in contrast to non-dislocated topic marking (e.g. with the *about* test), that does not. That is, different means of syntactic, morphological or prosodic marking that are commonly classified as topic marking, can mark different kinds of topics in a single language (Frascarelli and Hinterhölzl 2007: 88).

Cross-linguistically, there are some commonalities concerning the relation between the position of a topic and its interpretation. In a typological survey of information structuring in various languages, Givón (1983: 9) discusses three types of topics that he defines both by the pragmatic relation they code, and their typical positional realization:

1. Chain-initial topics are characteristically newly-introduced, newly-changed

or newly-returned topics, they are often continued as topics in the later discourse context.

2. Chain-final topics are typically continuous topics in terms of the preceding discourse context, but they are non-persistent in terms of the succeeding discourse context.
3. Chain-medial topics combine these properties and characteristically continue a topic, but they are not as persistent as chain-initial topics.

Similar classifications have been made for example by Erteshik-Shir (2007), Frascarelli (2018), and Frascarelli and Hinterhölzl (2007), and while these proposals differ in their details, all these scholars usually make a distinction between some form of continued (or familiar, backgrounded) topics and shifted (or switch) topics which roughly correspond to type (2) and type (1), respectively, in Givón's classification. For example, Frascarelli (2018) distinguishes *A-Topics* which combine *aboutness* with the property of being newly introduced or reintroduced and changed to, from *G-Topics* which refer to given information. G-Topics can either be used for topic maintenance with respect to a newly introduced A-Topic, or as background topics.

The following two examples illustrate shifted topics and familiar topics in discourse. Example (10) from Tzotzil exhibits a shift topic: In the first clause, a girl from San Andres is introduced in the discourse as focus. In the second clause, the narration turns to the girl which is now marked with a topic marker *a*, a determiner *ti*, and an enclitic *-e*. Aissen (1992: 51) argues that Tzotzil topics marked by the topic marker and the enclitic are always new or shifted topics, but cannot refer back to given topics.

(10) Tzotzil (Aissen 1992: 50–51)

Something had landed at the foot of the tree, they went to look. There was a straw mat. ‘Hell, what could it be? Come on, let’s untie the straw mat!’ the two men said to each other. They untied it. You know what? –

TZEB SAN-ANTREX la te staik un. A ti tzeb san-antrex
 girl San.Andres CL there they.found ENC TOP DET girl San.Andres

un-e, iyik’ik la ech’el un.

ENC-ENC they.took CL away ENC

‘It was a GIRL FROM SAN ANDRES that they found there. They took **the San Andres girl** with them.’

As example for a familiar topic consider the following interview sequence from Italian in example (11).¹ Speaker B in their turn introduces *la conferma* ‘the confirmation’ which is resumed by speaker A. Finally, speaker B repeats *la conferma* in a right-peripheral position. Frascarelli and Hinterhölzl (2007: 93) show that in this last repetition the constituent is realized with a L* tone and separate from the intonational phrase of the preceding sentence. They argue that the tonal realization is characteristic only of familiar topics in Italian, and that shift topics and contrastive topics are realized differently from familiar topics.

¹Translation and gloss by me.

(11) Italian (Frascarelli and Hinterhölzl 2007: 93)

B. *io dovevo studiare le regole qui e lì fare*
 1SG must.PST.1SG study.INF DET rule.PL here and there do.INF
solo esercizio, invece mi aspettavo di trovare dei
 only exercise while 1SG.DAT expect.PST.1SG to find.INF DET
punti a cui far riferimento ogni volta per vedere
 point.PL at which do.INF reference all.PL time for see.INF
la regola, questo mi è mancato praticamente per
 DET rule this 1SG.DAT AUX lack.PTCP practically for
avere la conferma di ricordare tutto insomma
 have.INF DET confirmation to remember.INF all in.short

‘I was supposed to study rules here and to the exercises there, while I expected to find some outline I could refer to, at any point, to check the relevant rule, this is what was lacking for me, to have the **confirmation** that I could remember everything.’

A. *comunque quelle domande ti davano la*
 however DET.PL question.PL 2SG.DAT give.PST.3PL DET
conferma che avevi capito
 confirmation that AUX.PST.2SG understand.PTCP

‘However those questions gave you the **confirmation** that you had understood.’

B. *ma ... magari non me la- non riesco a*
 but maybe NEG 1SG.ACC DEM NEG manage.1SG to
darmela da sola la conferma
 give.1SG.DAT.3SG.ACC alone DET confirmation

‘Well, maybe I cannot... I cannot give myself the **confirmation** on my own’

Besides the difference between shift and familiar topics in various languages, the interpretation of topics can differ in smaller details as well: A case in point,

Roberts (2019) discusses various differences between Japanese and Korean topic marking. While the languages are similar as regards most functions and distribution of the topic marking morphemes, they differ in sensitivity to the givenness of the constituent they mark. This manifests for example in the fact that the Korean topic marker *-nun* is acceptable with (non-specific) indefinites like ‘some man’, but the Japanese topic marker *-wa* is not (Roberts 2019: 398–400). Ultimately, this raises the question what topics have in common at all. Some scholars argue that the notion of topic is too vague altogether. Roberts (2019: 393) states: “So though there may be many common factors, in the end it appears that what these various conventions realize is a family of closely related notions, rather than a syntactic or pragmatic universal”. This of course effects the method for investigation of topics in a given language.

3.2.3 Diagnosing topics

The most common test for topics is the *about* test, as discussed above. Other proposed tests include the interpretation of pronouns following multiple possible antecedents (Reinhart 1981: 70–71). Furthermore, stories with figures recurring in different roles have been stated to be apt tools for the documentation of topic maintenance and change in a given language (Aissen 2015).

We saw above that topics can “do different things”, and that different tests can target these different kinds of topics: Both the *as for* test and the *about* test target topics in English, but “as for” implicates a contrast between the topic and other entities whereas “about” does not. That is, even in a single language, similar tests for topicality can target slightly different phenomena. This issue is exacerbated in cross-linguistic investigation when it is not even clear from the outset whether a test like the *about* test can be applied to a given language at all, what types of topics it targets, and what properties topics in a given language have at all (cf. Roberts 2019: 390–392). Due to these issues, Büring (2016: 83) suggests: “An investigation of a suspected topic-marking in a given language should therefore start, not by assuming that, since it is a topic, it has a particular cluster of properties, but by establishing their exact specific properties independently, among them:”

- what items can be so marked
- whether the marking can
 - establish a new discourse referent as the aboutee² for the following, or
 - establish an existing discourse referent as the new aboutee [...], or
 - refer to an established aboutee throughout its tenure as ‘what the passage is about’
 - ... or do something altogether different
- whether elements that meet that description have to be so marked, or merely may be,
- whether the same marking can serve other pragmatic functions
- whether there are other tests (than occurrence with that marking) to establish the status marked by it

Ultimately these considerations mean that for every language different tests for topicality may be more apt than others, and that possible topic marking strategies have to be tested for various factors before concluding that the given strategy can indeed mark topics. In Section 7.1, I argue that the sentence-initial position is a position for topics in Mari by showing that placing a constituent in the sentence-initial position has effects which are connected to topicality: non-subjects can occupy a left-peripheral position in a topic progression; placing a constituent in the clause-initial position increases its acceptability as antecedent for a following null subject; the sentence topic can be changed with the help of a specific topic-marking phrase, *manmašte*, in the left periphery; new topics can be established via left dislocation. The analysis of the Mari suffix ŽE in Section 8 is more detailed and will proceed closer along Buring’s questions.

²An individual introduced into the discourse context and being currently ‘under discussion’ (Buring 2016).

3.3 Focus

3.3.1 General considerations

The second information structural notion central to this thesis is focus. One definition of focus is provided by Krifka (2008: 248):

- (12) A property F of an expression α is a focus property iff F signals (a) that alternatives of (parts of) the expression α or (b) alternatives of the denotation of (parts of) α are relevant for the interpretation of α .

The (a) part of the above definition is concerned with *expression* focus affecting for example pronunciation or the choice of words. This type of focus will not be dealt with in this thesis. The case in (b) is called *denotation* focus by Krifka, and in this case focus introduces alternatives on the level of denotations which play a role in the interpretation of the constituent containing the focussed expression. This is a definition in the spirit of Alternative Semantics (Rooth 1985). Its central claim is that focus indicates the presence of alternatives relevant for the interpretation of linguistic expressions. Thus, a focused constituent is chosen from a set of alternatives (Krifka 2008).

The most widespread test for focus is the answer to a *wh*-question (cf. van der Wal 2016: 265) and the alternative semantic treatment of focus can easily be explained with such an example. In example (13) there is a question-answer pair. The correspondence to the *wh*-word *who* in the question, is the noun *Peter* in the answer. The phrase that replaces the *wh*-element is in focus because it indicates that there are other relevant alternative cookie stealers, from among which the referent of *Peter* is chosen:

- (13) (Krifka 2008: 250)
Who stole the cookie? – PETER stole the cookie.

The *wh*-diagnostic is somewhat independent of the exact definition of focus. An alternative to the alternative semantic treatment, Lambrecht (1994: 213) defines focus as “the semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition”, i.e. what is new in a sentence. This is also the case in example (13) above. What is important for

the purposes of this thesis is that the same *wh*-question diagnostic is compatible with different definitions of focus. A working approach to focus can even make do without reference to a certain definition of focus at all by just using some widespread diagnostics. For example, Büring (2009: 178) simply states that “certain (semantic or pragmatic properties of certain) contexts systematically trigger focus”. Büring remarks that there is no generally agreed on definition on these contexts, however, and indeed a large number of different tests and diagnostics have been employed to test for different types of focus. For this reason, the discussion about “establishing which linguistic strategies a language employs to express focus and what sort of focus these strategies express” (van der Wal 2016: 259) is a central issue in research on information structure. Since the purpose of this thesis is descriptively oriented, I will limit my attention to different proposed diagnostics for focus, and will not discuss the merits of one approach to focus over the other. Before discussing tests and diagnostics for focus, I will clarify what types of focus can actually be tested this way. Focus types can be differentiated by their size and by their type.

Focus is associated with a focus domain which scopes either over the whole clause, a predicate, or an argument. Syntactically, these domains approximately map to the clause, the verb phrase, and an argument of the verb phrase, respectively. Thus, foci can be of different sizes. Using the *wh*-diagnostic, this can easily be shown. As stated above, it is a common assumption that the part of an answer corresponding to the *wh*-word in the question is in focus. In example (14), the same answer is presented in contexts where the *wh*-word targets different parts of the sentence and the focus is of a different size in each answer. So in ex. (14a), the *wh*-word has as its target the VP, and the whole VP is in focus. The *wh*-word in example (14b), on the other hand, targets only the object, so in the answer, the object is in focus, and so on. The example also shows that the focus of the a sentence is largely dependent on the context it is uttered in.

(14) (Erteshik-Shir 2007: 28)

- a. What did John do? – He WASHED THE DISHES.
- b. What did John wash? – He washed THE DISHES.

- c. Who washed the dishes? – JOHN washed them.
- d. What happened to the dishes? – JOHN WASHED THEM.
- e. What happened? – JOHN WASHED THE DISHES.
- f. What did John do with the dishes? – He WASHED them.

Besides the different sizes of foci, foci may also be differentiated by the sentential role of the focus marked constituent. This is not a question related to their interpretation: an object focus is interpreted as having relevant alternatives to the object just like a subject focus is interpreted as having relevant alternatives to the subject. However, different focus marking strategies may be involved in marking different sizes or grammatical roles of foci, so the differentiation is relevant for grammatical descriptions of foci. For example, in SOV languages with a pre-verbal focus position, the object occupies the preverbal position in both a verbal focus word order and in the unmarked word order, whereas the subject does not. In this case, subject and object focussing necessarily involve different strategies.

3.3.2 Types of focus

Besides their size (or sentential role), foci can also be classified by their type, that is, the overall semantic-pragmatic effect that is evoked by a certain focus-marking strategy. The main effect of focus (according to Alternative Semantics) is the indication of a relevant alternative, and other types of focus are the outcomes of additional pragmatic and semantic factors, including mirativity, unexpectedness, and others (see van der Wal 2016: 262–264 for an overview). With respect to the functions of the Common Ground – on the one hand containing a set of shared propositions and entities, and on the other hand containing information about communicative intents – Krifka (2008) and Krifka and Musan (2012) distinguish two different uses of focus, namely semantic and pragmatic uses of focus. Semantic usage of focus relates to the content of the Common Ground and thus to the factual information, affecting truth conditions, and pragmatic usage of focus relates to the communicative intents of speakers. Pragmatic focus is used to guide communication along these intents. For example, a *wh*-question indicates

that the communicative goal of the questioner is to inquire about the constituent corresponding to the *wh*-constituent.

In an influential paper, É. Kiss (1998) distinguishes two types of focus that correspond to the two uses, identificational focus (also called contrastive focus, narrow focus) and information focus (also called wide focus):

(15) (É. Kiss 1998: 245)

An identificational focus represents a subset of the set of contextually or situationally given elements for which the predicate phrase can potentially hold; it is identified as the exhaustive subset of this set for which the predicate phrase actually holds.

(16) (É. Kiss 1998: 246)

If a sentence part conveys new, nonpresupposed information marked by one or more pitch accents – without expressing exhaustive identification performed on a set of contextually or situationally given entities, it is not an identificational focus but a mere information focus.

They differ, among others, in the following ways: (1) The identificational focus expresses exhaustive identification; information focus merely marks the nonpresupposed nature of the information it carries, (2) certain types of constituents, universal quantifiers, *also*-phrases, and *even*-phrases, for example, cannot function as identificational foci; but the type of constituents that can function as information focus is not restricted, (3) the identificational focus does, information focus does not, take scope (É. Kiss 1998: 248). Importantly, identificational focus affects the truth conditions of a sentence – the predicate holds only of the exhaustive set identified by the focus, and not of other elements (cf. Krifka 2008). É. Kiss (1998) argues that these two types of focus are fundamentally different and are expressed by different strategies in some languages, including Hungarian. Her diagnostics are shortly discussed in the next section.

3.3.3 Diagnosing focus

In example (14) above, different questions elicited foci of different sizes in the same clause. As a counterpart to that example and introduction to the diagnostics-

section, consider ex. (17). Here, all sentences (17a-f) elicit the utterance (18) where the direct object is focussed. The same happens when (18) in turn is continued by the utterance in (19) (Büring 2009: 178). That is, the focus is always the same size (and in the same sentential role), but it appears in different contexts: as answer to a *wh*-question, to a cleft question, to an alternative question, as correction to a polar question, as second part of the answer to an implicit multiple *wh*-question (*Who puts what in his pasta sauce?*), as contrast in juxtaposition, and as correction.

(17) (Büring 2009: 178–179)

- a. What do you put in your pasta sauce?
- b. What was it that you put in your pasta sauce?
- c. Do you put tarragon or thyme in your pasta sauce?
- d. Do you put tarragon in your pasta sauce?
- e. Daniel puts tarragon in his pasta sauce.
- f. First, I put tarragon in my pasta sauce, then . . .

(18) I put THYME in my pasta sauce.

(19) . . . not tarragon.

Wh-questions can be used to elicit focus of different sizes, such as shown in example (14) above (van der Wal 2016: 265–269). Alternative questions are used to select a phrase from among given alternatives. (*Do you want tea or coffee? – (I want) tea.*) The selection of one alternative over the other excludes the other alternative. Correction works similarly; if a statement is corrected, the contrastive part is in focus. The false constituent is contrasted with the alternative true constituent. One alternative is excluded and the strategy is thus compatible with exclusive focus. Lastly, since the meaning of focus is dependent on the immediate utterance context, the context can be used to create scenarios which are compatible only with certain types of focus. For example, phrases like ‘among others’ or ‘for example’ entail that there are other viable alternatives to a referent and thus are not compatible with an exhaustive focus strategy. Similarly, an

incomplete statement can be corrected with a negative reply, negating the exhaustive aspect of the statement. The answer presents the larger set to which the predicate applies. The context also comes to aid in presentational focus, a subtype of information focus (Krifka 2008: 251). This kind of focus indicates a covert question around which discourse is structured, such as *What was there?*

(20) (Krifka 2008: 251)

Once upon a time, there was a PRINCESS.

Ideally, multiple diagnostics are used for the investigation of focus in a given language given that cross-linguistically, different contexts for focus could result in different kinds of focus in different languages, such as answer focus or correction focus (Büring 2009: 179).

As an example for diagnostics aiming for different types of focus, consider the following test by É. Kiss (1998: 251). She provides a test to tease apart identificational focus and information focus, attributing the test to Donka Farkas. It includes a dialogue about buying a hat where the first turn contains the focus marked constituent, and depending on the felicitousness of the answer, it can be determined what type of focus is expressed in the first turn. Consider examples (21) and (22). In both first turns, it is expressed that Mary picked for herself a hat, and the *hat* is focussed, albeit in different positions in Hungarian. The continuation then asserts that Mary picked a coat as well. The purpose of the continuation is to negate exhaustivity. This is why this interpretation is only available in example (21a) which includes an identificational focus. The dialogue in ex. (22) is odd, since in the first turn, no exhaustive focus is expressed, but in the second turn, exhaustivity is negated where there is none. The meaning ‘too’ is incompatible with exhaustivity.

(21) Hungarian (É. Kiss 1998: 251)

- a. *Mari EGY KALAPOT nézett ki magának.*
 Mary a hat.ACC picked PFX herself.DAT
 ‘It was a HAT that Mary picked for herself.’

- b. *Nem, egy kabátot is ki nézett.*
 no a coat too PFX picked
 ‘No, she picked a coat, too.’

(22) Hungarian (É. Kiss 1998: 251)

- a. *Mari ki nézett magának EGY KALAPOT.*
 Mary out picked herself.DAT a hat.ACC
 ‘Mary picked herself a HAT.’
- b. %*Nem, egy kabátot is ki nézett.*
 no a coat.ACC too out picked
 ‘%No, she picked a coat, too.’

A final, central point to be kept in mind when using various diagnostics is made by van der Wal (2016: 294):

“[...] showing that a certain strategy is used in a sentence with a certain interpretation (e.g., exhaustive), or is compatible with / felicitous in a certain interpretively limited context, does not necessarily mean that the strategy is a dedicated marker for this interpretation. It may be that the strategy indeed encodes focus, but is underspecified as to the precise interpretation (the type of focus) it encodes, or it may be that the interpretation is a side-effect of some other mechanism [...]”.

This also means that if a strategy is *compatible* with a certain interpretation, this does not yet *exclude* other interpretations. For the investigation of focus in a given language this means that a given strategy of focus marking should ideally be tested in different contexts to clarify its exact function.

The considerations above make clear that for a thorough description of focus in a given language, one should use different diagnostics to check for the same means of focus marking. Among the most common diagnostics for focus, independent of exact formulation, are: answers to *wh*-questions and alternative

questions and corrections. These and other tests can target foci of different sizes, and the right context can target identificational or information foci.

3.4 Contrastive topics

The last information structural category that shall be discussed here is the contrastive topic. As an illustration consider first the two following examples from Büring (2016). In example (23a), the focus in the answer is interpreted exhaustively. The answer yields the inference: “It was her who did the kicking out, and it was me who got kicked out; no one else kicked anyone out”. In a related context, the same elements receive a different interpretation. The answer in example (23b) can be continued by “... whereas someone else wants to kick so-and-so out”. The word *she* in example (23b) is regarded as a contrastive topic, or marked with “contrast”.

- (23) a. (Did you kick her out? –) SHE(F) kicked ME(F) out!
 b. (Who do they want to kick out? –) SHE(CT) wants to kick ME(F) out.

In this thesis, I will follow the approach to contrastive topics described in Büring (2003, 2016). In this approach, the notion of Contrastive Topic is independent of the definition of Topics, contrary to other approaches that view contrast as separate category that can be a feature of both topics or foci (e.g. É. Kiss 1998; Neeleman and Vermeulen 2013), or contrastive topics as combination of topic and focus (Krifka 2008; Krifka and Musan 2012).

Büring (2003, 2016) argues that, while in an alternative semantic treatment focus relates a declarative sentence to alternative propositions, a contrastive topic in combination with a focus relates it to alternative questions. Büring (2016: 66) devises a rule for the interpretation of contrastive topics with three requirements:

- (24) For a sentence S^{CT+F} to be felicitous, there must be at least one question meaning in S^{CT+F} ’s CT-value which is
- a. currently pertinent, and
 - b. logically independent of [the ordinary meaning of] S^{CT+F} and
 - c. identifiable.

These requirements are a conventional implicature triggered by CT marking. As an illustration, consider again example (23b). The answer in this sentence could not be uttered as answer to the question in (23a). This is because there is no possible question which is independent of the question meaning of the original question. The context in example (23a) is a polar question, so what we are interested in is whether he kicked her out, or she him. This issue is completely resolved. There is no possible question “Who did *y* kick out?”, and the answer is thus in violation of requirement (b), independence (Büring 2016).

Contrastive topics are overtly marked by a characteristic rise-fall-rise prosody in English. Furthermore, in an English CT+F structure, the contrastive topic may be fronted across the focussed constituent whereas this is not possible with double focus. Consider example (25a): the context allows for a pair-list reading: Who recorded the song in which year? So possible pertinent questions for the answer in ex. (25b) are “Who recorded it in ’82?” or “Who recorded it in ’78?”, etc.

- (25) a. (Churchmoth recorded this song in eighty-three.)
 b. And MUCKENSTURM(F) recorded it in SEVENTY-TWO(CT) / And in SEVENTY TWO(CT), MUCKENSTURM(F) recorded it.

This is not possible in an F+F answer – the context makes it clear that the answer must be a F+F sequence this time, since for pragmatical reasons we know that a song can be written only once. Therefore there are no other pertinent questions like “When did Batiston write it?” or “Who wrote it in eighty-two?”:

- (26) a. (Churchmoth wrote this song in eighty-three.)
 b. No, MUCKENSTURM(F) wrote it in SEVENTY-TWO(F). / #No, in SEVENTY-TWO(F), MUCKENSTURM(F) wrote it.

Besides prosodic or syntactic marking, contrastive topics may be marked by topic marking morphemes or special contrast morphemes in languages like Korean or Guaraní. Notably, Meadow Mari, too, has a morpheme which can be employed to mark contrastive topics, namely the 3SG possessive suffix *žE* which is discussed at length in Section 8.

Typical contexts to evoke contrastive topics are multiple *wh*-questions (ex. 27a), single *wh*-questions containing plurals (ex. 27b), or implicational topics (ex. 27c). This third type again indicates additional questions such as “Where was the cook?”.

(27) (Büring 2016)

- a. (Which guest brought what? –) FRED(CT) brought the BEANS(F).
- b. (Where do your siblings live? –) My SISTER(CT) lives in STOCKHOLM(F).
- c. (Where was the gardener at the time of the murder? –) The GARDENER(CT) was in the HOUSE(F).

4 Strategies of information structure marking

The previous sections were concerned with defining the object of inquiry and reviewing a number of diagnostics for topic and focus. In the next section I will shortly discuss the actual linguistic strategies that encode some dimensions of information structure in the languages of the world, and discuss what implications this has for the current study. I will first summarize very shortly which strategies can be used to mark focus, cross-linguistically, before turning to the combination and the work load of different strategies in one single language. Topic marking strategies will briefly be touched upon. The Uralic languages will also be discussed in a separate subsection to facilitate comparison of information structural marking in Mari with the marking in other closely related Uralic languages.

4.1 Strategies of focus marking cross-linguistically

While information packaging is probably a phenomenon universal to human language, the coding of information structural categories is subject to cross-linguistic variation. Information structural categories interact with various grammatical phenomena such as intonation, morphology, or word order (Erteshik-Shir 2007: 2). Naturally, this concerns the expression of focussing as well. The alternative

semantics definition of focus in (12) does not refer to a specific linguistic strategy to code focus, so it is compatible with different ways of marking focus.

Cross-linguistically, focus is variously marked by prosodic phrasing, constituent order, special focus morphemes, or not at all (Büring 2009). Büring distinguishes the following types of focus marking in languages – boundary languages, edge languages, mixed languages, particle languages, and non-marking languages. He argues that these different kinds of coding can be traced back to a single principle going back to Truckenbrodt (1995). It states that focus is always realized by structural prominence, which itself ultimately comes down to prosodic prominence. Prosodic prominence deviating from the unmarked prosodic mapping in a language can be achieved by altering the prosodic structure of syntactic phrases, e.g. turning a focussed constituent into the head of a prosodic constituent where another constituent was the head first; inserting a prosodic boundary at the right edge to make it more prominent; or deleting a boundary at the right edge. Besides these adjustment strategies, there are non-prosodic strategies to make focussed constituents structurally prominent as well. The most important strategy is syntactic adjustment where a phrase is moved into a prosodically prominent position. This happens in languages that mark focus with constituent order (Büring 2009).

In boundary languages, the first type, focus is “marked by insertion of a prosodic phrase boundary to the left or right of the focus” (Büring 2009: 183). Edge languages, on the other hand, are characterized by non-standard constituent order marking focus. Edge languages typically minimize the material between the focus and the relevant prosodic phrase edge. This means that, while boundary languages alter the prosodic structure to mark focus, edge languages can only do this in a limited way or not at all. A sub-type of edge languages are relaxed-edge languages where focus is not realized exactly at the edge but close towards an edge. Most commonly, this type is found in languages with a rather strict verb-final syntax. Here, focus appears not at the right edge of an IP, but pre-verbally. The reason for this is apparently the fact that since the verb cannot be moved away from the final position for focus if it is not focussed itself, other focussed constituents must occupy the position closest to the IP edge, which is then directly before the verb (Büring 2009; Hyun-Oak Kim 1988). In strict-position lan-

guages, finally, focussed constituents are obligatorily put into a special syntactic position reserved for focussed constituents.

According to Büring (2009), a language can be classified as strict-position language only if there is no alternative in-situ strategy. For example, in many SOV languages the immediately preverbal position is often described as focus position, but at the same time it is also the position for unmarked objects. In such a case, the appearance of focussed objects in this position is no evidence for a strictly preverbal focus strategy because the position is underspecified in this respect (cf. Herring and Paolillo 1995: 168 on this issue in Sinhala). Similarly, van der Wal (2016: 294) states that “not all strategies that can be used for focus are indeed inherently focus strategies”. These issues are relevant for Mari as well, as it is an SOV language with focus in-situ or in the preverbal position (see Section 6). This means that the preverbal position is only an information-structurally “loaded” position, in Büring’s terms. That is, focussed constituents *can* be put in the preverbal position, but this is no evidence that there is a designated syntactic position for all focussed constituents.

A given language usually has a number of different strategies at its disposal for information structural marking, e.g. for focus marking, including a subset of the means mentioned above – prosodic phrasing, constituent order, special focus morphemes, or special constructions like clefts. The strategies may not always be used for equal purposes in different languages, however. For example, É. Kiss (1998: 267–271) argues on the basis of data from Hungarian, English, Romanian, Catalan, Greek, Arabic, and Finnish that deviations from the canonic word order are more likely to mark identificational focus than information focus (see also Skopeteas and Fanselow 2010: 170). Although the mapping between information structural notions and marking strategies is not absolute, some studies have found tendencies that more marked strategies may be preferably used to code more complex information structural notions. Based on data from Georgian, Hungarian, American English and Québec French, Skopeteas and Fanselow (2010: 190) devise a markedness scale of structural complexity for focus marking:

- (28) Scale of Structural Complexity (Skopeteas and Fanselow 2010: 190)
in situ < reordering < clefting

Expressed in simple terms they predict that more specific instances of focus (such as focus which excludes other alternatives) require a more marked strategy to encode focus in a given language. In turn, this scale predicts that, “if focus may be expressed *in situ* (through prosodic prominence), then *ex situ* focus occurs only in a subset of the instances in which a constituent bears a focus feature” (Skopeteas and Fanselow 2010: 173). There is no absolute mapping between a strategy in the scale and the type of focus, cross-linguistically. Concerning clefts, for example, Skopeteas and Fanselow (2010) found that in French, “clefts occur whenever the subject is part of whatever focus domain”, while in English “clefts only occur in contexts that license an identificationally focused subject” (Skopeteas and Fanselow 2010: 194). In sum, languages may use different strategies of focus marking, and the structurally more complex marking strategies are usually employed for more complex types of focus.

Another parameter in which languages differ with respect to focus marking via morphosyntactic strategies, is the basic word order of the language. In many languages that permit a large number of word order variants (e.g. German, Russian, Japanese, Hindi, and many others), the choice between these variants is “rarely truly optional and is often motivated by information structure.” (Erteshik-Shir 2007: 125). Languages where information structure has a strong effect on word order are called discourse-configurational. The word order in such languages is “free in terms of the order of the arguments and the verb, but it is fixed with respect to focus structure” (Erteshik-Shir 2007: 85).

It has long been recognized that the position for argument focus is connected to the basic word order type of a language (e.g. Dezső 1978; Hyun-Oak Kim 1988). In SVO languages, the position is usually the immediately *postverbal* position while the typical position for focussed constituents in SOV languages is the immediately *preverbal* position (Dezső 1978: 7–9; Hyun-Oak Kim 1988: 147; Herring 1990). In both cases, proximity to the verb is an important factor. Another study found that in a sample of 59 SOV languages, focus was in the preverbal position in 21 languages and in the sentence initial position in 20 languages; 2 languages had a postverbal focus position (Czypionka 2007). Importantly, at least 25 of the 59 SOV languages also allowed focus marking *in situ*, and Czypionka (2007: 443) remarks that “the proportion of languages with optional *in situ* focus marking

is likely to be higher than this”. In the SVO languages in this study, focussed constituents were mostly found in the sentence initial position.

4.2 Strategies of topic marking cross-linguistically

As is the case with focus, languages can employ different strategies for topic marking as well: Languages may mark topical constituents by morphological or prosodic markers, distinguished syntactic positions, or scrambling. In Section 3.2.2, we already saw examples for special clitics marking shift Topics in Tzotzil (example 10) and a L* tone contour marking familiar topics in Italian (example 11). In Japanese and Korean, topical constituents are marked with the enclitics *-wa*, and *-nun*, respectively.

(29) Korean (Erteshik-Shir 2007: 98)

I chayk-un Swunja-ka sassta.
this book-TOP Swunja-NOM bought

‘As for this book, Swunja bought (it).’

Hungarian, on the other hand, is well known for its designated position for topical constituents in the left periphery (see Gyuris 2012). The topic field precedes the focus position. Two diagnostics are commonly employed: focus is known to immediately precede the verb in Hungarian, and if a constituent fills this position, all “verbal modifiers”, among them verbal prefixes, have to move to a position behind the verb. In example (31), the order of prefix and verb is inverted because *Marit* is focussed and has to occupy the immediately preverbal position. Sentence adverbials are not allowed to appear after the constituent in the focus position, so they can be used as diagnostic for topics. In example (30), the sentence adverbial *szerencsére* can appear in any position between the sentence initial and the focus position – after topical *Marit* as well – whereas in example (31) it cannot follow focal *Marit* (Gyuris 2012).

- (30) Hungarian (Gyuris 2012: 176)

(Szerencsére) János (szerencsére) Marit (szerencsére) meglátta a
 luckily John luckily Mary.ACC luckily PFX.saw the
buliban.
 party.INE

‘Luckily, John_T saw Mary_T at the party.’

- (31) Hungarian (Gyuris 2012: 176)

(Szerencsére) János (szerencsére) Marit (*szerencsére) látta meg a
 luckily John luckily Mary.ACC luckily saw PFX the
buliban.
 party.INE

‘Luckily, John_T saw Mary_F at the party.’

Non-topical or non-focal constituents consequently appear marked differently from topics or foci. In SOV languages there is a tendency for backgrounded constituents to appear in the postverbal position. Such constituents are intonationally unremarkable and “tend to express highly predictable, thematic information; accordingly, continuous topics may appear sentence-finally” (Herring 1990: 67). For example, Khalkha Mongolian, a SOV language, allows postverbal constituents in colloquial speech. Such constituents “represent afterthoughts, providing or repairing missing information in the clause” (Öztürk 2013: 272). Uyghur is another example for an SOV language that behaves this way. Postverbal constituents in Uyghur “denote backgrounding, i.e. they express supplementary and given information recoverable from discourse. Topic- or focus-related elements, on the other hand, typically appear in the preverbal domain” (Öztürk 2013: 284). Note however, that the postverbal field in SOV languages is very likely not derived uniformly (cf. Pregla 2023).

4.3 Information structure marking in the Uralic languages

From the viewpoint of information structuring, the Uralic languages differ considerably from each other, as a consequence of large structural differences (Surányi

et al. to appear), possibly also as a result of belonging to different typological areas which are also partly defined by their main contact languages. In the family as whole, various strategies for information structural marking can be encountered, including “differential object marking, differential object agreement, diathetic operations (passive, ‘dative shift’), focus- or topic-marking particles, nominal determination by definite articles, definite declension, and possessive suffixes, as well as clefting” (Klumpp and Skribnik 2022: 1018, see also de Smit and Janda 2023).

Notwithstanding their great differences, a few common traits can be observed. Topics are sentence-initial throughout the Uralic language family. Foci are expressed in different ways but pattern alike in related or geographically close languages; in Hungarian and the Ob-Ugric languages, focus is placed immediately before the verb, while in Finnic, and partly, the Permic languages (e.g. Udmurt and Komi) focus is in a clause-final position (Klumpp and Skribnik 2022: 1021; Surányi et al. to appear). One study on Hungarian, Finnish, Estonian and Udmurt (Surányi et al. to appear) found that all languages can encode topic and focus via word order changes. Topics can be marked by fronting of some type (not necessarily to the sentence-initial position), and focus can be encoded by putting constituents into a preverbal position in Hungarian and Udmurt, or into a postverbal position in Estonian and Finnish, and partly Udmurt as well.

Specific morphological means of information structure marking such as particles or clitics are not common in the family, although in several languages, possessive suffixes can be employed for information structure marking. In the languages of the Volga-Kama area (e.g. Mari, Komi, Udmurt), possessive suffixes are frequently employed for contrastive purposes (which sets them apart from other Uralic languages in this respect, Nikolaeva 2003) although the details of their usage differs (see Nikolaeva 2003, A. Simonenko 2014 and Klumpp and Skribnik 2022: 1025–1026 for comparative descriptions). For example, the Komi enclitic contrastive particles *=tə* and *=sə* can be qualified as special particles for coding information structural values. They originate from 2SG and 3SG possessive suffixes, respectively, which merged with the accusative. Their function is described as contrastive and thus marking both contrastive focus and contrastive topics, although the whole range of functions is not well understood yet (Klumpp and Skribnik 2022: 1022). In the Ob-Ugric and Samoyedic languages, on the other

hand, possessive suffixes can be employed to express various relations between two individuals in the discourse context or in narrative situations (Janda 2015; Körtvély 2010; Siegl 2021; Zayzon 2015, Wagner-Nagy 2018: 463–467). For Northern Mansi, Janda (2015: 250–254) argues that possessive suffixes may be employed for a topic shift. The 3SG possessive suffix in Mari is discussed in more detail in Section 8.

As discussed in section 4.1, the focus position correlates with the basic word order – in SOV languages it is usually the preverbal position that hosts focussed constituents, in SVO languages the postverbal position. It has been argued that word order change as a result of language contact can also influence the placement of focus. One example for this development is Udmurt (Asztalos 2020; Tanczos 2010). Typical for an SOV language, Udmurt has an immediately preverbal focus position. However, one study found that focus is not restricted to this position alone, but can also occur postverbally and in a preverbal position not directly adjacent to the verb (Asztalos 2020). The details of focus placement in Udmurt are rather intricate and the possible position of focussed constituents depends on their argument role, case marking, and whether they are pronouns or not (see Asztalos 2020; Pregla 2023). In any case, postverbal focus was largely accepted by speakers, although postverbal or sentence-final focus is rather untypical for SOV languages. On the other hand, Russian, the most important contact language of Udmurt, also has a sentence-final focus position. It has been argued before that Udmurt is currently undergoing a change from an SOV to an SVO language (Asztalos 2018, 2021; Asztalos, Gugán, and Mus 2017), and based on her findings, Asztalos (2020) argues that Russian has influenced focus placement in Udmurt in the same process as well.

In comparison to the other Uralic languages spoken in the vicinity of Mari (Udmurt and Komi), Mari seems to be rather conservative regarding the SOV word order. In a study comparing translations of a novel from Russian into Mari, Udmurt and two varieties of Komi, Bradley, Kellner, and Partanen (2018) showed that both Komi varieties predominantly used the VO order, Udmurt predominantly OV, and that the usage of OV was especially frequent in Mari. That is, at least literary varieties of Mari are conservative with regards to the SOV order. Another study based on the word order in the representation of dialogue speech

in Mari literature found that in about 30% of all sentences, the word order deviated from SOV, but the corpus for the given study consisted only of about 300 dialogues (Yegorova 1981: 275). The influence of Russian on Mari is strong like on any minority language in Russia today, however, and inferences can be found in various syntactic domains, e.g. reverse direction of the head-dependent relations or word order shifting away from strictly verb-final clauses to SVO (e.g. Anduganov 2001). The same study found some examples for word order variation in newspapers as well (Anduganov 2001: 73–74). For example, Anduganov provides example (32), claiming that it is indicative of a shift to a slow head-initial syntax in Mari since the verb is followed by an illative adverbial. The example is presented in isolation, however, and it is not taken into account that constituents can appear in a postverbal position if backgrounded as is the case in other SOV languages (see section 7.2 on Mari).

(32) (Anduganov 2001: 73)

Kastene, indeš sagat gutlašte [...], mij-en šu-na
in.the.evening nine clock around come-CVB arrive-PST1.1PL
avtovokzal-aš.
bus.station-ILL

‘In the evening around nine o’clock we arrived at the bus station.’

4.4 Summary

This section gave a short overview about different strategies of expressing focus and topic, and the implications for this study. In the languages of the world, focus can be marked by prosodic phrasing, constituent order, special focus morphemes, or not at all. Languages usually mix strategies, however, and in that case more marked strategies may be preferably used to code more complex information structural notions. Most importantly, a language should only be classified as a language marking focus by moving constituents into one designated position (that is, as strict position language) if there is no alternative in-situ strategy for focus. As regards constituent order as means of focus marking, it is well known that

the basic word order in a language correlates with its position for focussed constituents. In SVO languages, the position is usually in the immediately postverbal position while the typical position for focussed constituents in SOV languages is the immediately preverbal position. The preverbal position is, however, usually neither reserved for focussed constituents, nor necessarily a designated focus position. Objects can be found in the preverbal position in neutral sentences in a SOV language as well, which means that the preverbal position of a focussed object can be epiphenomenal. Non-focal and non-topical constituents are usually located in the postverbal field in SOV languages.

5 Method

In this section I shortly discuss different methods to document topic and (mostly) focus strategies, with emphasis on fieldwork, and discuss the method used for data collection in the following chapters of this thesis. The issue of how to diagnose topic and focus with various tests was already discussed in Sections 3.2.3 and 3.3.3. The diagnostics themselves are of course not limited to a certain speech situation; *wh*-questions, for example, can occur in natural speech or be presented in semantic judgment tasks. So far, the data in this work has come from various sources; already existing work on Mari with elicited data, corpus data, questionnaires, or semantic judgments.

The corpus data in this work comes from different sources. The largest Mari corpus is the Mari Corpus in the Giellatekno/Divvun infrastructure (*CLM*) (Trosterud et al. 2023). The 57.3 million tokens are from a variety of sources including fiction and non-fiction texts. Unfortunately, the corpus does not allow to show the context of a sentence, which makes it less suitable for gathering data about discourse-dependent information structural phenomena. For this reason, the data in the section about Topics (Section 7) mostly comes from two other sources; the Corpus of Contemporary Written Literary Meadow Mari (*CCWLM*) (Arkhangelskiy 2023a), developed and maintained by Timofey Arkhangelskiy, and a small corpus of Mari texts by Serdobol'skaya and Toldova (2012). The Corpus of Contemporary Written Literary Meadow Mari has a size of 5.53 million tokens and mostly consists of contemporary newspaper texts, and context can be shown up to four

sentences before and after the target sentence. The Mari texts in Serdobol'skaya and Toldova (2012) in turn are a small corpus of a few texts recorded in the year 2000. Despite their short length, they are a valuable source of contemporary spoken Mari for which there are not many transcribed sources. The section about the suffix *žE* relies heavily on that small corpus. Other sources of contemporary spoken Meadow Mari are the Spoken corpus of Meadow Mari (*SCMM*) (Volkova et al. 2023) which includes texts recorded in the years 2000–2004 and 2018 in the village Staryj Torjal. For this thesis I did not systematically go through the texts in the corpus, but in future investigations, that corpus will likely play an important role in addition to the texts in Serdobol'skaya and Toldova (2012). Another source of contemporary Mari is the Corpus of Social Media in Meadow Mari (*CSMM*) (Arkhangelskiy 2023b) which consists of 3.59 million words of written text from the social network “VKontakte”.³ In this corpus it is more difficult to find longer stretches of coherent text due to the nature of online communication, however, which makes it less suitable for the investigation of information structure.

To gather additional data for strategies of focus marking and the suffix *žE* in Mari, I worked with a consultant. My consultant is a female L1 speaker of Meadow Mari in the age group 30–40, originally from the central Meadow Mari dialect area. She holds a university degree. The elicitation sessions were conducted online. My consultant has a very good command of English, which was used as a meta language in the elicitation sessions. One obvious drawback in using only one speaker's data is that the judgments solely reflect that single one speaker's own grammar, and the judgments may not necessarily align with the ones from other speakers. This issue is a central concern in empirical linguistic research. Mari is no exception, and some issues regarding rather divergent judgments of various grammatical phenomena have been noted before. These also concern some domains of the language which are relevant for information structure, for example, the preferred word order in clauses involving a possible contrastive topic (Georgieva 2022a: 5), the interpretation of the 3rd person singular possessive suffix (Georgieva 2022a: 6), or the preference vs. dispreference

³A description of the Corpus of Social Media in Meadow Mari is provided in Arkhangelskiy (2019).

for *wh*-movement (Georgieva 2022b: 1). Nevertheless, the data gathered here can be used as first approximation towards an account of some information structural notions and their expression in Mari, and can serve as the basis for further fieldwork with other speakers.

Coming back to the fieldwork methods, different methods can be employed to document different strategies of focus marking: observing focus constructions in naturally occurring speech; focus constructions as part of a literacy-based elicitation or as part of a non-literacy-based elicitation, (e.g. with the help of pictures) (see Aissen 2015: 2). All these strategies came to be used in the elicitation sessions. The largest portion of the elicitation can be classified as non-literacy-based elicitation for which I used different stimulus kits: Storyboards from Totem Field Storyboards (Grubic 2014; Littell 2010a,b; TFS Working Group 2011), a few pictures from Benjamin Bruening’s Scope Fieldwork Project (Bruening 2012), and the Questionnaire on Information Structure (Skopeteas, Fiedler, et al. 2006). Examples for the former two stimulus kits are presented in Figure 1 and Figure 2.



Figure 1: “Anna is not giving food to a ghost, she is giving food to her friend!” (Grubic 2014)



Figure 2: “The person who is holding the rifle is Mihail” (Bruening 2012)

Figure 1 is part of the Totem Field Storyboards storyboard *Lakka and the Ghosts* (Grubic 2014). A storyboard is a series of pictures visualizing a story for elicitation. The story contains situations which are designed to elicit constructions that the researcher is interested in, several times. Using the storyboard as visual aid, the story is first told in the metalanguage (English in this case) until the consultant knows the story well enough to re-tell it in their own language (Meadow Mari in this case). The story is recorded and transcribed. Later, the transcription can also serve as basis to elicit negative data. For example, could a certain situation be described with another construction or not? One advantage of the storyboard task is that it provides a more or less natural context to elicit the constructions in question, and story-telling involves more natural speech (Grubic 2022: 5–7). The pictures from the *The Scope Fieldwork Project* (Bruening 2012) served as visual aid to conjure up situations where a certain property is either true for only one individual (e.g. holding a rifle, as in Figure 2) or for all individuals.

A visual aid facilitates the elicitation and production task for both the consultant and the researcher, minimizing potential problems with the presentation of

the context. Ideally, there is no interference from the contact language, and not from the researcher's choice of words in drawing up the context (Grubic 2022: 7). Thus, picture-based elicitation tasks circumvent possible interference from a meta-language.⁴ A drawback to picture-based elicitation is its rigidity, however: it is hardly possible to draw materials for all elicitation purposes, and the elicitation is limited to the content of the pictures. The picture-based tasks should therefore be supplemented with explicit questions that allow to bring up new test items on the spot (Matthewson 2004: 378–379). Therefore I used semantic judgment tasks in addition to the picture tasks.

Focus is dependent on the context in which the sentence containing it is uttered. Focus size in the answers depends only on the question, that is, the context in which the answer is uttered. The most straightforward way of a judgment task for focus is therefore to present a certain context and ask whether in that given context a certain focus construction is acceptable or not. The reverse method is what van der Wal (2016: 264–265) calls “context conjuring”: a sentence containing (a special type of) focus is presented, and the speaker is asked to think about a context the sentence could be uttered in (van der Wal 2016: 264–265). In the elicitation sessions conducted for this thesis, the storyboards and other visually aided elicitation proved to be the most fruitful, both in terms of the actual linguistic data produced with their help and in terms of how fluently the session went. Therefore most of the elicited data presented in the section about focus originated in picture-based elicitation tasks.

6 Focus in Mari

This section deals with strategies of information focus, identificational focus, and focus negation in Mari. I discuss these strategies on the basis of existing data from the literature and newly elicited data.

⁴See Matthewson (2004: 394–398) for discussion about meta vs. object language in elicitations.

6.1 Information focus

In this subsection I discuss several strategies of expressing information focus in Mari: (1) constituent focus and (2) full clausal focus in answers to *wh*-questions, (3) presentational focus, (4) correction focus, (5) verbal focus and (6) focus in polar questions.

A basic insight found in the literature on Mari is the observation that constituents can be “emphasized” by prosodically stressing them (Yegorova 1981: 274; Riese et al. 2019: §12.3; Il’in 1986), although often the interpretation of these constituents is not further elaborated upon. As mentioned in the introduction, the often used term *logicheskoye udareniye* (“logical stress”) likely refers to constituents with a focus-like function (Asztalos 2020: 18). The only studies which deal with sentence intonation and prosody in Meadow Mari that I’m aware of are Baitchura (1988) and Il’in (1986), but they do not systematically deal with the prosody of focus.⁵

The neutral word order is SOV in Meadow Mari (see Section 2.2). Previous research indicates that focussed constituents in Mari, including both subjects and objects, can either occupy the preverbal position or the same position as in the unmarked clause (Nelson, Vedernikova, and Bradley to appear: 7, Saarinen 2022: 463, Vilkuna 1998: 195). Additionally, some studies have found that contrastively focussed constituents may occupy a clause-final position as well (Vilkuna 1998: 195; Pregla 2023). For example, Vilkuna (1998: 195) states that the pre-verbal position in Mari is a frequent position for focussed constituents, but that they are also permitted in other positions. In her data on subject focus, there are sentence-initial subject foci, preverbal subject foci and postverbal subject foci.

(33) (Vilkuna 1998: 195)

- a. *Ača-(že) ergə-žə-m kər-en.*
father-PX3SG son-PX3SG-ACC hit-PST2.3SG
‘The father hit his son.’

⁵For word-level prosody of Meadow Mari see furthermore Lehiste et al. (2005) and F. Rozhanskiy (2013) and references therein, and for Hill Mari see Zhornik and Rozhanskiy (2021), and references therein.

- b. *Uke*, (AVA-ŽE) *tudə-m* (AVA-ŽE) *kər-en*
 no mother-PX3SG 3SG-ACC mother-PX3SG hit-PST2.3SG
 (AVA-ŽE).
 mother-PX3SG
 ‘No, his/the MOTHER hit him.’

6.1.1 Constituent focus

My consultant preferred prosodic in-situ marking of constituent focus, but she also accepted preverbal foci if presented with them. In *wh*-questions, she preferred *wh*-words asking for different parts of the sentence – nominative subjects, accusative objects, dative objects, or PP – in the left periphery.⁶ (Multiple *wh*-questions are discussed in Section 8.3.) In the answer, the word order was the unmarked SOV in all cases. Focus was marked by prosody only: the intonation rose on the stressed syllable of the focussed constituent in either answer.

(34) (elicited)

- a. *A kö kočk-aš kond-en?* – *ur jod-o*.
 and who eat-INF bring-PST2.3SG squirrel ask-PST1.3SG
 ‘And who brought (something) to eat? – asked the squirrel.’
- b. *MASKA kočk-aš kond-en* – *krab kalasə-š*.
 bear eat-INF bring-PST.3SG crab say-PST1.3SG
 ‘The BEAR brought (something) to eat – said the crab.’

⁶ *Wh*-words may stay in situ or be moved to the left periphery of the clause. The distribution of the two types and the reasons for choosing one strategy over the other are not entirely clear. Nelson, Vedernikova, and Bradley (to appear) attribute the choice to register (the in-situ strategy being more frequent in spoken language, and the movement strategy a feature of the literary language), Georgieva (2022b: 1) notes that there is speaker variation, and Vilkuna (1998: 207) simply remarks that *wh*-elements “frequently but not obligatory occupy the pre-verbal, verb-adjacent position”. My consultant furthermore remarked that there was a difference in politeness between placement possibilities of interrogative pronouns. The reasons for one placement over the other are likely attributable to a number of these factors, but I am not going to discuss this any further. In any case, my consultant preferred fronting all *wh*-words (see examples (34) and (35) and the discussion there).

(35) (elicited)

- a. *Mo-m maska kond-en? – ur jod-o.*
what-ACC bear bring-PST2.3SG squirrel ask-PST1.3SG
‘What did the bear bring? – asked the squirrel?’
- b. *Maska KOČKƏŠ-ƏM kond-en – krab vašeštə-š.*
bear food-ACC bring-PST2.3SG crab answer-PST1.3SG
‘The bear brought the FOOD – answered the crab.’

(36) (elicited)

- a. *A kö-lan tide čəla pajrem ələ? – ur*
and who-DAT this whole party be.PST1.3SG squirrel
jod-ən.
ask-PST2.3SG
‘But who was this whole party for? – asked the squirrel.’
- b. *Təj ot pale? Tide čəla pajrem MƏLAM, KRAB-LAN –*
2SG NEG.2SG know.CNG this all party 1SG.DAT crab-DAT
krab vašeštə-š.
crab answer-PST1.3SG
‘Don’t you know? This whole party was for ME, THE CRAB – answered the crab.’

(37) (elicited)

- a. *A kö-lan verč tide čəla pajrem-əm əšt-en-ət? – ur*
and who-DAT for this all party-ACC do-PST2-3PL squirrel
jod-o.
ask-PST1.3SG
‘But who did they throw this whole party for? – asked the squirrel.’

- b. *A tāj ot šarne? – krab kəčkəraļe. Tide čəla*
 but 2SG NEG.2SG remember.CNG crab yell.PST1.3SG this all
məJ-əN verč!
 1SG-GEN for

‘But don’t you remember? – yelled the crab. It (was) all for ME!’

The second strategy of constituent focus is placing the focussed constituent in a preverbal position, illustrated with data from Pregla (2023: 90). Preverbal subjects are only acceptable when focussed, and not in neutral contexts:

(38) (Pregla 2023: 90)

- Ikeče jal-əšte joča-lan jomak-vlak-əm KOVA /*
 recently village-INE child-DAT tale-PL-ACC grandmother
**kova ojlə-š.*
 grandmother tell-PST1.3SG

‘GRANDMA / *Grandma recently told stories to the children in the village.’

Verbs can appear further left in the clause to allow focussing a constituent in preverbal position as well. That is, the focus-verb string can either appear in the clause-final position or clause medially.

(39) (Pregla 2023: 138)

- Poškudo TENGEČE ERDENE puə-š joča-lan pört-əštö*
 neighbor yesterday morning give-PST1.3SG child-DAT house-INE
pəras-əm.
 cat-ACC

‘The neighbor gave a cat to the child in the house YESTERDAY MORNING.’

6.1.2 Full clausal focus

Full clausal focus after all-new questions involved only the canonical word order without any special prosodic marking:

(40) (elicited)

Mo təšte? – ÜDƏRAMAŞ JOŠKAR ARAKA-M JÜ-EŠ.

what here woman red wine-ACC drink-3SG

‘What’s (going on) here? – THE WOMAN IS DRINKING RED WINE.’

(41) (elicited)

Mo təšte kaj-a? – RVEZE MEČE DENE MOD-EŠ.

what here go-3SG boy ball with play-3SG

‘What’s going on here? – THE BOY IS PLAYING WITH A BALL.’

(42) (elicited)

Mo lij-ən? – TUDO JAL-əŠTE KOL-EN.

what be-PST2.3SG 3SG village-INE die-PST2.3SG

‘What happened? – HE DIED IN THE VILLAGE.’

6.1.3 Presentational focus

A subtype of information focus, presentational focus indicates a covert question in the discourse. In this context, constituent foci may appear in the postverbal position in Mari. Such postverbal “presentational” subjects are typical as openings of folk tales, as in example (43), recorded between 1916 and 1918, but similar instances can be found in modern texts as well. Example (44) is a spoken language example recorded in 2000, and example (45) is from a contemporary newspaper text about an excellent student. After describing their motivation to teach Mari, the student’s name is introduced again with presentational focus, in a postverbal position. This example is discussed with some more discourse context on page 84.

(43) Eastern Meadow Mari, Petrušino (Beke 1938: 257)

Ožnê ik tšerk-an jal-əštê ul-mažê POP.

once one church-ADJ village-INE be-NMLZ priest

‘Once upon a time in a church village there was a PRIEST.’

- (44) (Serdobol'skaya and Toldova 2012: 712)

Memnan pört vokterne ələ TƏGAJ KLAT.
 2SG.GEN house besides be.PST1.3SG such storage

‘Next to our house there was a STORAGE.’

- (45) CCWLM, Newspaper *Marij El* (27.05.2006)

Tə ojarthem-əm şot-əş nal-ən paşa-m əşt-a GALINA
 this peculiarity-ACC count-ILL take-CVB work-ACC do-3SG Galina

ALEKSEEVNA.

Alekseevna

‘It’s GALINA ALEKSEEVNA working, taking this peculiarity into account.’

6.1.4 Correction focus

Correction focus is realized just like other types of information focus. Both subjects and objects remained in the same position as in the unmarked clause, and the focussed constituents received a rising intonation on the stressed syllable.

- (46) (elicited)

a. *Məj təj deč-at şukə-rak kogələ-m əşt-en kert-am.*
 1SG 2SG from-ADD much-CMPR pie-ACC do-CVB can-1SG

‘I can bake more pies than you.’

b. *Uke, məj şukə-rak kogələ-m əşt-en kert-am.*
 no 1SG much-CMPR pie-ACC do-CVB can-1SG

‘No, I can bake more pies.’

- (47) (elicited. In the garden, is the woman hitting a girl?)

Uke, sad-əşte üdəramaş RVEZƏ-M şel-eş.
 no garden-INE girl boy-ACC hit-3SG

‘No, in the garden, the girl is hitting a BOY.’

Correction focus with the focussed constituent in a position different from the position in the context clause was judged licit but dispreferred (see also Pregla 2023: 88). The following examples show this phenomenon with correction focus in situ versus postverbally. More examples of correction focus with negation are discussed in sections 6.1.5 and 6.3.

(48) (elicited)

- A. *Ivuk kniga-m lud-ən.*
 Ivuk book-ACC read-PST2.3SG
 ‘Ivuk read a book.’
- B. *Uke, GAZET-əM lud-ən.*
 no newspaper-ACC read-PST2.3SG
 ‘No, (he) read a NEWSPAPER.’
- B.?? *Uke, Ivuk lud-ən GAZET-əM.*
 no Ivuk read-PST2.3SG newspaper-ACC
 ‘No, Ivuk read a NEWSPAPER.’

(49) (elicited)

- A. *Kniga-m Ivuk lud-ən.*
 book-ACC Ivuk read-PST2.3SG
 ‘Ivuk read a book.’
- B. *Uke, ANUŠ lud-ən.*
 no Anuš read-PST2.3SG
 ‘No, ANUŠ read (it).’
- B.?? *Uke, kniga-m lud-ən ANUŠ.*
 no book-ACC read-PST2.3SG Anuš
 ‘No, Anuš read the book.’

6.1.5 Verbal focus and verum focus

Verbal focus is marked either by prosody or by the verb occupying a clause initial position. In the first strategy, the word order is the canonical SOV and the verb is prosodically marked. In a polar question, the intonation rises on the stressed vowel of the sentence-final verb if the verb is focussed (Il'in 1986: 93), other than in broad focus questions where the intonation falls again (see Section 6.1.6). In contrast to a neutral clause where the sentence intonation falls on the stressed vowel of the preverbal constituent, the intonation falls on the stressed syllable of the verb in final position (cf. Il'in 1986: 88–89):

(50) (elicited)

Tudo jal-ašte IL-EN? – Uke, tudo jal-ašte KOL-EN.
 3SG village-INE live-PST2.3SG no 3SG village-INE die-PST2.3SG

‘Did s/he LIVE in the village? – No, s/he DIED in the village.’

The second strategy is characterized by fronting the verb, leaving other constituents in a postverbal position (cf. Chkhaidze 1941: 29). If a finite verb is focussed, be it the main lexical verb, or a negation verb, it can appear fronted in a left peripheral position. In a verbal complex, the negative auxiliary and the main verb cannot be separated, but modal verbs can be separated from the lexical verb they select (Georgieva, Salzmann, and Weisser 2021). This construction is commonly associated with correction focus. Subjects, objects and adverbs are located after the verb in examples (51–54) and remain in their relative order as in the neutral clause. Note however, that even in the case of verbal focus via verb fronting, place adverbials may still precede the focussed verb without being focussed themselves (ex. (54), see Section 7.1.3 on the position of adverbials in the left periphery). See also Georgieva, Salzmann, and Weisser (2021: 463) and Pregla (2023: 138) and section 7.2.3 for similar effects of placing constituents into a postverbal position.⁷

⁷Note that in example (52), both the negation verb and the lexical verb (*aš jüktö*) are focussed, but the small caps are poorly visible on *aš*.

(51) (elicited)

Tenij māj Budapest-aške š-am kaj. – KAJ-aš-ač tāj
this.year 1SG Budapest-ILL NEG.PST.1-1SG go.CNG go-PST1-2SG 2SG
tenij Budapest-aške!
this.year Budapest-ILL

‘I did not go to Budapest this year. – You DID GO to Budapest this year!’

(52) (elicited)

Vrač joča-lan em-am jüktə-š. – aš
doctor child-DAT medicine-ACC give.to.drink-PST1.3SG NEG.PST1.3SG
JÜ-KTÖ vrač joča-lan em-am!
give.to.drink.CNG doctor child-ACC medicine-ACC

‘The doctor gave the child medicine to drink. – The doctor DID NOT give the child medicine to drink!’

(53) (Trosterud et al. 2023), News text

O-M KERT māj tidə-m niguze ašt-en.
NEG-1SG can.CNG 1SG this-ACC no.way do-CVB

‘I CAN NOT do this in any way.’

(54) (elicited, Context: Someone argues it is allowed to smoke here.)

A təšte OK LIJ tamak-am šupš-aš.
but here NEG.3SG be.CNG tobacco-ACC suck-INF

‘But you CAN’T smoke here.’

6.1.6 Focus in polar questions

A final issue concerns constituent focus marking in polar questions, where prosody interacts with the polar question marking particle *mo*. As marker of polar questions asking for the truth value of the sentence, the particle is placed sentence-finally and can be used after both affirmative and negative sentences:

- (55) (Saarinen 2015: 345)
Pij kudal-eš mo? – Uke.
 dog run-3SG Q no
 ‘Is the dog running? – No.’

- (56) (Saarinen 2015: 345)
Pij ok kudal mo? – Uke.
 dog NEG.3SG run.CNG Q no
 ‘Is the dog not running? – No.’

The prosody of polar questions and of focus in polar questions has received some attention in the literature. In a study about sentence intonation in Mari, Baitchura (1988) notes that in a polar questions like in ex. (57), the intonation of the sentence is generally falling, with intermediate higher tones on *taj*, the stressed second vowel of *gazETəm*, and an intense rise on the penultimate syllable, *nalAT* right before the question particle *mo*, where it falls again. Riese et al. (2019: §12.4.2.1) have made similar observations; In example (58), intonation rises on the first syllable of the predicate, *šočan* ‘born’ and falls on *mo*.

- (57) (Baitchura 1988: 122–125)
A taj iктаž-mogaj gazet-əm nal-at mo?
 but 2SG INDEF-what.kind.of newspaper-ACC take-2SG Q
 ‘And do you receive any newspaper?’

- (58) (Riese et al. 2019: 421)
Tudo jal-əšte šoč-ən mo?
 3SG village-INE be.born-PST2.3SG Q
 ‘Was she born in the village?’

Furthermore, Riese et al. (2019: §14.4.2.1) found that if a single constituent is focussed in a polar question, intonation starts to rise on the stressed vowel of that constituent and falls from there to *mo* (Riese et al. 2019: 421). In this case the particle may be placed sentence-finally or directly after the focussed constituent,

and this has an effect on the prosodic realization. In example (59) the intonation starts to rise on the first syllable of the NP *jaləšte* ‘in the village’ and falls before *mo*. The same happens when the particle is placed sentence-finally with *jaləšte* focussed (ex. 60), but in this case the intonation contour continues over the verb, and falls again before *mo*.

(59) (Riese et al. 2019: §14.4.2.1)

Tudo JAL-əŠTE mo šoč-ən?
 3SG village-INE Q be.born-PST2.3SG
 ‘Was she born in the VILLAGE?’

(60) (Riese et al. 2019: 421)

Tudo JAL-əŠTE šoč-ən mo?
 3SG village-INE be.born-PST2.3SG Q
 ‘Was she born in the VILLAGE?’

Combining the observations that intonation starts to rise on the focussed constituent in a polar question, and that intonation likewise rises on the (stressed vowel of the) verb before *mo*, one could expect focus syncretism between broad sentential focus and verbal focus in example (58). But this is not the case (ex. 61): According to my consultant, it is not possible to use *mo* when focussing the verb despite using the standard prosodic contour with a rise on the stressed, second syllable of *ilen* and a fall before *mo*. Instead, to focus the verb, the question must be formulated bare without *mo*, with a rise on *ilen* and no subsequent fall (see Section 6.1.5). The verb in the answer is stressed as well, but with a globally falling intonation. This example is in direct contrast to example (50) where the verb is focussed.

(61) (elicited)

a. *Tudo jal-əšte il-en* *mo?*
 3SG village-INE live-PST2.3SG Q
 ‘Did s/he live in the village? / #Did s/he LIVE in the village?’

- b. (– #*Uke*, *tudo jal-əšte* *KOL-EN*.)
 no 3SG village-INE die-PST2.3SG
 ‘(– No, s/he DIED in the village.)’

6.2 Identificational focus

To express identificational focus, Mari utilizes the same marking as employed for information focus or a pseudocleft. Furthermore, the phrase *molo ... koklašte* makes a focus necessarily identificational. Pseudoclefts are only felicitous in situations in which a sentence provides the exhaustive extension of the predicate. Consider the following examples. When the context indicates that there is only one individual for which the predicate holds, usage a pseudocleft is felicitous (this situation is depicted in Figure 2 on page 41), as is the phrase *molo ajdeme koklašte*:

- (62) (elicited. In a situation where only Mihail is holding a rifle:)

Kö pəçal-əm kuč-a, tide MIHAIL.
 who rifle-ACC hold-3SG this Mihail
 ‘It is MIHAIL who is holding the rifle.’

- (63) (elicited. In a situation where only Jəvan is holding a bottle:)

Molo ajdeme koklašte JəVAN butəlka-m kuč-a.
 other person among Jəvan bottle-ACC hold-3SG
 ‘Among the people it is JəVAN who is holding the bottle.’

When the context indicates that a predicate holds of more than one individual in a given situation, usage of a pseudocleft is not felicitous. This is shown in example (64) where the predicate “holding a bottle” holds of all people. In the same context the phrase *molo ajdeme koklašte* ‘among the people’ can not be used, either, because it entails that from the set of people, only one is doing something particular (ex. 65).

(64) (elicited. In a situation where three people are holding one bottle each:)

#*Kö butəlka-m kuč-a, tide JəVAN.*

who bottle-ACC hold-3SG this Jəvan

‘(The person) who is holding the bottle is JəVAN.’

(65) (elicited. In a situation where three people are holding one bottle each:)

#*Molo ajdeme koklašte JəVAN-əN kid-əštə-že butəlka.*

other person among Jəvan-GEN hand-INE-PX3SG bottle

‘Among the people, the bottle is in JəVAN’s hands.’

It should be noted, however, that my consultant considered the pseudocleft a marked construction of the literary standard which possibly was introduced under influence of Russian, and not of their own spoken variety. Besides the specialized pseudocleft construction, prosodic marking as with information focus can also express identificational focus in the same context:

(66) (elicited. In a situation where only Mihail is holding a rifle:)

MIHAIL pəčal-əm kuč-a.

Mihail rifle-ACC hold-3SG

‘MIHAIL is holding the rifle.’

Similar facts are reported by Andreas Pregla (p.c.) who also describes that some instances of (postverbal) contrastive focus are associated with a reading of “this, and not something else” as well. Therefore, there is a syntactic construction specialized for expressing exhaustive focus, but other focussing strategies can express the same meanings as well. Constituent focus is therefore underspecified for the type of focus.

6.3 Focus negation

Single constituents can be negated with a negation particle *ogəl*. This negator transparently contains the inflected negation verb of the present 3SG and the copula in the connegative form (*og-əl* ‘NEG.3SG-be.CNG’), but does not show any

signs of agreement (see Georgieva, Salzmann, and Weisser 2021). It can occur as constituent negator in all clauses. A homophonous form is employed in compound past tense negation where it agrees with the subject of the clause. The constituent negator can negate most parts of the sentence, including NP in subject and object position, PP, and infinitives. Constituent negation appears in correction focus contexts, most productively in contrastive pairs, but also in isolation (Georgieva, Salzmann, and Weisser 2021: 490). Therefore, *ogəl* can be used as specialized focussing particle. Negated subjects can either appear after an object which is elided in a following clause, or in the sentence-initial position. Other constituents appear in the preverbal position.

- (67) (Georgieva, Salzmann, and Weisser 2021: 491)

Olma-m Təj ogəl, a PETR kočk-o.

apple-ACC 2SG NEG but Petr eat-PST1.3SG

‘Not YOU but PETER ate the apple.’

- (68) (elicited)

Təj ogəl olma-m kočk-əč, a PETR kočk-o.

2SG NEG apple-ACC eat-PST1.2SG but Peter eat-PST1.3SG

‘Not YOU but PETER ate the apple.’

- (69) (elicited. Context: In the room, is the man kicking a table?)

Uke, pölem-əšte pörjeŋ ÜSTEL-əM ogəl, PÜKEN-əM čumal-eš.

no room-INE man table-ACC NEG chair-ACC kick-3SG

‘No, in the room the man is not kicking a TABLE, he is kicking a CHAIR.’

- (70) (Georgieva, Salzmann, and Weisser 2021: 490)

Təj šAXMAT dene ogəl, a šAŠKE dene mod-əč.

2SG chess with NEG but checkers with play-PST1.2SG

‘You played not (with) CHESS but (with) CHECKERS.’

The different interpretational effects of clausal negation and constituent negation are illustrated by the following examples. Clausal negation reverses the truth

value of the whole sentence whereas constituent negation implies that the sentence is true of an alternative to the negated constituent. The preferred interpretation of a clause like example (72) is that the predicate holds of an alternative to the negated constituent – in the example not of the ghost but of a friend (the second clause in example (72) is optional and not necessary for this interpretation). Note also the difference with respect to the form of the lexical verb – in example (71) the verb ‘feed’ appears in the connegative form because the negative auxiliary is inflected whereas in example (72) the lexical verb is inflected.

(71) (elicited)

Anna končəmaš-əmə ok pukšo.
 Anna ghost-ACC NEG.3SG feed.CNG
 ‘Anna doesn’t feed a/the ghost.’

(72) (elicited)

Anna KONČƏMAŠ-ƏM ogəl pukš-a (– JOLTAŠ-ƏŽƏ-M pukš-a).
 Anna ghost-ACC NEG feed-3SG friend-PX3SG-ACC feed-3SG
 ‘Anna doesn’t feed a GHOST (– she feeds HER FRIEND).’

Apart from constituent negation with *ogəl*, single constituents can be prosodically marked for focus and interact with negation in the same way. Consider examples (73a–b). Sentence (73a) shows regular clausal negation with the whole sentence in the scope of negation. As mentioned above, the form *ogəl* is not a constituent negator here, but part of the verbal paradigm as negator of compound past tenses. Sentence (73b) on the other hand, shows that by focussing the pronoun *tudo*, negation targets only this pronoun. The clause implies that the sentence is true of a focus alternative. Whether there is any distributional or interpretational difference compared to constituent negation with *ogəl* in this case is a question for further research.

(73) (Saarinen 2015: 347)

- a. *Tudo məj-əm uʒ-ən og-əl.*
3SG 1SG-ACC see-CVB NEG.3SG-be.CNG
'S/he has not seen me.'
- b. *TUDO məj-əm uʒ-ən og-əl.*
3SG 1SG-ACC see-PST2.3SG NEG.3SG-be.CNG
'S/HE has not seen me (= but someone else has).'

6.4 Summary

This section described different types of focus in Mari: information focus (constituent focus as answer to *wh*-questions, full clausal focus, presentational focus, correction focus, verbal focus, focus in polar questions), identificational focus, and focus negation.

The main findings of this section about focus in Mari can be summarized as follows: Prosody plays an important role for focus marking and most instances of constituent focus can be marked in situ by prosody alone. This is true for subject, object, verbal and PP focus both as answer to a *wh*-question and in correction contexts. The preverbal position can be employed for focus as reported in the literature, but for constituents other than objects this strategy seems to be employed rather rarely in speech. The preverbal position appears to be a more marked position in any case, as placing subjects there after an all-new question is dispreferred. In polar questions, prosodic marking of a focussed constituent (except verbs) can be accompanied by the polar question particle directly following the constituent. Verbs are either focus marked prosodically in situ, or by having the verb precede other constituents in the clause. Identificational focus can be expressed by a special pseudocleft construction, but prosodic constituent focus is felicitous in the same context for identificational focus as well. Presentational focus differs from other constituent foci since the focussed constituent appears in a postverbal position there. There are no strategies specific of all focus sentences. Importantly, almost all types of focus discussed in this section do not require a

special morphosyntactic realization but can be marked solely by prosodic means. It is a matter of further research to see whether prosodic marking and syntactic marking of different foci have different interpretational effects, and what the division of labor between the strategies is.

With regard to focus and word order typology, Mari behaves like a rather typical SOV language. Although focussed non-objects are placed in a preverbal position only rarely, focussed constituents *can* appear in this position. In this regard, the preverbal position is what Buring (2009) calls a structurally “loaded” position, but not a fixed focus position. Furthermore, the possibility to focus the verb by having it precede other constituents is typical of SOV languages as well. In any case, Mari is not a strict position language with regard to focus placement, and prosody plays an important role for focus marking. Finally, the prediction by the markendess scale of structural complexity of focus marking (Skopeteas and Fanselow 2010) is borne out: regular constituent focus is simply marked in situ; the preverbal position likely has other restrictions as to when non-object constituents may appear there; and finally, pseudoclefts as structurally most complex type of focus marking are employed only for identificational focus.

7 Topics in Mari

This section deals with topics and phenomena related to (non-)topicality in Mari. The section has two main parts: Section 7.1 establishes the left periphery as a place for topical constituents using different diagnostics. In section 7.2, on the other hand, I turn the attention to the various reasons why constituents can be placed in the postverbal field, a position known to host backgrounded or non-topical constituents in SOV languages. Contrastive topics are discussed separately in Section 8.3.

7.1 The left periphery

The unmarked word order in Mari is SOV, and the sentence-initial position is the position for subjects in the unmarked clause where subject and aboutness-topic usually coincide. It is well known from the literature on Mari word order that

deviations from this word order are used for information structural marking (see Section 2.2). In this section I discuss the role of the left periphery as place for topics. I show that placing a constituent in the left periphery, and specifically in the sentence-initial position has a number of effects attributed to topicality:

1. Non-subjects as topics in a topic progression are preposed and located in the left periphery
2. Placing a constituent in the clause-initial position increases its acceptability as antecedent for a following null subject
3. The topics discussed in the first two subsections can be preceded by adverbials in the left periphery
4. The sentence topic can be changed with the help of a specific topic-marking phrase, *manmašte*, in the left periphery
5. New topics can be established via left dislocation

7.1.1 Non-subjects as topics

In a clause uttered in a neutral context, subjects usually occupy a position in the left periphery, and the subject functions as sentence topic as well. Therefore, the occurrence of subjects in a left-peripheral position alone does not provide any evidence that the left periphery is a position for topics. Other parts of the sentence, like objects, do provide this kind of evidence. In Mari we find that if a non-subject becomes the topic of a clause in the sense that the clause is “about” the non-subject, it can be placed in the sentence-initial position to express the progression to this new topic. In these cases, the topics necessarily have a corresponding gap in the clause, other than left dislocated topics (see Sections 7.1.4 and 7.1.5).

The following examples show topical accusative and dative objects in the sentence initial position. In example (74), the first sentence tells about what happened after a concert and introduces ‘200 000 rubles’ (underlined). The next sentence describes what happened to that money, and the accusative marked phrase “this money” occupies the sentence-initial position before the subject of

the clause, in an OSV order. Another short stretch of discourse showing the same phenomenon is shown in example (75). In this example, a festival is introduced, and as a topic in the next clause it occupies the position before the subject.⁸

(74) CCWLM, Journal *Kidšer* (26.03.2016)

Kodšo ij-ən murəzo Svetlana Osipova Seŋəmašən XXX idaləkše
 last year-INS singer Svetlana Osipova Culture Palace of the 30th
 lümeš tüvəra polat-əšte, *poro kuməl-əm*
 Anniversary of the Victory in Europe-INE goodwill-ACC

ončəktə-šo koncert-əm ertar-en ələ. Tunam, koncert
 show-PTCP.ACT concert-ACC carry.out-CVB AUX then concert
ertarə-me dene kok šüdö tüzem tenge pogən-en.
 carry.out-PTCP.PASS with 200 000 rubles be.gathered-PST2.3SG

Tide šijvundə-m murəzo onkologij čer dene čerlanə-še
 this money-ACC singer oncology illness with be.sick-PTCP.ACT
joča-vlak-əm eml-aš kusar-en.
 child-PL-ACC cure-INF donate-PST2.3SG

‘Last year, singer Svetlana Osipova played a charity concert in the *Culture Palace of the 30th Anniversary of the Victory in Europe*. In the course of the concert 200 000 rubles were gathered. **This money**, the singer donated to the treatment of children with cancer.’

⁸Note that there is no passive in the Mari example, so the preposed object is in fact there due to the topic progression and not due to a passive.

(75) CCWLM, Journal *Kidšer* (12.04.2018)

Kugeče kečə-n Joškar-Ola-šte 'Kugečə-lan u tuvər' nacional'nəj
easter day-INS Joškar-Ola-INE easter-DAT new shirt national

kostjum festival' ikəmše gana ert-en. Tide pajrem-əm
costume festival first time pass-PST2.3SG this celebration-ACC

Marij tür rüder den Marij jumənjüla ušem-ən
Mari embroidery center with Mari jumənjüla organization-GEN

Joškar-Ola-se obščinə-že organizovatl-en-ət.
Joškar-Ola-ADJ commune-PX3SG organize-PST2-3PL

‘On Easter Sunday the national costume festival ‘A new shirt for Easter’ was held in Joškar-Ola for the first time. **This celebration** was organized by the Mari embroidery center and the Joškar Ola commune of the Mari *jumənjüla* folk belief organization.’

The following example shows a dative object in sentence-initial position preceding a nominative subject:

(76) CCWLM, Newspaper *Kugaría* (02.09.2016)

Aleksandr Kuznecov vič ikšəv-an ješ-əšte Joškar-Ola-šte
 Aleksandr Kuznecov five child-ADJ family-INE Joškar-Ola-INE
šoč-ən-kuš-ən. Ača-že Ivan Mixajlovič
 be.born-PST2.3SG-grow-PST2.3SG father-PX.3SG Ivan Mixajlovič
20 ij utla Marij radiomexaničeskij texnikum-əm
 20 year more Mari radiomechanical technical.college-ACC
vujlat-en, respublikə-šte volejbol-əm vijəŋdəmaške kugu
 lead-PST2.3SG republic-INE volleyball-ACC development.ILL big
nadər-əm pəšt-en. [...] Aleksandr-lan ača-že
 contribution-ACC put-PST2.3SG Aleksandr-DAT father-PX3SG
primer lij-ən, tudo iziñek volejbol dene
 role.model be-PST2.3SG 3SG from.childhood volleyball with
mod-ən [...].
 play-PST2.3SG

‘Aleksandr Kuznecov was born and grew up in a family of five children in Joškar-Ola. His father Ivan Mixajlovič led the Mari radiomechanical technical college for more than twenty years and made a great contribution to the development of volleyball in the republic. [...] To Aleksandr, his father was a role model, he played volleyball from childhood [...].’

Another case is presented in example (77). The text introduces ‘other books than textbooks’, and in the next sentences, and the new sentence topic is a *subset* of those books (‘belletristic books’; ‘these books’):

(77) CCWLM, Newspaper *Marij El* (27.05.2005)

Avtor čən palemd-a: pətartəš 10-15 ij-əšte škol-əško
 author correctly remark-3SG previous 10-15 year-INE school-ILL
učebnik deč molo kniga-že, čən-ak, og-eš-ak
 textbook from other book-PX3SG really-EMPH NEG-3SG-EMPH
tol. Amal-že pale: səlnəmut-an kniga-m avtor-vlak
 come.CNG reason-3SG clear literature-ADJ book-ACC author-PL
ške kūšeš aļe sponsor polšəmo dene vele lukt-ət. A
 own at.expense or sponsor help with only publish-3PL but
təgaj knig-am pisatel'-vlak en ondak užal-aš tərš-at, a
 such.kind.of book-ACC writer-PL SUP before sell-INF try-3PL but
pölek šotəšto bibliotekə-la-ške pu-aš kuməl-əšt uke.
 gift thought.INE library-PL-ILL give-INF mood-PX3PL NEG.EXIST

‘The author correctly remarks: in the last 10-15 years basically no books other than textbooks have come into the schools. There is an apparent reason for this: **Belletristic books** the authors only publish on their own expense or with the support of sponsors. And **such books**, the writers first of all try to sell, and they are not inclined to donate them to libraries.’

7.1.2 Antecedents for null subjects

Mari is a consistent null subject language, which means that null subjects can occur in any person-number combination, tense, mood, and clause type (Bradley and Hirvonen 2022). The question of how null subject reference in different languages is resolved is subject to ongoing research (see the papers in the volume by Dalmi, Tsedryk, and Cegłowski 2022). One approach to null subject licensing and interpretation ties (third person) null subjects to the topichood of entities in the discourse (Frascarelli 2018; Holmberg 2010). In simple terms, proponents of this approach argue that in consistent null subject languages, null subjects are resolved with respect to the Aboutness Topic. A comparable diagnostic for topics was already proposed by Reinhart (1981: 70) as well.

On the basis of the approach to null subject licensing via topichood, Bradley and Hirvonen (2022) tested the interpretation of 3sg null subjects in Mari where possible overt antecedents varied with regard to their position in a preceding clause. Consider examples (78–79).

(78) (Bradley and Hirvonen 2022: 299)

Jelu₁ tengeče Serge₂ dene kutər-en. ande Ø_{1/2} saj-ən
Jelu yesterday Serge with speak-PST2.3SG now good-INS
mal-en kert-eš.
sleep-CVB can-3SG

‘Yesterday Jelu₁ talked with Serge₂. Now s/he_{1/2} can sleep well.’

(79) (Bradley and Hirvonen 2022: 299)

Serge₂ dene Jelu₁ tengeče kutər-en. ande Ø_{1/2} saj-ən
Serge with Jelu yesterday speak-PST2.3SG now good-INS
mal-en kert-eš.
sleep-CVB can-3SG

‘With Serge₂, Jelu₁ talked yesterday. Now s/he_{1/2} can sleep well.’

Both examples involve a null subject in the second clause, and two possible overt antecedents (*Jelu* and *Serge*) in a preceding clause. The context clauses only differ in their word order – in example (78), subject and clause-initial constituent coincide in *Jelu*, but in example (79), the non-subject PP *Serge dene* ‘with Serge’ is the clause initial constituent. The respondents of an online survey were asked about their preferred interpretation of the null subject and were asked whether they think Jelu, Serge, or someone else can sleep well (with multiple answers permitted). In example (78), 79.1% preferred a reading of the null pronoun coreferent with *Jelu* and 20.9% with *Serge*, whereas in example (79), only 54.3% preferred the reading coreferent with *Jelu*, 37% with *Serge*, and 8.7% with someone else altogether. The sentence was judged ambiguous by 15.2% of participants who chose both Jelu and Serge as possible antecedents, whereas this was only the case with 10.3% in the first sentence. These results indicate that the interpretation of 3sg

person null subjects is dependent on multiple factors. Subjecthood seems to be the strongest factor to make a constituent viable as antecedent for a following null subject, but topicness is another factor to be taken into account – since in example (79), the clause-initial constituent *Serge dene* was (to a small degree) more likely chosen as antecedent for the 3SG null subject than in (78), the clause initial position is likely a topic position, following Frascarelli’ and Holmberg’s theory of null subject resolution.

It should also be noted that subject pronouns are not the only topical constituent that can be continued by null pronouns. In the following example, a girl is the topic in the first clause. The following clause is about what her father did, and the girl is continued as implied object with a null pronoun. In the last clause, both the girl and the father are referred to by null pronouns. The presence of both causer and causee is implied by the causative verb *əštəkte-* ‘do-CAUS’. The extent of this phenomenon is unclear to me.

- (80) Newspaper *Marij El* (<https://газетамарийэл.рф/?p=11962>, accessed 29.06.2023)

Ješ-əšte üdər joča-vlak kokla gəč tudo₁ en kugu-žo
family-INE girl child-PL among from 3SG SUP tall-PX3SG

lij-ən, sandene ača-že₂ Ø₁ škol-əš tunem-aš
be-PST2.3SG therefore father-PX3SG school-ILL learn-INE

kolt-en og-əl, Ø₂ Ø₁ surt koklašte paša-m
send-CVB NEG.3SG-be.CNG farm amidst work-ACC

əštə-kt-en.
do-CAUS-PST2.3SG

‘In the family, among the girls, she₁ was the tallest one, therefore her father₂ did not send (her₁) to school, (he₂) made (her₁) work at the farm.’

7.1.3 Adverbials in the left periphery

Topics in the left periphery can be preceded by adverbs or PP. Already Chkhaidze (1941: 30) noted that time and place adverbials can precede subjects in a neutral

clause. In example (47), repeated here as (81), the inessive marked *sadāšte* ‘in the garden’ precedes the sentence topic and subject *ūdāramaš* ‘the woman’.

(81) (elicited. In the garden, is the girl hitting a girl?)

Uke, sad-āšte ūdāramaš rvezā-m šel-eš.

no garden-INE woman boy-ACC hit-3SG

‘No, in the garden the woman is hitting a boy.’

Such adverbials can precede fronted non-subject topics as well. In the first sentence of example (82) an organization is introduced. As sub-part of the organizational work, the ‘director’s work’ is taken up as sentence topic, again as accusative object in the initial position. In the last sentence, the referent is continued as topic in initial position before the subject. This is shown by usage of the proximal demonstrative *tə*, referring to the work, as well. The sentence topic here is preceded by the adverbial *tačəse kečəlan*.

(82) CCWLM, Journal *Kidšer* (29.03.2016)

Tide tale üdaramaš da Jurij Tanəgin-ən təršəmə-ž-lan
 this talented woman and Jurij Tanəgin-GEN effort-PX3SG-DAT
köra 2005 ij 9 dekabır'-əšte HMAO-Jugra-šte kuməšo marij
 because 2005 year 9 december-INE KhMAO-Yugra-INE third Mari
mer organizacij čumərg-en. Vujlatə-me paša-m
 social organization be.formed-PST2.3SG lead-PTCP.PASS work-ACC
Jurij Tanəgin tünal-ən, ikmənar žap gəč tudə-m Larisa
 Jurij Tanəgin begin-PST2.3SG some time from 3SG-ACC Larisa
Timčenko almašt-en. Tačə-se kečə-lan tə soməl-əm
 Timčenko replace-PST2.3SG today-ADJ day-DAT this duty-ACC
murəzo Šakmaeva Svetlana šukt-a.
 singer Šakmaeva Svetlana lead-3SG

‘Thanks to the efforts of this skilled woman and Jurij Tanəgin, on 9th December 2005 in the Khanty-Mansi Autonomous Okrug – Yugra, the third Mari social organization was formed. **The director’s work** was begun by Jurij Tanəgin, after some time he was replaced by Larisa Timčenko. Today, **these duties** are fulfilled by the singer Svetlana Šakmaeva.’

Note also example (77) where the time and place adverbials *pətartəš 10-15 ijəšte* ‘in the past 10–15 years’ and *školəško* ‘into the schools’ precede the topic, and example (112), in which a PP precedes the sentence topic.

7.1.4 Topic marking phrases

The topic can be changed in a clause by a designated phrase for topic marking – *manmašte* (*man-ma-šte* ‘say-PTCP-INE’) ‘concerning, as regards, about’. This strategy is not too commonly employed, but is an ideal diagnostic due to its unambiguous overt nature of topic marking. According to my consultant, the phrase is less acceptable with concrete entities or things and more acceptable with abstract topics like situations or immaterial things. But even if the *manmašte* test does not test for all topics, it can help to diagnose for at least some of them. Recall

also the remark by Reinhart (1981: 64) regarding the *as for* test: “when this test is applicable, it is clear that the NP in question can be the topic. But when it is not, we cannot, with certainty, infer the opposite.”

The topics used with *manmašte* are new sentence topics and can make recourse to something which was introduced into the discourse before (marked by underlining if relevant). Together with the phrase which *manmašte* modifies, *manmašte* itself usually occupies the very first position in a clause. These topics may be fronted as part of the clause without intonational breaks or resumptive elements (ex. 83), or – more frequently – appear outside of the core clause as left dislocated topics, (ex. 84–86). In the latter function there can be a resumptive element as well: in example (86), the 3SG pronoun *tudo* refers back to the dislocated topic:⁹

⁹In example (86), the title of the play, “Orjeṇ melna”, refers to the pancakes traditionally baked by the bride on the first day after the wedding.

(83) (elicited)

Tide situacija manmašte māj nīm-at kalas-en o-m
this situation about 1SG nothing-ADD tell-CVB NEG-1SG
kert.
can.CNG

‘As regards this situation I can’t tell you anything.’

(84) CCWLM, Newspaper *Marij El* (01.09.2007)

Miġa-n paša gəč pörtəl-meškə-že, təge jol ümbalne
Miġa-GEN work from return-CVB.POST-PX3SG such foot on
šogəlt-eš, ig-an čəvə-la pörd-eš. Paša
busy.oneself-3SG offspring-ADJ chicken-COMP turn-3SG work
manmašte, Sergej nīmənar-at jara ok šinče,
about Sergej no.amount.of-ADD idle NEG.3SG sit.CVB
püken-la-m, kuxnə-la-n škaf-əm əštəl-eš.
chair-PL-ACC kitchen-PL-GEN cupboard-ACC do-3SG

‘Before Miġa returns from work, she’s busying herself, bustling around like a chicken with younglings. **Speaking about work**, Sergej doesn’t sit around idly either, he’s building chairs, and kitchen cupboards.’

(85) CCWLM, Newspaper *Marij El* (14.10.2003)

Čən, tide tühəltəš oškəl vele, no čəple tühəltəš. ... Kniga-šte
true this beginning step only but excellent beginning book-INE
čəla saj manmašte, šinča-lan pernə-še ekšək-əm-at
all good about eye-DAT strike-PTCP.ACT mistake-ACC-ADD
ončəktə-de og-eš lij, šon-em.
show-CVB.NEG NEG-3SG be.CNG think-1SG

‘True, this is only an initial step, but an excellent one at that. ... **Speaking about that all is in the book**, I think it will not go without pointing out some glaring mistakes as well.’

(86) CCWLM, Newspaper *Marij El* (04.06.2003)

– [...] A molan ‘*Salika-m*’, ‘*Orjeṇ melna-m*’, tošto repertuar gəč
 but why Salika-ACC bride pancake-ACC old repertoire from
molo spektakl-əm og-ət kondo? Posnak ilalše-vlak
 other performance-ACC NEG-3PL bring.CNG especially elderly-PL
turgəžlan-at. – **Salika manmašte, tudo** scena gəč ok
 worry-3PL Salika about 3SG scene from NEG.3SG
korəṇdalt, möṅgešla.
 be.removed.CNG on.the.contrary

‘– But why don’t they continue ‘*Salika*’, ‘The bride’s pancake’ or other performances from the old repertoire? Especially the elderly are worried about this. – As for ‘*Salika*’, it will not be removed from the scene, on the contrary.’

7.1.5 Left dislocation

Topics in Mari can occur in the left periphery outside of the core clause as result of left dislocation. In the previous section about the topic marking phrase *manmašte* ‘concerning, as regards, about’ we saw some examples for this type of topic already. Left dislocated topics can appear without *manmašte*, too. In the clause following the topic there may be a resumptive element (*tušto* in ex. 87, *təšte* in ex. 88), or not (ex. 89). In the latter example it can be clearly seen that the topic is not part of the core clause: the topic is in the nominative case like a subject, but the verb does not agree with it. Instead, the verb shows 3PL morphology which is typical for generic statements (cf. Bradley and Hirvonen 2022).

(87) (elicited)

Kande kava, tušto üdəramaš pörjeṇ-əm šel-eš.
 blue sky there woman man-ACC hit-3SG

‘The blue sky, a woman is hitting the man there.’

- (88) SCMM, Биография А. Н. Ф. (часть 1)

A tunam jal-əštə-že, kö təšte en pervəj televizor-žo,
 and then village-INE-PX3SG who here SUP first television-PX3SG
kö-n lij-ən əle?
 who-GEN be-PST2.3SG AUX

‘And back then in the village, who had the first television here, whose was it?’

- (89) (Serdobol’skaya and Toldova 2012: 731)

Marij televidenij-že tidə-m-ak vele ončəkt-at kuze ubiratl-at.
 Mari television-PX3SG this-ACC-EMPH only show-3PL how harvest-3PL
 ‘Mari television, they only show such stuff, how they harvest crops.’

7.1.6 Summary

This section discussed the left periphery position as position for topics in Mari. Generally, sentence topics occupy the clause initial position, which is easy to observe when topical non-subjects precede subjects, e.g. in an OSV word order. These fronted topics can also be preceded by adverbials. Placing a non-subject constituent in the clause-initial position slightly facilitates its selection as possible antecedent for a null subject in the following clause, which is predicted by accounts which explain the resolution of null subject reference with topichood of a possible antecedent. The topic can also be changed by an overt topic-marking phrase like *manmašte* ‘concerning, as regards, about’ which, together with the new topic, occupies a left-peripheral position. These phrases are typically left dislocated topics that are not part of the core clause. Left dislocated topics can also occur without *manmašte*. Summarizing, we note that the left periphery is the field where topics are placed. The next section deals with the postverbal field, and the various constituents which are found in that position. These include familiar topics as well.

7.2 Postverbal elements

Above I argued that the left periphery, among which is the clause-initial position, is the field for topical constituents, while the preverbal position can be a position for focal constituents. Other prefocal elements do not seem to receive a specific information structural interpretation in sentences with multiple constituents in the prefocal field (Pregla 2023: 89). Leaving out the intricacies of positional variants in this field, I will examine another clausal field in some detail, namely the postverbal field. This is a purely descriptive term and does not imply a certain analysis for a derivation of these word orders. The postverbal position serves a number of different purposes, which I will discuss one after another:

1. In the spoken language, afterthoughts and completions of the preceding clause appear postverbally
2. Familiar topics may appear in the postverbal position
3. Constituents can “altruistically” appear in the postverbal position to allow focussing of the verb or another preverbal constituent
4. Prosodically heavy constituents can be extraposed to a right-peripheral position
5. Some usages of postverbal constituents are typical for certain genres

Some scholars have noted a tendency towards a neutral SVO order in spoken Mari, contrary to the rather conservative written SOV standard (see section 4.3), so it is unclear how register specific the following data are.

7.2.1 Afterthoughts and completions

Typical for spoken language, the postverbal position serves as position for afterthoughts or completions of the preceding clause. These constituents usually add some additional information to the content of the clause. They do not necessarily require an overt resumptive element like the demonstrative *tide* in example (89), repeated here as (92); the correlate of a subject in the main clause has the option to be null (ex. 90, 132, 133) as Mari is a consistent null subject language

(see section 7.1.2 and Bradley and Hirvonen 2022). These postverbal elements are presumably deaccented or realized with an intonational break, but information on prosody was not marked in the data.

- (90) (Serdobol'skaya and Toldova 2012: 739) (It was Z.'s wedding but she didn't invite us. She invited nobody.)

Ńigö-m üž-ən og-əl ala-mo tægaj üdər.
nobody-ACC invite-CVB NEG.3SG-be.CNG INDEF-what such girl

‘She didn't invite anybody, whatever, that girl.’

- (91) (Serdobol'skaya and Toldova 2012: 740) (You're never coming along!)

Ĵal-aštə-že ala-gušto il-et kilometr tora-šte
village-INE-PX3SG INDEF-where live-2SG kilometer distance-INE

memnan deč.

1PL.GEN from

‘You live in the countryside somewhere, kilometers away from us.’

- (92) (Serdobol'skaya and Toldova 2012: 731) (They showed on TV how they make hay, and some other stuff.)

Marij televidenij-že tidə-m-ak vele ončəkt-at kuze ubiratl-at.
Mari television-PX3SG this-ACC-EMPH only show-3PL how harvest-3PL

‘Mari television, they only show such stuff, how they harvest crops.’

7.2.2 Familiar topics

The postverbal position can also host familiar topics which were aboutness-topics in preceding clauses. In examples (93) and (94) the postverbal constituent in both cases refers to an entity which was previously introduced – *Kuženər* in ex. (93)¹⁰ and the bridge in ex. (94).

¹⁰Note that in this example, the order ADD-GEN in the forms *S.-atən* and *kö-t-ən* is unexpected as *-at* is a clitic which attaches after case endings. I do not have an explanation for this order.

(93) (Serdobol'skaya and Toldova 2012: 728)

A. *Kužeŋer-že ške-že poselok jöršan saj ogəl.*
 Kužeŋer-PX3SG self-PX3SG village exceedingly good NEG

‘Kužeŋer itself as a village is not nice at all.’

B. *Aj [...] mij-en-na ələ S.-at-ən kö-t-ən*
 oh go-PST2-1PL AUX S.-ADD-GEN who-ADD-GEN

süan-əškə-že Kužeŋer-əš?
 wedding-ILL-PX3SG Kužeŋer-ILL

‘Oh, didn’t we go to S.’s, (or) to whose wedding, to Kužeŋer?’

(94) (Serdobol'skaya and Toldova 2012: 711)

Jadək-Belak küvar jəmake pur-en šinč-ən-na, ənde
 Jadək-Belak bridge to.under enter-CVB sit.down-PST2-1PL now

rveze-šaməč-et lišem-ət peš-ak memnan küvar
 boy-PL-PX2SG approach-3PL very-EMPH 1PL.GEN bridge

voktek-et.

to.besides-PX2SG

‘We went under the bridge at Jadək-Belak and sat down, now the boys are getting closer, very much towards our bridge.’

New aboutness topics are illicit in a right dislocated position:

(95) (elicited. Have you heard anything about Serge?)

**Pialeš, u paša-m mu-ən, Serge.*
 fortunately new work-ACC find-PST2.3SG Serge

‘Fortunately, (he) has found a new job, Serge.’

7.2.3 Altruistic postverbal elements

Constituents may appear in the postverbal position for “altruistic” reasons if focus is on the finite verb (see section 6.1.5) or on a remaining preverbal constituent

(see section 6.1.1). Verbal focus and preverbal focus via verb fronting appear to be results of the same operation. The postverbal field is then filled by the other constituents which are not part of the focus (Pregla 2023). The postverbal placement of constituents after various kinds of verb fronting is therefore likely epiphenomenal and not directly connected to any information structural status of the postverbal constituents themselves. The relevant examples from previous sections are repeated here.

Verb fronting can well be illustrated by sentences with complex verb clusters involving modal or auxiliary verbs and the inflected negation auxiliary (cf. Georgieva, Salzmann, and Weisser 2021). Modal verbs like *kert-* ‘can’ select a converbial form of the lexical verb, e.g. *əšten* ‘do.CVB’. The inflected negation auxiliary in turn triggers the connegative marking on the modal verb. The unmarked linear order in such a verb complex is lexical verb – negation auxiliary – modal verb, and the whole complex appears in sentence-final position (ex. 96). Crucially, the negation auxiliary and the modal verb can only move together and have to stay adjacent, whereas they can be separated from the lexical verb (Georgieva, Salzmann, and Weisser 2021). If the finite negation auxiliary is fronted for focussing purposes, the modal verb follows it to the sentence initial position. Compared to the unmarked order, all other constituents (subject, object, adverb and lexical verb) end up in the postverbal field (ex. 97):

(96) (elicited)

Məj tidə-m ɲiguze əšt-en o-m kert.
 1SG this-ACC no.way do-CVB NEG-1SG can.CNG
 ‘I cannot do this in any way.’

(97) CLM, News text

O-M KERT məj tidə-m ɲiguze əšt-en.
 NEG-1SG can.CNG 1SG this-ACC no.way do-CVB
 ‘I CANNOT do this in any way.’

In the case of verbal focus via fronting, the verb does not necessarily have to be the leftmost constituent, however. Consider example (98). The sentence in (98a)

does not have a canonical subject and the sentence-initial position is taken by a place adverbial. The auxiliary *lij-* ‘be, become’ selects an infinitive complement (*šupšaš*) which precedes the auxiliary in the unmarked order. When the order of auxiliary and VP is reversed (ex. (98b), it is the negation that is stressed, not the place adverbial, even though it appears in the preverbal position. Place adverbials can precede other constituents fronted for reasons of topicality (see Section 7.1.3), so this is in principle expected. In this example, my consultant also prosodically stressed the finite verbal complex.

(98) a. (Riese et al. 2019: §7.3.3.3)

Təšte tamak-əm šupš-aš ok lij.
 here tobacco-ACC suck-INF NEG.3SG be.CNG

‘You can’t smoke here.’

b. (elicited, Context: Someone argues it is allowed to smoke here.)

A təšte OK LIJ tamak-əm šupš-aš.
 but here NEG.3SG be.CNG tobacco-ACC suck-INF

‘But you CAN NOT smoke here.’

The same kind of dislocation is visible as effect of preverbal focus via verb fronting; other constituents are left behind in a postverbal field. In example (99b), the effects of focussing the adverbials ‘yesterday morning’ via verb fronting are clearly visible: the other constituents appear postverbally and preserve the ordering of the unmarked clause (99a).

(99) (Pregla 2023: 138)

a. *Poškudo tengeče erdene joča-lan pört-əštö pərəs-əm*
 neighbor yesterday morning child-DAT house-INE cat-ACC

puə-š.
 give-PST1.3SG

‘The neighbor gave a cat to the child in the house yesterday morning.’

- b. *Poškudo TENGEČE ERDENE puə-š joča-lan pört-əštö*
 neighbor yesterday morning give-PST1.3SG child-DAT house-INE
pəras-am.
 cat-ACC

‘The neighbor gave a cat to the child in the house YESTERDAY MORNING.’

7.2.4 Heavy XP shift

Some constituents seem to appear in the postverbal position without discernible information structural effect due to their heavy prosodic weight. These include heavy adverbials and lists:

(100) CCWLM, Newspaper *Marij El* (28.03.2007)

Ves marij mer pašajeŋ teve tægaj šonəmaš-əm
 another Mari social worker well such.kind.of thought-ACC
lukt-ən kalasə-š: 'Polučaets'a täge, čto memnan
 lead.out-CVB speak-PST1.3SG it.seems such that 1PL.GEN
predok-na-vlak peš ušan lij-ən-ət.' Rušla-marla
 ancestor-PX1PL-PL very intelligent be-PST2-3PL in.Russian-in.Mari
kuktəštmo gəč mo-m-gənat uməl-en-da, šon-em.
 mixture from what-ACC-INDEF understand-PST2-2PL think-1SG
Tægaj-rak mutlanəmaš ert-en marij kalək-na-n
 such.kind.of-CMPR talk go-PST2.3SG Mari people-PX1PL-GEN
tüŋ mer organizacij-žə-n jubilej šot-an
 main social organization-PX3SG-GEN jubilarly thought-ADJ
konferencij-əštə-že.
 conference-INE-PX3SG

‘Well, another Mari social worker put out such a thought: [russified Mari]
 ”It seems like our ancestors were very intelligent.” I guess you have un-
 derstood something from this Russian-Mari mess. Such was the talk that
 went in the jubilarly conference of our main Mari community organiza-
 tion.’

(101) CCWLM, Newspaper *Marij El* (16.07.2005)

Mutat uke, tide otrasl-lan vijay-aš-əže kundem-əštə-na
of course this department-DAT strengthen-INF-PX3SG land-INE-PX1PL
kugu ver ojaralt-ən da ončəkə-žə-m-at
big place be.allocated-PST2.3SG and in.future-PX3SG-ACC-ADD
lij-eš. Təšk-ak pur-at kalək-vlak kokla-se kəl,
be-3SG to.here-EMPH enter-3PL people-PL among-ADJ tie
massovəj informacij sredstvo-vlak, nauka, iskusstvo, mer
mass information means-PL science arts social
organizacij-vlak-ən paša-št da mol-at.
organization-PL-GEN work-PX3PL and other-ADD

‘Of course, in order to strengthen the status of this department it has been allocated an important position in our country, and will have this status in the future as well. It will include **international relations, mass media, science, arts, community work et cetera.**’

7.2.5 Genre-specific usage

A large number of the corpus examples with postverbal nominal constituents were attributable to genre specific usage like in matrix clauses that were postposed after embedded quotes in newspaper articles (ex. 102) or in biographies (ex. 103).

(102) CCWLM, Newspaper *Marij El* (02.02.2005)

Surt kajək da sösna šal-əm šukə-rak əšt-en
 farm bird and pig meat-ACC much-CMPR make-CVB
nal-šaš-lan ul-šo čəla kuat-əm-at paša-š
 take-INF.FUT-DAT be-PTCP.ACT all force-ACC-ADD work-LAT
ušə-man, kalas-en ministr.
 join-NEC say-PST2.3SG minister

‘In order to yield more poultry and pork all forces have to be adduced to work, said the minister.’

(103) CCWLM, Newspaper *Marij El* (26.05.2007)

Təgak lij-ən doktor dissertacij-əm aralə-me
 also be-PST2.3SG doctor dissertation-ACC defend-PTCP.PASS
deč vara-t.
 after-ADD

‘So it happened after defending her doctoral dissertation.’

7.2.6 Summary

In this section I investigated the reasons for placing constituents in the postverbal position. Some of the configurations have a clear information structural motivation which boils down to the postverbal field as being a position for non-prominent constituents. This became apparent from the examples for afterthoughts in section 7.2.1 and for familiar, given topics in section 7.2.2. Similarly, constituents may appear in the postverbal field to allow focussing of *another* constituent like the verb or a preverbal constituent. Finally, the examples with heavy XP shift in section 7.2.4 and genre-specific usage in section 7.2.5 showed no clear information structural motivation for placing constituents in a postverbal position. What was called the “postverbal field” in this section is to be understood as purely descriptive term, in any case. The various unrelated reasons for placing constituents in the postverbal field suggest that these phenomena are not

a unified class, and left peripheral clause-internal constituents are syntactically different from left dislocated constituents.

Summarizing all of section 7, we note that the left periphery and sentence-initial position is the field for aboutness topics. The postverbal position, on the other hand, is reserved for non-prominent constituents placed there for various reasons. The position hosts backgrounded topics and afterthoughts as well as extraposed, heavy constituents.

It is clear that this list provides only a first overview about various reasons for constituent placement related to information structure in Mari, and many questions for subsequent research remain open. These concern the middle field, for example: the middle field positions between the initial constituent and the verb can variably be filled by focussed constituents or any other constituent, and the intricacies of constituent order in this field are a topic for additional research. The other main concern is the prosodic realization of the various postverbal elements – are they prosodically detached from the main clause or not? Is a resumptive element in the main clause necessary?

This initial survey also leaves out more complex cases like the following. In example (104), observe a sentence with VOS order. Here, the verb is preposed whereas the subject *Marij samərək teatr* follows the object and occupies the clause-final position. This constituent subsequently becomes the new topic. The verb likely precedes the object to distinguish the construction from VP focus with OV order, or it is focussed itself. It is unclear however, what causes the OS order – the subject could be a presentational focus. In any case, these are the questions for further research on Meadow Mari information structure.

(104) CSMM, Newspaper *Marij El* (30.04.2019)

Kreml'-əšt-ak 'Məj marla ojl-em' akcij ert-en.

Kremlin-INE-EMPH 1SG in.Mari speak-1SG event go.by-PST2.3SG

Koklašte marij estradə-n murəzo-vlak ške mastarlək-əšt
among.others Mari estrade-GEN singer-PL own mastery-PX3PL

dene kuməl-əm nölt-en-ət. A mučəšl-en pajrem
with spirit-ACC raise-PST2-3PL but finish-PST2.3SG party

programmə-m Marij samərək teatr. Artist-vlak Z. Dolgova-n
program-ACC Mari youth theater artist-PL Z. Dolgova-GEN

‘Neməč vate’ p’esə-ž počəš šəndə-me postanovkə-m
German wife play-PX3SG after put-PTCP.PASS staging-ACC

ončəkt-en-ət.

show-PST2-3PL

‘Right there in the Kreml’ the event ‘I speak Mari’ was held. Among others, the Mari estrade singers raised the mood with their mastery. But it was the Mari youth theater that finished the festival program. The artists showed a staging based on Z. Dolgova’s play ‘The German wife’.

A similar problem is posed by example (45), repeated here with more discourse context as example (105). In this example, the referential distance between the first introduction of a referent and its subsequent appearance as a possible familiar topic in the postverbal position can be considerably large. The following example is from a text presenting an excellent student, Galina Alekseevna. In the text, the student is mentioned many sentences before the clause in question and constitutes a discourse topic. After describing the motivation for her work, the student is mentioned again in a postverbal position. In a similar manner to the previous example, the postverbal position still allows a sentential topic shift since a null subject is employed for reference to the student. In section 6.1.3 I argued that this is a case of presentational focus. The question how these structures interact are matters for further research.

(105) CSMM, Newspaper *Marij El* (27.05.2006)

Čüčkədən-ak ača-ava-n marla peleštə-mə-št-lan-at
 often-EMPH father-mother-GEN in.Mari speak-PTCP.PASS-PX3PL-DAT-ADD
üdər-rveze-vlak rušla vašəšt-at. Tə ojərtem-əm
 daughter-son-PL in.Russian answer-3PL this peculiarity-ACC
šot-əš nal-ən paša-m əšt-a Galina Alekseevna. Šočmo
 count-ILL take-CVB work-ACC do-3SG Galina Alekseevna native
mut-ən jam-žə-m urok-əšto da utla-rak-še
 word-GEN beauty-PX3SG-ACC lesson-INE and more-CMPR-PX3SG
urok deč vara-se ješartəš paša-m šukt-əmo
 lesson from after-ADJ additional work-ACC conduct-PTCP.PASS
godəm šiž-aš tunəkt-a.
 during feel-INF teach-3SG

‘Often, even when their parents talk to them in Mari, the kids answer in Russian. It’s Galina Alekseevna working, taking this peculiarity into account. She teaches to perceive the beauty of the mother tongue in lessons and especially when she does extra work after lessons.’

8 The 3SG Possessive Suffix *žE* as Marker of Contrast and Topic Shift

8.1 Introduction

Mari employs possessive suffixes for a variety of functions (see e.g. Alhoniemi 1993: 75–76). The prototypic function of a possessive suffix in Mari is to mark affiliation between entities, especially various kinds of possession. In this latter usage, possessive suffixes mark person and number of the possessor on the possessed (but the possessor usually remains covert in neutral sentences). In example (106a), the features of the 1SG possessor are marked with the *m*-suffix on ‘father’, and the father’s 3SG feature is marked with the *žə*-suffix on ‘daughter’, and so

forth. Besides the usage as marker of possession, possessive suffixes also mark person and number of a postposition's complement on the postposition, or the agent in non-finite clauses (ex. 107). The possessive suffix usually precedes suffixes of structural cases (nominative, genitive, accusative) but follows suffixes of local cases (illative, lative) (see Guseva and Weisser 2018; Hammer and Bradley 2022; Luutonen 1997 for discussion).

(106) (Riese et al. 2019: §2.3.2)

- a. *Məjən joltaš-em – paləme süretče.*
 1SG.GEN friend-PX1SG famous artist
 'My friend is a famous artist.'
- b. *Teve təšte memnan pört-na.*
 see here 1SG.GEN house-PX1PL
 'Here is our house.'

(107) (Bradley and Hirvonen 2022: 302)

- Təj-ən kajə-mek-et, pel ij gəč üdər-na*
 2SG-GEN go-CVB.ANT-PX2SG half year from daughter-PX1PL
šoç-o.
 be.born-PST1.3SG
 'After you left, our daughter was born half a year later.'

Besides possessive marking in a narrow sense as shown in the examples above, some possessive suffixes are used in a range of other discursive functions, most prominently the 3SG suffix, and less so the 1SG, 2SG and 3PL suffixes (Tuzharov 1987: 65–71). For an illustrative example with the 3SG possessive suffix in a discursive function, see ex. (108). Even without further context it is rather evident that there is no canonic possession involved in this example because the suffix appears on the nouns 'nature' and 'circumstance'. Instead, the suffix fulfills a discursive function, in this case contrasting those two nouns. This section is about the discursive usages of the 3SG possessive suffix.

(108) (Serdobol'skaya and Toldova 2012: 728)

No pürtüs-še peš saj, a uslovij-že saj ogəl
 well nature-PX3SG very good but circumstance-PX3SG good NEG
il-aš.
 live-INF

‘Well, nature is very pretty (there), but the living conditions are not.’

As for its phonetic realization, the suffix has a number of allomorphs that appear as a result of processes like voicing assimilation (e.g. *-že* vs. *-še*), avoidance of consonant clusters (e.g. *-əže* vs. *-že*), vowel harmony (e.g. *-že* vs. *-žö*) (Alhoniemi 1993: 72; Riese et al. 2019). Furthermore, word-final /e/, /o/, /ö/ generally do not bear stress in Mari apart from some lexemes, and these unstressed vowels are commonly reduced to [ə] in many dialects¹¹ (Alhoniemi 1993: 24), and in rapid speech, this reduced vowel can be omitted altogether. As a shorthand for all these variants, I will refer to the suffix as *žE*, written in small capital letters.

As concerns its distribution in discursive functions, the possessive suffix can be attached to almost any constituent – not only to nouns, or adjectives, but also to pronouns, quantifiers, infinite forms, finite verbs (Alhoniemi 1993: 76; Andreeva 2014; Tuzharov 1987: 67) and it can even be stacked onto another possessive suffix (see Section 8.2). In essence, the suffix in this function does not seem to follow any selectional requirements and simply attaches to the right of a single constituent or phrase. For example, observe how *žE* attaches to a PP *deč posna* in example (109) or to a whole converbial clause marked with the negative converb *-de* in example (110).¹²

(109) CSMM, VKontakte

[Araka deč posna]-že veselitl-aš ok lij mo?
 vodka without-PX3SG have.fun-INF NEG.3SG be.CNG Q

‘Can’t you have a good time without alcohol?’

¹¹In some examples taken from Serdobol'skaya and Toldova (2012), the vowel of the suffix is not the expected harmonic variant. As Natalia Serdobolskaya (p.c.) informs me, the vowels are hard to distinguish in the recordings, but nevertheless it seems that in some cases speakers use the elsewhere variant [e] even in cases of expected harmony: *köže* pro *köžö* in ex. 131.

¹²Thanks to Micke Mayer for drawing my attention to example (110).

(110) *CLM*, Fiction text

[*Təgaj pis'ma-lan vašestə-de*]-ž-at *niguze ok lij.*
such letter-DAT answer-CVB.NEG-PX3SG-ADD no.way NEG.3SG be.CNG
'One can't just not respond to a letter like that, either.'

In its discursive function, the 3SG suffix may also be stacked onto a possessive suffix in canonical possessive function (see Pleshak 2019; Georgieva 2022a; Tuzharov 1987: 67). According to A. Simonenko (2014), in the following example, the 1SG possessive suffix marks that the cow belongs to the speaker, and the 3SG suffix marks the cow's membership in the set of all cows belonging to the speaker. This example illustrates that "genuine" possessive suffixes like the 1SG suffix, and the 3SG suffix do not occupy the same structural position here and are indeed, different suffixes.

(111) (A. Simonenko 2014) (Context: I have four cows. I want to sell one of them:)

Uškal-em-že šiz-eš, što māj tudə-m užal-em.
cow-PX1SG-PX3SG feel-3SG that 1SG 3SG-ACC sell-1SG
'A cow of mine (among those) feels that I'm going to sell her.'

As regards the discursive functions of *žE*, different analyses have been proposed to capture its full usage range. The suffix has been described as marker of definiteness (e.g. Alhoniemi 1993: 75-76), identifiability (Nikolaeva 2003), specificity (Georgieva 2022a; A. Simonenko 2017), contrast (Riese et al. 2019: §9.2.5; Tuzharov 1987: 67-68; Nikolaeva 2003; A. Simonenko 2014), as marker establishing reference to previously mentioned individuals in the discourse (Riese et al. 2019: §9.2.5; Saarinen 2022: 445), or some other kind of emphasis (Riese et al. 2019: §9.2.5; Tuzharov 1987: 67). In the following section I will summarize and evaluate the previous proposals related to definiteness and specificity before discussing some corpus data that illustrate functions of *žE* in the wider discourse context which have not attracted much attention so far. I argue in Section 8.4 that *žE* can function as a topic marker for shift topics in addition to its contrastive function by showing that these two functions can be teased apart in the right context. The

analysis is followed by discussion and some problematic examples which suggest influence from Russian discourse particles on the usage of the suffix as well.

8.2 Previous approaches

In this section I will review some previous analyses of žE, starting with the analysis as marker of definiteness (e.g. Alhoniemi 1993: 75-76, see Nikolaeva 2003). A definiteness analysis is too restrictive for žE, however, and does not explain a number of the suffix’s characteristics, as Nikolaeva (2003) and A. Simonenko (2014) have shown. Some typical contexts for markers of definiteness are global or local uniqueness, usage in anaphoric contexts, or referring to an individual in a previously mentioned superset. The suffix can attach to specific indefinite pronouns like *ala-kö* ‘somebody, anybody’ which is already indicative of the fact that žE cannot plausibly mark definiteness:

(112) CLM, News text

žušö-šaməč kokla gəč ala-kö-žö jüləšö sorta-m
 drunk-PL among from INDEF-who-PX3SG burning candle-ACC
spirt-an stakan-əš joŋələš kamvozəkt-en [...]
 alcohol-ADJ glass-ILL by.mistake drop-PST2.3SG

‘Among the drunk, someone by mistake dropped the burning candle into a glass containing alcohol.’

As regards its distribution, the suffix appears on constituents other than nouns, which are not associated with definiteness as they lack referents which could be definite (see ex. 109–110 above). In the other typical contexts for definite expressions like the uniqueness of the referent in a “local” context (113) or uniqueness of the referent in a “global” context (114) usage of the suffix is not licit, either:

(113) (A. Simonenko 2014, in a room with one door)

*Petəre-za omsa-(*žə)-m!*
 close-IMP.2PL door-PX3SG-ACC

‘Close the door!’

(114) (A. Simonenko 2014)

*ʃara šinča dene kečə-š-(*še) onč-aš ok lij.*
bare eye with sun-ILL-PX3SG look-INF neg.3SG be.CNG

‘One shouldn’t look at the sun with unprotected eyes.’

An antecedent does not license the anaphoric use of *žE*, according to (A. Simonenko 2014) (ex. 115), but Georgieva (2022a) reports that one consultant accepted the anaphoric usage of *žE* in that example. Furthermore, the consultant accepted usage of the suffix in an elliptical structure (ex. 116). In the latter example, the groups of three and two pancakes each are known from the context, and the numerals refer to them.

(115) (A. Simonenko 2014)

*Vasja kniga-m nal-ən. Tače tudo (tide) kniga-(*žə)-m*
Vasja book-ACC buy-PST2.3SG today he that book-PX3SG-ACC
lud-eš.
read-3SG

‘Vasja bought a book. Today he is reading (that) book.’

(116) (Georgieva 2022a: 6) (Context: Mom baked pancakes. She put three (of them) into the fridge and two (of them) on the table.)

%*Maša kokət-šə-m kočk-ən, a Taňa kumət-šə-m.*
Maša two-PX3SG-ACC eat-PST2.3SG and Taňa three-PX3SG-ACC

‘Masha ate the two (from the table), and Tanya the three (from the fridge).’

According to A. Simonenko (2014), the only necessary and sufficient condition for licensing *žE* is membership of the marked expression’s referent in a previously introduced group, or a group that is an alternative set associated with contrastive focus. Examples for picking a referent from a previously introduced group are given in ex. (117–118). In the former example, omission of the possessive suffix would render the reference of *ik kniga* ‘one book’ ambiguous, not linked to the previously introduced books, according to A. Simonenko (2017). The sentence would result in a non-specific meaning ‘any book’ instead of the partitive reading

‘one of those books’. The same usage is licit in elliptic structures as well (example 119), where usage of the suffix is preferred, according to (Georgieva 2022a). The difference between example (119) and the previously shown example (116) is that in example (116), the groups of two and three pancakes each are already introduced in the context whereas in example (119), these groups are not known.

(117) (A. Simonenko 2017)

Maj kum kniga-m nal-an-am. Ik kniga#(-žə)-m Kost’a-lan
 1SG three book-ACC buy-PST2-1SG one book-PX3SG-ACC Kost’a-DAT
pölekl-em.
 give.present-1SG

‘I bought three books. I will give one of them to Kost’a.’

(118) (Riese et al. 2019: §9.2.5.)

Memnan indeš čave ulo. Iktə-žə ala-kušan
 1PL.GEN nine chicken EXIST one-PX3SG INDEF-where.LAT
jom-an.
 get.lost-PST2.3SG

‘We have nine chickens. One (of them) has gone off somewhere.’

(119) (Georgieva 2022a: 5) (Context: Mom baked pancakes.)

Tudo kumat-(šə)-m xolodilnik-aš šəndə-š, a kokət-(šə)-m
 3SG three-px3sg-acc fridge-ILL put-PST1.3SG and two-px3sg-acc
üstembake.
 onto.the.table

‘She put three (of them) into the fridge and two (of them) on the table.’

8.3 Contrastive topics

In some of the examples above, ŽE appears as a marker of contrastive topics. This function has been noted by Yurayong (2020: 115) and Georgieva (2022a) be-

fore. As an illustration, consider an example in which two individuals are introduced and then marked as contrastive topics by *žE*. In example (120), a man and a woman are introduced in the first sentence and then contrasted in the second one.

(120) (Alhoniemi 1993: 75)

Pört-aštö pörjeŋ den üdaramaš mutlan-en šinč-at. Teve
 house-INE man with woman talk-CVB sit-3PL behold
ÜDARAMAŠ-ƏŽE peš kutər-a, PÖRJEŊ-ŽE peš koləšt-eš.
 woman-PX3SG very speak-3SG, man-PX3SG very listen-3SG

‘In a house there sit a man and a woman, talking. Behold, the WOMAN speaks a lot, the MAN listens carefully.’

Examples of this kind of contrastive usage are frequently found across genres in Mari. Also note that in some cases similar meanings can be expressed by adding *žE* only to the second noun (see example 145), or omitting it altogether (see below).

A typical context eliciting contrastive topic – focus sequences is a question containing multiple *wh*-words (see Section 3.4). In Mari, there is no special morphosyntactic strategy to answer such a question. Consider example (121). This sentence was elicited with the help of a picture that showed a man holding a spoon and a woman holding a knife. The questions were variably formulated as all new question (122a), as multiple constituent question (122b), as coordinated question for agent focus (122c) and as coordinated question for patient focus (122d). Irrespective of the question prompt, the answers varied only little. The word order in the answer is the canonical SOV, with some optional elided constituents. My informant did not prefer to mark the contrastive topics with a possessive suffix.

(121) (elicited)

ÜDARAMAŠ küzə-m (kuč-a), (a) PÖRJEŊ sovla-m.
 woman knife-ACC hold-3SG but man spoon-ACC

‘The WOMAN(CT) (is holding) a knife, (and) the MAN(CT) a spoon.’

(122) (elicited)

- a. *Mo təšte kaj-a?*
what here go-3SG
‘What is happening here?’
- b. *Kö mo-m kuč-a?*
who what-ACC hold-3SG
‘Who is holding what?’
- c. *Kö sovla-m kuč-a, a kö küzə-m kuč-a?*
who spoon-ACC hold-3SG but who knife-ACC hold-3SG
‘Who is holding the spoon, and who is holding the knife?’
- d. *Mo-m üdəramaš kuč-a, mo-m pörjeŋ kuč-a?*
what-ACC woman hold-3SG what-ACC man hold-ACC
‘What is the woman holding (and) what is the man holding?’

Answers to questions containing a plural, which are a typical context for contrastive topics as well, elicit the same response:

(123) (elicited. Where do your siblings live?)

- IZA-M Moskva-šte, a ŠÜŽAR-EM Joškar-Ola-šte.*
older.brother-1SG Moscow-INE but little.sister-1SG Joškar-Ola-INE
‘MY OLDER BROTHER(CT) (lives) in Moscow, and MY LITTLE SISTER(CT) (lives) in Joškar-Ola.’

Contrastive topics *can* be marked with the marker ŽE, but marking is optional. In situations where the set of entities is explicitly introduced before, my consultant did not express a preference for using or omitting the suffix without further context:

(124) (elicited. What food do your pets like to eat?)

PIJ-EM(-ŽE) pəzə-m kočk-aš jörat-a, pəRəS-EM(-ŽE) kol-əm
 dog-PX1SG-PX3SG meat-ACC eat-INF love-3SG cat-PX1SG-PX3SG fish-ACC
kočk-aš jörat-a.
 eat-INF love-3SG

‘My DOG(CT) likes to eat meat, (and) MY CAT(CT) likes to eat fish.’

In a context where the set is not known and the suffix is used, it typically implies the presence of a contrasted individual, but the contrast does not have to be known to the speaker:

(125) (elicited)

Maša kočkəš-əm jamdəl-en mo? – Šür-žə-m tudo šolt-en.
 Maša food-ACC prepare-PST2.3SG Q soup-ACC 3SG cook-PST2.3SG
Ves kočkəš-əm tudo jamdəl-en aʎe uke – o-m
 other food-ACC 3SG prepare-PST2.3SG or NEG.EXIST NEG-1SG
pale.
 know.CNG

‘Did Maša prepare food? – The soup she cooked. Whether she cooked other food, I don’t know. (= But some other food was cooked as well.)’

Summarizing the discussion so far, we find that ŽE is not licensed by the typical contexts for definiteness – local or global uniqueness – and there is interspeaker variation in admissibility of the suffix in anaphoric usage.

Marking a contrast between two referents is an additional very frequent usage context of the suffix. A. Simonenko (2014, 2017) argues that the only possible usage context of ŽE in a non-possessive function are contexts with a superset antecedent. Similarly, Georgieva (2022a) (following A. Simonenko 2017) argues that -ŽE is licit in nominal elliptical structures if “the remnant and the antecedent are understood as belonging to a set” with a part-whole relation, where “the remnant is usually interpreted as specific in the sense of Enç (1991)”. A. Simonenko (2017) argues that, in contrast to “maximally quantifying determiners” known

from Indo-European languages like Romance, Germanic, or Albanian, Mari žE belongs to a class of “non-maximally quantifying determiners”. Such determiners do not presuppose that there is only one individual with the properties denoted by the noun, rather, they “relate the denotation of the noun to some contextually specified individual”. In the case of žE, this relation between the antecedent and the denotation of the head noun is possession or inclusion, but precisely not identity. Importantly, the suffix is for the most part not obligatory in the superset contexts which means that žE is not (yet) a grammaticalized marker of definiteness or specificity.¹³ In sum, the suffix žE most productively functions as marker of contrast to or selection from a superset.

8.4 žE as marker of topic shift

The analyses discussed above do not specifically deal with the usage of žE in the wider discourse context. I will argue that besides marking contrast, žE can be used as marker for shift topics. A similar analysis has also been drawn up for the Hill Mari 3SG possessive suffix in Khomchenkova (2022).

The first indicator of a topic marking function is the distribution of the suffix. In most cases in discursive usage, žE appears attached to a constituent in the left periphery or altogether first position of the clause, the position for other topical expressions (see Section 7.1), without categorical restrictions. This distribution suggests that the suffix preferably attaches to topical constituents (and contrastive topics). Consider example (126) where the žE-marked object in clause-initial position precedes an overt pronominal subject in an OSV order:

¹³However, see Georgieva (2022a: 6) for an example where the omission of the suffix on a numeral resulted in an infelicitous reading of an elliptical structure for one speaker, making this claim more difficult.

(126) (elicited)

A *Irina kok tarelka-m vele mužk-ən šuktə-š.* **Kumšo**

but Irina two plate-ACC only wash-CVB manage-PST1.3SG third

tarelka-žə-m *tudo mužk-ən əš* *šukto.*

plate-PX3SG-ACC 3SG wash-CVB NEG.PST.3SG manage.CNG

‘But Irina only managed to wash two plates. A third plate she did not manage to wash.’

In most cases in the dialogue material from Serdobol’skaya and Toldova (2012) which I examined in detail, the suffix marks a topic shift. In Section 3.2.2, I defined shift topics as topics that are newly introduced, newly changed, or newly returned topics, following the classifications by Frascarelli (2018) and Givón (1983). Below, I discuss some examples from the small corpus in Serdobol’skaya and Toldova (2012) and elicited examples that show *žE* as topic marker. The examples were chosen to highlight different properties of *žE*, often occurring in the same stretch of discourse.

First, consider example (127). The participants of the dialogue talk about A’s trip to the countryside. In their first turn, A introduces springs (*pamaš*) into the discourse. In B’s turn, the springs become the new topic and are marked with *žE*. In the next turn, however, A introduces a completely new topic, namely wild strawberries. To mark the sudden change of topic, the strawberries are marked with *žE*. We see that in this example *žE* marks two shift topics (*pamašəže*, *sněgəže*), and these differ only in terms of whether the referents of these expressions had been introduced into the discourse before marking them with *žE*; only *pamaš* was mentioned before. This shows that the topics marked with *žE* can either be completely new, or introduced as new sentence topics. In any case *žE* marks a topic shift.

(127) (Serdobol'skaya and Toldova 2012: 729)

- A. *Nəl keče mo lij il-əš-na; vot mo saj-že pamaš*
 four day what be live-PST1-1PL well what good-PX3SG spring
šuko ulo tušto, srazu pamaš gəč vüd-əm numal-ət.
 many EXIST there at.once spring from water-ACC carry-3PL

‘We were there for four days, or so; you know, what’s best, there’s lots of springs, people just go right away to get some water from the springs.’

- B. *Memnan pamaš-əže nim-at kod-ən og-əl*
 1PL.GEN spring-PX3SG nothing-ADD stay-CVB NEG.3SG-be.CNG
čəla ənde pətar-en-ət; keč ik pamaš-əm erəkt-en
 all now end-PST2-3PL maybe one spring-ACC clean.out-PST2.3SG
əl̩e.
 AUX

‘At our place, **springs**, there is nothing left, they have all run dry; maybe one should be cleaned out.’

- A. *Snegə-že vet čot šuko ulo, šoldəra*
 wild.strawberry-3SG see very much EXIST large
möre gaj.
 garden.strawberry like

‘(And) **wild strawberries**, you know, there’s lots of, like large strawberries from the garden.’

In the next example, (128), the speakers are talking about A’s trip to the Kužeņer region. A inquires whether B has been there, too, and introduces Kužeņer in their turn again. In A’s next turn, A explicitly shifts the topic to the *village* of the same name by marking it with ŽE. Then, both discuss a wedding, before in the last turn of this dialogue example, A returns to the Kužeņer topic again. In the last turn, A first contrasts nature (*pürtüsše*) and (living) conditions (*uslovijže*), marking them with ŽE, which is typical of this construction. But then, the topic is shifted back

to Kužer which the following sentence is about, and the shift is marked overtly with ŽE. Even though Kužer had been mentioned in the first two turns of this excerpt, and had been a topic of the discourse before this portion of dialogue, it is not marked with ŽE before it is made a topic later. Repetition of a topic is therefore not a sufficient condition for marking with ŽE. Also note that Kužer as continued topic is situated in the postverbal position in A's first turn and B's second turn.¹⁴

(128) (Serdobol'skaya and Toldova 2012: 728–729)

A. *Už-ən-at, mij-en-at Kužer-əš?*
see-PST2-2SG go-PST2-2SG Kužer-ILL

‘Did you see... Did you go to Kužer?’

B. *Kužer-əš təge-že uke. Tak pur-en lekt-ən-na. Možet*
Kužer-ILL so-PX3SG no well enter-CVB leave-PST2-1PL maybe

už-ən-am.

see-PST2-1SG

‘To Kužer, so, no. Well we passed through. Maybe I saw it.’

A. *Kužer-že ške-že poselok jöršan saj ogəl.*
Kužer-PX3SG SELF-PX3SG village absolutely good NEG

‘Kužer itself, the village is not nice at all.’

B. *Aj. . . mij-en-na ələ S.-at-ən kö-t-ən süan-əškə-že*
oh go-PST2-1PL AUX S.-ADD-GEN who-ADD-GEN wedding-ILL-PX3SG

Kužer-əš?

Kužer-ILL

‘Oh. . . didn't we go to S.'s, or to whose wedding, to Kužer?’

¹⁴Note that in this example, the order ADD-GEN in the forms *S.-at-ən*, *kö-t-ən*, and *El.-at-ən* is unexpected as *-at* is a clitic which attaches after case endings. I do not have an explanation for this order.

A. *Mogaj S. memnan gəč mo?*

what.kind.of S. 1PL.GEN from Q

‘What S.? From our area?’

B. *Aj, El.-at-ən aka-že L.-n süan-əškə-že. Počəš*

oh E-ADD-GEN older.sister-PX3SG L.-GEN wedding-ILL-PX3SG after

jü-aš mij-en-na ələ.

drink go-PST2-1PL AUX

‘Oh, to E.’s elder sister L.’s wedding. After that we went drinking.’

A. No *pürtüs-še peš saj, a uslovij-že saj*

well nature-PX3SG very good but circumstance-PX3SG good

ogəl il-aš. Tora Kužəjer-že, klub-əšt-at

NEG live-INF far Kužəjer-PX3SG club.house-PX3PL-ADD

ala-mogaj pu pört.

INDEF-what.kind.of wood house

‘Well **nature** (there) is really nice, but the **conditions** are not good, to live (there). The remote **Kužəjer**, even their club house is just some wooden shed.’

In some examples of this sort, the usage of ŽE as marker of shift topics seems to be tied to unexpectedness or adversative contexts (see also Khomchenkova 2022: 162 for Hill Mari). In the beginning of the narration which the following example is taken from, a shed (*saraj* ~ *levaš*) is introduced. The shed is mentioned a few times, but only in the last clause it becomes the sentence topic and is marked with ŽE:

- (129) (Serdobol'skaya and Toldova 2012: 724) (Once, there was a brick factory besides this spring. In addition to the brick factory, they built a shed near the spring.)

No ikana šorgəkt-en jogə-mo šošo vüd saraj-əm
 but once beat.down-CVB flow-PTCP.PASS spring water shed-ACC
mušk-ən naŋgaj-en. Tiddeč vara levaš-əm ves
 wash-CVB take.away-PST2.3G from.this then shed-ACC other
vere čoŋ-en-ət. Revolucij žap-əšte kerməč jamdələmaš
 to.place build-PST2-3PL revolution time-INE brick production
čarn-en-ət. Saraj-že šuko žap eše šinč-en.
 stop-PST2.3PL shed-PX3SG much time still sit-PST2.3SG

‘But one time the wildly flowing water in spring washed away the shed. After that they built the shed in a different place. After the Revolution they stopped making bricks. (But) **the shed** was still standing for a long time.’

In many cases, ŽE appears on left dislocated topics. Consider example (130). Before this dialogue part, the speakers had been talking about being recorded by Mari television while doing farm labour and watching the program afterwards. Speaker A describes the content of the program using a generic statement with a 3PL null subject (‘they showed’). Speaker B then reintroduces Mari television as topic in the left periphery, marking it with ŽE.

- (130) (Serdobol'skaya and Toldova 2012: 730–731)

A. *No no šudə-m kəškə-mə-m da mo ončəkt-en-ət,*
 and and hay-ACC throw-PTCP.PASS-ACC and what show-PST2-3PL
tunam professor-vlak-əm vašlij-ən-na.
 then professor-PL-ACC meet-PST2-1PL

‘And, and, they showed how (we) threw about some hay and what not, then we met the professors.’

B. *Marij televidenij-že tidə-m-ak vele ončəkt-at kuze*
 Mari television-PX3SG this-ACC-EMPH only show-3PL how
ubiratl-at.
 harvest-3PL

‘Mari television – they only show such stuff, how people harvest crops.’

A similar case of left dislocation can be observed in example (131). The speakers are talking about student life in a shared apartment and B’s roommates. speaker A in their first turn introduces as a person *L.*, asking a question about them. Again, the marked constituent is in the nominative case but the verb does not agree with it. Then, in B’s first turn, another person *A.* is marked with *žE* and contrasted to person B speaking (that is, it is not speaker B and person *L.* living together, but person *A.* and person *L.*). Note that person *A.* is not the sentence topic in B’s turn and although marked with *žE*, the constituent does not occupy the clause-initial position. I suggest that this is characteristic of contrastive usages of *žE* but not necessarily of topic marking *žE*. Instead, person *L.* is the topic, as also indicated by the choice of pronominal reference strategies; the person is referred to by the 3SG personal pronoun *tudo* in B’s first turn, and by a null pronoun in A’s next turn. Then again, in speaker A’s last turn, they shift the topic to person *A.* again. Person *A.* is in a left dislocated position, not syntactically integrated into the clause.

(131) (Serdobol’skaya and Toldova 2012: 732–733)

A. *A kö-že L.-še pərļa o-da ile mo?*
 but who-PX3SG L.-PX3SG together NEG-2PL live.CNG Q
 ‘But **who**... what about *L.*, don’t you live together?’

- B. *Uke tudo A.-že dene ala-kuš ves ver-eš*
 no 3SG A.-PX3SG with INDEF-where.LAT other place-LAT
kaj-en-at maneš.
 go-PST2-3PL apparently
 ‘No, s/he and A. went together to a new place, apparently.’
- A. *Mo kelš-en og-əl adak?*
 what like-CVB NEG.3SG-be.CNG then
 ‘What didn’t (s/he) like, then?’
- B. *Šukə-n ul-əna da.*
 many-INS be-1PL PTCL
 ‘There are (too) many of us (in the apartment).’
- A. *A.-že tide podružkə-žo mo?*
 A.-PX3SG that female.friend-PX3SG Q
 ‘(And what about) A. – (is) this his/her friend?’
- B. *Podružkə-žo no pərļa tunem-ən-at škol-əšto.*
 female.friend-PX3SG um together study-PST2-3SG school-INE
 ‘His/her friend, um, they went to school together.’

In the previous texts I have shown examples for different properties of ŽE – its contrastive usage, and as marker of shift topics which in turn can appear as typical sentence topics in the left periphery of the clause or as left dislocated topics. Importantly, there are also topics which ŽE *cannot* mark, namely familiar topics.

As discussed in Section 3.2.2, most scholars distinguish at least two types of topics – shift topics and familiar topics. We saw in Section 7.2 that the postverbal field is generally a position for backgrounded, familiar constituents, including familiar topics. The shift topic marking analysis of ŽE proposed in this thesis, together with the facts about the information structural interpretation of constituents in the postverbal position, predicts that constituents marked with ŽE cannot appear as topics in a postverbal position. This is because shift topic marking is pragmatically infelicitous for familiar topics.

Consider example (132) for which such a context was set up. In this example, the first speaker introduces a cat as topic. Reference to the cat is continued with the 3SG personal pronoun *tudo* in the next turns. Some turns later, the cat can appear postverbally in the clause as familiar topic that the speakers return to (recall that new aboutness topics were not licit in this position, see ex. 95). With only case and regular possessive marking on the constituent, the sentence is felicitous, but additional marking with *žE* was not accepted by my consultant.

(132) (elicited)

- A. *Pərəs-et dene mo lij-ən?*
 cat-PX2SG with what be-PST2.3SG
 ‘What happened to your cat?’
- B. *Tudo ala-kušan jom-ən.*
 3SG INDEF-where.LAT get.lost-PST2.3SG
 ‘It has gone off somewhere.’
- A. *No, tudə-m kəčal-ən-at mo?*
 well 3SG-ACC search-PST2-2SG Q
 ‘Well, have you looked for it?’
- B. *Da, i mu-ən lukt-ən-am, pərəs-em-əm /*
 yes and find-CVB lead.out-PST2-1SG cat-PX1SG-ACC
**pərəs-em-žə-m.*
 cat-PX1SG-PX3SG-ACC
 ‘Yes, and I did find (him), my cat.’

The same can be seen in another example with a similar context (ex. 133). Here, a little sister is introduced as topic in the first turn. The topic is continued with null subjects and can appear postverbally as familiar topic. Again, marking with *žE* was not accepted by my consultant here. Note as well that an overt resumptive element is not necessary for the dislocated constituent in either of the examples, be it a dislocated object (ex. 132) or a subject (133).

(133) (elicited)

A. *Šūžar-et kodšo olimpiadā-šte učastvovatl-en, čən?*
little.sister-PX2SG previous olympics-INE take.part-PST2.3SG true

‘Your little sister took part in the last olympics, didn’t she?’

B. *Tuge.*
yes

‘That’s right.’

A. *A mogaj konkurs-əšto učastvovatl-en?*
and what.kind.of competition-INE take.part-PST2.3SG

‘And what competition did she partake in?’

B. *Ijmašte. Saj ver-əm nale, šūžar-em /*
swimming.INE good place-ACC get.PST1.3SG little.sister-PX1SG

**šūžar-em-že.*

little.sister-PX1SG-PX3SG

‘Swimming. She got a good place, my little sister.’

Note, however, that ŽE-marked constituents are not per se excluded from positions outside of the left periphery. In the left periphery, ŽE marks a topic shift, but other instances of ŽE that select an individual from a superset do not necessarily target topics. Contrastive ŽE appears to have some more positional freedom – it is not restricted to the left periphery as topical position. There, it can for example follow pronominal subjects (see ex. 119 and 131) or as a marker of contrasted objects simply appear in the preverbal position, the unmarked position for objects:

(134) (A. Simonenko 2014)

Memnan škol-na u, a tengeče ala-kö tide
our school-PX1PL new but yesterday INDEF-who that
okna-žə-m šalalt-en.
window-PX3SG-ACC break-PST2.3SG

‘Our school is new, but yesterday someone broke that [pointing to one window] window.’

(135) (Georgieva 2022a: 10)

A tāj nine kum počelamut kokla gəč kokət-šə-m vele
and 2SG those three poem among from two-PX3SG-ACC only
lud-ən-at.
read-PST2-2SG

‘And you have read only two from those three poems.’

Such examples suggest that, since žE-marking in this case is not immediately related to topic marking, the suffix should be able to appear in the postverbal position which is usual for non-topics or familiar topics. Indeed, žE-marking on a postverbal constituent can be rescued by introducing a salient alternative for the referent of the marked constituent in the right context. Consider examples (136) and (137). These are small contextual variations on the examples above. In example (136) not only a cat is introduced as topic, but a cat, and a dog as an alternative. They are both continued as topics. In the last turn, the dialogue participant returns to the cat, again in a postverbal position. Importantly, the dialogue is continued with the dog as contrasted alternative after the target sentence. In this context, marking the constituent with žE was judged acceptable by my consultant. The same can be observed in example (137): Not only the little sister is introduced as discourse topic, but a big sister as well. Again, in the last turn, the little sister may appear in the postverbal position, but žE-marking is possible in this context.

(136) (elicited)

- A. *Pəras-et-at, pij-et-at dene mo lij-ən?*
cat-PX2SG-ADD dog-2SG-ADD with what be-PST2.3SG

‘What happened to your cat and dog?’

- B. *Nuno ala-kušan jom-ən-ət.*
3PL INDEF-where.LAT get.lost-PST2-3PL

‘They have gone off somewhere.’

- A. *No, nunə-m kəčal-ən-at mo?*
well 3PL-ACC search-PST2-2SG Q

‘Well, have you looked for them?’

- B. *Da, i mu-ən lukt-ən-am, pəras-em-əm /*
yes and find-CVB lead.out-PST2-1SG cat-PX1SG-ACC
pəras-em-žə-m. A pij-em-əm eše mu-ən
cat-PX1SG-PX3SG-ACC but dog-PX1SG-ACC yet find-CVB
om-əl.
NEG.1SG-BE

‘Yes, and I did find (him), my cat. But I have not found my dog yet.’

(137) (elicited)

- A. *Šūžar-et, aka-t kodšo olimpiadə-šte*
little.sister-PX2SG big.sister-PX2SG previous olympiad-INE

učastvovatl-en-ət, čən?
take.part-PST2-3PL true

‘Your little sister and your big sister took part in the last olympics, didn’t they?’

- B. *Tuge.*
yes

‘That’s right.’

- A. *A mogaj konkurs-ašto učastvovatl-en-at?*
 but what.kind.of competition-INE take.part-PST2-3PL
 ‘And what competition did they partake in?’
- B. *Ijmašte. Saj ver-əm naľe, šüžar-em /*
 swimming.INE good place-ACC get.PST1.3SG little.sister-PX1SG
šüžar-em-že. No aka-m pətartəš ver-əm
 little.sister-PX1SG-PX3SG but big.sister-PX1SG last place-ACC
veľe naľe.
 only get.PST1.3SG
 ‘Swimming. She got a good place, my little sister. But my big sister only got the last place.’

These results indicate that ŽE as topic marker is incompatible with familiar topics in a postverbal position (ex. 133–132), but can in this position be rescued with a contrastive interpretation if no shift interpretation is available.

8.4.1 Interim summary – properties of ŽE

In this section I examined the discursive properties of the Meadow Mari 3SG possessive suffix ŽE. I showed that besides its well-known contrastive function and marker of identifiability, ŽE can mark shift topics by attaching to them in the left periphery. The suffix can mark both those topics that have been previously introduced, and those that have not. The only condition for felicitous topic marking with ŽE seems to be that it must mark the shift to a new topic, which the sentence is “about”. Left dislocated topics (ex. 130 and 131) are a clear indicator of the topic shift function since they serve no other purpose than explicitly introducing a new topic. Furthermore, the shift requirement is especially apparent in examples like ex. (128) where a constituent appears in some turns before it becomes a sentence topic and is marked with ŽE for the first time. Familiar topics are per se not compatible with ŽE. However the contrastive function can “rescue” its usage with non-shift topics, where for example, a familiar topic is acceptable when contrasted to an alternative from a previously introduced set (see

ex. 132–137). This suggests that in the contrastive function, žE is less positionally restricted. The obvious question which arises here is whether these functions can receive a unified analysis as contrastive topics, and in many cases, žE *does* mark contrastive topics which indicate the presence of pertinent alternative questions. In other cases, however, the suffix simply seems to mark the introduction of a new topic without a pertinent alternative question in the discourse, which is why the analysis as contrastive topic does not explain all contexts where žE appears.

In sum, I argued in the previous section that besides its known contrastive function, žE has a topic marking function as well. With the checklist of properties of Topic markers by Büring (2016) introduced in see Section 3.2.3, we can summarize the properties of the Meadow Mari 3SG possessive suffix žE in its topic marking function:¹⁵

- Distribution: All constituents (DP, VP, PP, AP), including indefinite pronouns and adverbs (e.g. *ala-kö* ‘someone’) can be marked.
- Marking with žE can:
 - establish a new discourse referent as the aboutee for the following (e.g. *snegaze* in ex. 127), or
 - establish an existing discourse referent as the new aboutee (e.g. *Ku-žerže* in ex. 128), but
 - not refer to an established aboutee throughout its tenure as ‘what the passage is about’, i.e. familiar topics (ex. 132, 133).
- Elements that meet that description do not *have* to be so marked, but merely *can* be.
- The same marking can be used for contrastive purposes.
- There are no other tests for the topic shift function as the usage of žE as topic marker is optional.

¹⁵Regarding the last point note that the topic marking function and contrast marking function can indeed be differentiated in the right context, see ex. (132–137).

So far, I have mainly investigated the necessary conditions for when a phrase *can* be marked with žE, in both topic marking and contrastive function, and it is unclear whether in these contexts the suffix ever obligatorily *must* be used, as in many cases its usage seems to be optional in the sense that the semantic well-formedness of a sentence does not depend on the usage of žE (cf. Nikolaeva 2003).¹⁶ Rather, the choice to mark a new topic with žE seems to be tied to the pragmatic importance of marking a topic, with factors influencing the choice including topic shift, contrast, or unexpectedness. The fact that žE frequently appears on left dislocated topics as topic marker is indicative of this as well.

In a comparative Mari view, my findings are very similar to those by Khomchenkova (2022) for the Hill Mari 3SG possessive suffix. Khomchenkova (2022) argues that the suffix is a marker for both shift topics and contrast. The function as shift topic marker can be observed in example (138) where it marks the new sentence topic (which was introduced before). Similarly to Meadow Mari, it can also mark the selection from a previously introduced superset, and in this contrastive function, the suffix has greater positional freedom. In example (139), ‘questions’ are introduced in the first clause, and one of the questions, marked with the 3SG possessive suffix is placed in a right peripheral position. Even indefinite NPs can be marked with žE, but only if contrasted. Whether there are any significant differences between Meadow Mari and Hill Mari žE is a matter for further research.

(138) Hill Mari (Khomchenkova 2022: 161)

Ik gänä lapka-š ke-mäk-em, kalpak-əm už-ən-am.
 one time store-ILL go-CVB.ANT-1SG hat-ACC see-PST2-1SG

Kalpak-šə ludə cäre-än ələ.
 hat-PX3SG grey color-ADJ be.PST1.3SG

‘One time when I went into the store, I saw a hat. **The hat** was grey.’

¹⁶Note however, that for one speaker in Georgieva (2022a: 6), usage of žE was indeed obligatory in anaphoric contexts and in elliptic contexts with an antecedent.

(139) Hill Mari (Khomchenkova 2022: 159)

ʃadmaš-vlä int'er'esnəj-vlä əl-ən-ət, tidä mämnan ros'ijä-n
question-PL interesting-PL be-PST2-3PL this 2PL.GEN Russia-GEN
konst'itucij-əm pälä-mä dono köldält-eš ik
constitution-ACC know-PTCP.PASS with be.tied-3SG one
ʃadmaš-əžə.
question-PX3SG

‘The questions were interesting, it was related to knowledge of our Russian constitution, **one question**.’

8.5 Other functions of žE and possible contact influence

To complete this section about the functions of žE, I will in the last part shortly discuss some other cases of the discursive usage of žE that cannot straightforwardly be explained by the topic or contrast analyses and probably require a different explanation altogether.

One factor complicating the delineation of the functions of žE is mentioned by Saarinen (2022: 466, Fn. 7) and Hammer and Bradley (2022: 62, Fn. 3): the most frequent, unrounded front vocalic allomorph of the possessive suffix, *že*, is homophonous with the Russian discourse particle *že* (же). Hammer and Bradley (2022) and Saarinen (2022) also claim that the Mari possessive suffix mirrors some usage contexts of the Russian discourse particle. In this section I will discuss a few cases where Russian influence might explain some of the properties of Mari žE found in contemporary data.

The Russian particle *že* is a multifunctional particle with a large usage range. Previous analyses have identified the following functions of *že* (see McCoy 2003 and references therein): thematic *že* for discourse organization; *že* as affective or modal clitic; as marker of contrast or contrastive focus; as marker of emphasis; as marker of (re-)activated information; as marker of a reference point in the activated “domain of reference”. The functions also go together with different syntactic properties. In her survey of Russian *že*, Padučeva (1988) distinguishes “initial *že*” from “neutral *že*”. The former is restricted to the position after the first

prosodic word or after the first syntactic phrase (cf. Valova and Slyusar’ 2017) and the latter takes any phonological word in the clause as host. Within the former group, Padučeva (1988) distinguishes between adversative meanings, argumentative meanings, and consequential meanings; in the latter between identity and exact specification. Two sentences illustrating the difference between argumentative and adversative rus. *že* are shown in example (140). In ex. (140a), *že* marks that speaker and hearer are both aware of a certain kind of reasoning in an “argument” and that this reasoning is uncontroversial (cf. Panov 2020) – being a combine harvester driver makes one like grain. In ex. (140b), however, *že* marks a contrast.

(140) Russian (Padučeva 1988: 27)

- a. *U menja ruk-i po hleb-u soskuči-l-i-s’.* **Ja**
 at 1SG.GEN hand-NOM.PL for grain-DAT long.for-PST-PL-REFL 1SG
že kombajner!
žE combine.harvester.driver
 ‘My hands were longing for grain. I drive a combine harvester, you know!’
- b. *On traktorist. Ja že – kombajner.*
 3SG.MSC tractor.driver 1SG *žE* combine.harvester.driver
 ‘He’s a tractor driver. I, on the other hand, drive a combine harvester.’

Examples for similar “argumentative” usages of Mari *žE* like in the Russian example (140a) can be found as well. Example (141) is set in the context that the narrator of the story got a woman pregnant. The woman wanted an abortion which the narrator did not initially want to agree on. Then, the narrator agrees to perform the abortion as he had a background in medicine. The suffix is attached to the personal pronoun in second position and marks that the reasoning is uncontroversial – due to the narrator’s medical education, he can perform an abortion. Neither is there an independent pertinent question such as ‘What did person *x* do?’ which would make *majže* interpretable as contrastive topic; nor is there an apparent topic shift. Identifiability marking cannot be a factor either,

since the referents of personal pronouns are always identifiable. The sentence is not primarily *about* properties of the narrator but rather the justification of his actions. The same can be observed in example (142); the speaker is explaining the reason for why there are no eggs, expressing this reasoning with *žE*. Importantly, the hearer is not required to share the knowledge with the speaker (ex. 143); the speaker only marks the reasoning as uncontroversial. Given the similarities to the usage of Russian *že* above, I suggest that usage of Mari *žE* in such a context is due to influence from Russian.

(141) CCWLM, Newspaper *Marij El* (31.03.2007)

Ik keč-ən tunem-me ver-ašk-em mij-en lekt-ən
 one day-INS study-PTCP.PASS place-ILL-PX1SG come-CVB leave-PST2.3SG
da, lüdəktəl-ən-ak, kutər-en savər-en. Məj-žE
 and frighten-PST2.3SG-EMPH speak-CVB turn-PST2.3SG 1SG-PX3SG
vrač-lan tunem-ən-am, avtomašina-m tarl-en, möngö
 doctor-DAT study-PST2-1SG car-ACC rent-CVB home
veləš kondə-š-əm da abort-šə-m-at ške-ak
 direction.ILL bring-PST1-1SG and abortion-PX3SG-ACC-ADD self-EMPH
aštə-š-əm.
 do-PST1-1SG

‘One day, she came to the place where I studied, threatening me, and made me change my mind. I was studying to become a doctor, **you see**, so I hired a car, brought her home and also performed the abortion on her myself.’

(142) (elicited)

Tače təlat nīmogaj munə-m kond-en o-m kert.
 today 2SG.DAT no.kind.of egg-ACC bring-CVB NEG-1SG can.CNG
Rəvəž-šE čəvə-na-m kočk-ən.
 fox-PX3SG chicken-PX1PL-ACC eat-PST2.3SG

‘I cannot bring you any eggs today. The fox has eaten our chicken, **you see**.’

(143) (elicited)

Tače kastene ava-t joča-vlak-na-m onč-en
 today in.the.evening mother-PX2SG child-PL-PX1PL-ACC watch-CVB
kert-eš mo? – Uke, ava-m-že otpusk-aš
 can-3SG Q no mother-PX1SG-PX3SG vacation-ILL
lekt-an.
 leave-PST2.3SG

‘Can your mother watch our children tonight? – No, my mother went on vacation, **you see**.’

In the “adversative” function where “initial *že*” is used to contrast two phrases, it appears only on the second contrasted phrase in Russian. In example (144) two adjectives are contrasted, and rus. *že* appears only on the second constituent, not on both like usually in Mari. My consultant also spontaneously produced sentences where *že* appears only on the second constituent (ex. 145 and 146).¹⁷ In this case, too, I suspect that this distribution is due to influence from Russian.

(144) (Valova and Slyusar’ 2017: 41)

Derevjann-yj dom-ik sgore-l, kamenn-ye že
 wooden-MSC.NOM house-DIM burn.down-PST.MSC stone-MSC.NOM.PL *že*
dom-a prostož-at ešče mnogo let.
 house-PL STAND.STILL-3PL still many year.GEN.PL

‘The wooden house has burnt down, but the stone houses will still stand for many years.’

(145) (elicited)

Tide Ivuk-an kində-že, tidə-že Anuš-an šokolad-še.
 this Ivuk-GEN bread-PX3SG this-PX3SG Anuš-GEN chocolate-PX3SG

‘This is Ivuk’s bread, this is Anuš’s chocolate.’

¹⁷The other instances of *že* in the first clause in example (145) are clearly employed for possessor marking.

(146) (elicited)

Kum butalka Jəvan-ən kid-əštə-že, a iktə-že Andrej-ən.
three bottle Ivan-GEN hand-INE-PX3SG but one-PX3SG Andrej-GEN
'Three bottles are in Ivan's hands, and one in Andrej's.'

In yet other contexts Meadow Mari žE seems to mirror the usage of another Russian discursive clitic, *-to*. This postposed clitic "can be regarded as a discourse marker that contributes to information structure through topicalisation and delimitation of a set of alternatives (i.e. focus and contrast), or expresses the interlocutors' state of knowledge such as familiarity or surprise" (see Yurayong 2020: 107 and references therein). For example, the particle *-to* can track sentence level topics. In example (147), speaker A introduces some new referents to the discourse (jacket, pants, cap), which are no discourse topics themselves. Speaker B then selects the pants as topic from these previously mentioned referents, making them a new topic (Grenoble 1998: 200). In this function, rus. *-to* resembles the shift topic function of mar. žE.

(147) Spoken Russian (Grenoble 1998: 200)

A. *na okno ravešivaj-et kurtk-u, štany, i šapk-u.*
on window.ACC hang-3SG jacket-ACC pants.PL and hat-ACC

Eto tolko priezžaj-et –
this.MSC only arrive-3SG

'He hangs his jacket, pants and hat on the window. He just arrives –'

B. *Štany-to otkuda?*
pants.PL-TO where.from

'Where do the pants come from?'

A. *A štany on vidimo sverhu nadevaj-et na velosiped-e*
 and pants.PL 3SG apparently on.top put.on-3SG on bike-PREP
ezdiť.
 ride.INF

‘Apparently he puts the pants on over his clothes when he rides his bike.’

Another context where Mari žE seems to mirror the usage of Russian *-to* are interrogative clauses where the suffix appears at the end of the clause and expresses surprise. The following examples show the same sentence in Mari and in Russian, where Mari žE corresponds to Russian *-to*.

(148) (Andreeva 2014: 10)

Kuze tægaj süret-əm süretl-en-at-še?
 how such picture-ACC paint-PST2-2SG-PX3SG

‘How did you paint such a picture?’

(149) Russian (Andreeva 2014: 10)

Kak tak-uju kartin-u narisova-l-to?
 how such-FEM.ACC picture-FEM.ACC paint-PST.MSC-TO

‘How did you paint such a picture?’

In sum, there are some contexts which lend themselves to a contact explanation. What exact role Russian could have played in influencing the functions of Mari žE is beyond the scope of this thesis as the parallels between Russian and Mari are part of a much larger Eurasian areal phenomenon: Possessive suffixes or postposed determiners can mark definiteness, topicality, or discourse prominence in a large area in Eurasia, spanning from the eastern Finnic languages like Veps up until the Turkic languages of the Volga area (Tatar, Bashkir, Chuvash), including Russian and the Uralic languages inbetween (Leinonen 2002). Additive clitics in a position after the first constituent are common markers of contrastive and shift topics in the world’s languages, too (Forker 2016).

Parallels between the particular Russian discourse particles and various clitics in other Finno-Ugric languages have been noted as well. The following is only a short list of what can be found in the literature. Russian *že* has widely been borrowed by Uralic languages, including all eastern Saami languages (see Juutinen 2019: 85–86): Kildin Saami as enclitic topic marker *=že ~ =še* ‘but; you know; you see; also’ (Kert 1971: 226; Rießler 2007: 238), Skolt Saami, Inari Saami (Valtonen, Fofonoff, and Juutinen 2018), Ter Saami (*že*) (Halász 1883: 40) and the Sea Saami variety of Northern Saami; Soikkola Ingrian (*že*) (Markus 2022: 85–86), Votic (Maytinskaya 1982: 139), Erzya Mordvin (*žo*) (Janda, Laakso, and Metslang 2022: 900, ex. 15), Moksha Mordvin (Blokland 2012: 5), Udmurt, and Komi (Zubova et al. 2022: 125) and Khanty (*še/se/že*) (Maytinskaya 1982: 139). Furthermore, the particle is found in other contact languages of Russian such as Russian Romani and eastern European Yiddish varieties (Panov 2020: 33–35). The number of languages that have borrowed Russian *že* illustrates that the particle is easily transferred in contact with Russian. As regards the borrowing not of the particle itself, but of the usage pattern, Khomchenkova (2022: 155) suggests that some usage contexts of *žE* in Hill Mari can be traced back to copying a pattern of Russian *-to*. Similar proposals exist for the pattern of the Komi possessive suffixes (Yurayong 2020: 110) or suffixes in eastern Finnic languages (Yurayong 2020).

In sum, it is very likely that the Russian discourse particle *že* has influenced the usage of the Mari suffix *žE* due to the similar phonological shape, and partly, its syntactic distribution because the suffix is very easily borrowed. However, while the functions of Russian *že* and Mari *žE* overlap to some extent, they are not identical. To express identity or exact specification, as Russian *že* can, for example, a clitic *-ak* is usually employed in Mari, as in ex. (101) (see Zakirova 2019 for functions of this suffix). Furthermore, Mari *žE* appears on both elements in contrastive usage, and seems to be more freely used than Russian *-to* or *že* for topic marking purposes. A definitive answer to the amount and the nature of contact influence on the functions of *žE* has to take into account the multitude of functions of Russian *že* and the larger areal patterns of discursive functions of second-position discourse clitics in general.

8.6 Summary

This section discussed the usage of the Mari 3sg possessive suffix in discursive usage. Previous research has shown that beside its usage in possessive contexts, the suffix has a number of different functions in the discourse: the most well-known is its contrastive function where the suffix selects one individual from a viable contrastive set. In this section I argued on the basis of data from corpora that the suffix can also mark shift topics. That is, if a speaker switches to a new sentence topic, the new topic can be marked with *žE*. Marking a topic shift with *žE* seems to be tied to the pragmatic importance of marking the new topic, with factors influencing the choice including a topic shift, contrast, or unexpectedness. The contexts for the contrast marking function and topic marking function often intersect but can be kept apart in the right context. It also appears that contrastive *žE* has more positional freedom than topic marking *žE*. Finally, the suffix has some additional discursive functions such as the enimitive that can likely be explained by contact with Russian.

9 Focus and Topic in Meadow Mari – Summary

This thesis is a first description of types of focus, topics, and contrastive topics, and their morphosyntactic coding in contemporary Meadow Mari. This section shortly summarizes the most important findings from the thesis.

Based on newly elicited data and corpus data I show in section 6 that prosodic in situ marking is an important strategy of constituent information focus marking in Mari, besides a preverbal focus strategy which is less frequently used. Constituent focus and correction focus work identically in this respect. An exception to the preverbal or in-situ placement of focus is presentational focus which can occur postverbally. Verbs can be focussed themselves by placing them in a left-peripheral position. Identificational focus can be marked by prosody alone, or by a more specific pseudocleft. In general, prosodic focus marking is underspecified for the type of focus. As regards the typology of focus marking, the existence of an in situ focus marking strategy disqualifies Mari as strict position language with a designated focus position, following Buring (2016). The preverbal position

has some relevance for the placement of focal constituents, however, as focal constituents *can* be placed there as well.

Topics are located in the left periphery of the clause in Mari. Non-subjects as topics can appear in a left-peripheral position if necessary for the topic progression, and a constituent in a clause is more likely to be chosen as antecedent for a null subject if in the leftmost position as topic. Outside of the main clause, Mari employs left dislocated topics for topic shift. These can be marked by an overt topic marking phrase *manmašte* as well, or with the 3SG possessive suffix *žE* as topic marker. The postverbal field, in turn, is a position in which constituents end up for various reasons: afterthoughts, completions of the preceding clause, and familiar topics can appear there as backgrounded, given elements; constituents appear postverbally as an epiphenomenon of verb fronting for focus purposes or as prosodically heavy constituents. As regards the typology of topic marking, Mari behaves like a typical rigidly head-final language.

Finally, I discuss the discursive functions of the Mari 3SG possessive suffix *žE*. As previous research has shown, one important function is the marking of contrast by selecting an individual from a previously introduced set. In this thesis I show that in addition to this function, the suffix is employed as marker of topic shift as well. Marking a topic shift overtly seems to be tied to the pragmatic importance of marking the shift. Furthermore, some functions of *žE* can likely be explained as result structural influence from the homophonous Russian discourse particle *že*.

References

- Aissen, Judith (1992). “Topic and Focus in Mayan”. In: *Language* 68.1, 43–80.
- (2015). “Documenting topic and focus”. Presentation on the 4 International Conference on Language Documentation and Conservation, University of Hawai’i. URL: <http://scholarspace.manoa.hawaii.edu/bitstream/handle/10125/25401/25401.pdf?sequence=2>.
- Alhoniemi, Alho (1988). “Das Tscheremissische”. In: *The Uralic Languages. Description, History and Foreign Influences*. Ed. by Denis Sinor. Leiden et al.: E.J. Brill, 84–95.
- (1993). *Grammatik des Tscheremissischen (Mari)*. Hamburg: Buske.
- (2010). *Marin kieliooppi*. Apuneuvoja suomalais-ugrilaisten kielten opintoja varten. Hilfsmittel für das Studium der finnisch-ugrischen Sprachen X. Helsinki: Suomalais-Ugrilainen Seura.
- Andreeva (2014) = Андреева, Л.А. (2014). “Функции посессивных суффиксов при глаголе в марийском языке”. In: *Вестник угроведения* 19.4. [Functions of possessive suffixes on the verb in Mari], 8–15.
- Anduganov (2001) = Андуганов, Юрий (2001). “Типологические изменения в синтаксе марийского языка”. In: *Congressus Nonus Internationalis Fenno-Ugristarum 7.–13.8.2000 Tartu. Pars IV. Dissertationes sectionum: Linguistica I*. Ed. by Tõnu Seilenthal. [Typological changes in the syntax of Mari]. Tartu, 72–77.
- Arkhangelskiy, Timofey (2019). “Corpora of social media in minority Uralic languages”. In: *Proceedings of the fifth Workshop on Computational Linguistics for Uralic Languages, Tartu, Estonia, January 7 - January 8, 2019*, 125–140.
- (2023a). *Corpus of Contemporary Written Literary Meadow Mari*. URL: http://meadow-mari.web-corpora.net/index_en.html (visited on 08/31/2023).
- (2023b). *Corpus of Social Media in Meadow Mari*. URL: http://meadow-mari.web-corpora.net/index_en.html (visited on 08/31/2023).
- Asztalos, Erika (2018). “Szórendi típusváltás az udmurt nyelvben”. [Word order type change in Udmurt.] PhD dissertation. Eötvös Loránd University.
- (2020). “Focus in Udmurt: Positions, Contrastivity, and Exhaustivity”. In: *Finnougric Languages and Linguistics* 9.1–2, 14–57.

- Asztalos, Erika (2021). "From head-final towards head-initial grammar. Generational and areal differences concerning word order usage and judgement among Udmurt speakers". In: *Language Contact in the Territory of the Former Soviet Union*. Ed. by Diana Forker and Lenore A. Grenoble. Amsterdam: John Benjamins, 143–182.
- Asztalos, Erika, Katalin Gugán, and Nikolett Mus (2017). "Uráli VX szórend: nyenyec, hanti és udmurt mondat szerkezeti változatok". In: *Nyelvelmélet és diakrónia 3*. Ed. by Katalin É. Kiss, Attila Hegedűs, and Lilla Pintér. [VX order in Uralic: sentence structures in Nenets, Khanty and Udmurt.] Budapest and Piliscsaba, 30–62.
- Baitchura, Uzbek (1988). *Instrumental-Phonetic Studies in Čeremis*. Vol. V–VI, 1987–88. Annali del Dipartimento di Studi dell'Europa Orientale. Sezione Linguistico-Filologica. Napoli.
- Bartholomä, Ruth and Monika Schötschel (2012). "Language Policy and Linguistic Reality of the Mari in the Republics of Mari El, Tatarstan, and Bashkortostan (Russian Federation)". In: *Finnisch-Ugrische Mitteilungen* 36, 1–39.
- Beke, Ödön (1938). *Tscheremissische Märchen, Sagen und Erzählungen*. Gesammelt und herausgegeben von Ödön Beke. Vol. 76. SUST. Helsinki: Suomalais-Ugrilainen Seura.
- Bereczki, Gábor (1984). "Die Beziehungen zwischen den finnougri-schen und türkischen Sprachen im Wolga–Kama–Gebiet". In: *Nyelvtudományi Közlemények* 86, 307–314.
- Blokland, Rogier (2012). "Borrowability of Pronouns: Evidence from Uralic". In: *Finnisch-Ugrische Mitteilungen* 35, 1–34.
- Bradley, Jeremy and Johannes Hirvonen (2022). "Null subjects in Mari". In: *Null Subjects in Slavic and Finno-Ugric. Licensing, Structure and Typology*. Ed. by Gréte Dalmi, Egor Tsedryk, and Piotr Ceglowski. Vol. 145. Studies in Generative Grammar. Boston and Berlin: de Gruyter, 281–306.
- Bradley, Jeremy, Alexandra Kellner, and Niko Partanen (2018). "Variation in word order in permic and mari varieties: a corpus-based investigation". In: *Языковые контакты народов Поволжья и Урала. XI Международный симпозиум. Чебоксары, 21–24 мая 2018 г. Сборник статей*, 238–244.

- Bradley, Jeremy and Jorma Luutonen (2023). “Mari”. In: *The Uralic Languages. Second Edition*. Ed. by Daniel Abondolo and Riitta-Liisa Valijärvi. New York: Routledge, 527–575.
- Bruening, Benjamin (2012). *The Scope Fieldwork Project*. URL: <https://udel.edu/~bruening/scopeproject/scopeproject.html> (visited on 09/26/2023).
- Büring, Daniel (2003). “On D-Trees, Beans, and B-Accents”. In: *Linguistics & Philosophy* 26.5, 511–545.
- (2009). “Towards a typology of focus realization”. In: *Information Structure*. Ed. by Malte Zimmermann and Caroline Féry. Oxford: Oxford University Press, 177–205.
- (2016). “(Contrastive) Topic”. In: *Oxford Handbook of Information Structure*. Ed. by Caroline Féry and Shin Ishihara. Oxford: OUP, 64–85.
- Burukina, Irina (2022). “On dative subjects and agreement with infinitives licensed by an external P head”. In: *NELS 52: Proceedings of the Fifty-Second Annual Meeting of the North East Linguistic Society*. Ed. by Pratley Breanna. Amherst, MA: GLSA, 105–118.
- (2023a). “External merge in spec,CP: Complementizers projecting an argument”. In: *Syntax* 26, 85–105.
- (2023b). “On the syntax of postpositional phrases in Mari: Choosing between two structures”. lingbuzz/007081.
- Chafe, Wallace L. (1976). “Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View”. In: *Subject and Topic*. Ed. by Charles N. Li. New York: Academic Press, 25–55.
- Chkhaidze (1941) = Чхаидзе, М. П. (1941). *Синтаксис лугово-восточного мари́йского языка*. [Syntax of Meadow Mari]. Москва: Государственное учебно-педагогическое издательство наркомпроса РСФСР.
- Czypionka, Anna (2007). “Word Order and Focus Position in the World’s Languages”. In: *Linguistische Berichte* 212, 439–454.
- Dalmi, Gréte, Egor Tsedryk, and Piotr Ceglowski, eds. (2022). *Null Subjects in Slavic and Finno-Ugric. Licensing, Structure and Typology*. Vol. 145. Studies in Generative Grammar. Boston and Berlin: de Gruyter.
- Den Dikken, Marcel (2018). “Case, P and Number in Mari possessive noun phrases”. Manuscript.

- De Smit, Merlijn and Gwen Eva Janda (2023). “Definiteness in Uralic”. Second Edition. In: *The Uralic Languages*. Ed. by Daniel Abondolo and Riitta-Liisa Valijärvi. New York: Routledge, 979–1006.
- Dezső, László (1978). “Towards a typology of theme and rheme: SOV languages”. In: *Wortstellung und Bedeutung*. Akten des 12. Linguistischen Kolloquiums Pavia 1977. Ed. by Maria-Elisabeth Conte, Anna Giacalone Ramat, and Paolo Ramat. Vol. 1. Tübingen: Niemeyer, 3–11.
- É. Kiss, Katalin (1998). “Identificational Focus Versus Information Focus”. In: *Language* 74, 245–273.
- Enç, Mürvet (1991). “The semantics of specificity”. In: *Linguistic Inquiry* 22.1, 1–25.
- Erteshik-Shir, Nomi (2007). *Information Structure. The Syntax-Discourse Interface*. Oxford: Oxford University Press.
- Forker, Diana (2016). “Toward a typology for additive markers”. In: *Lingua* 180, 69–100.
- Frascarelli, Mara (2018). “The interpretation of *pro* in consistent and partial null-subject languages. A comparative interface analysis”. In: *Null Subjects in Generative Grammar*. Ed. by Federica Cognola and Jan Casalicchio. OUP, 211–239.
- Frascarelli, Mara and Roland Hinterhölzl (2007). “Types of topics in German and Italian”. In: *On Information Structure, Meaning and Form. Generalizations across languages*. Ed. by Kerstin Schwabe and Susanne Winkler. Amsterdam and Philadelphia: John Benjamins, 87–116.
- Georgieva, Ekaterina (2022a). “Fake possessives, partitives and definites sharing the same D”. Handout at GLOW 45, 28.4.2022.
- (2022b). “Wh-expressions and sluicing in Mari”. Handout for CIFU XIII, 25.8.2022.
- Georgieva, Ekaterina, Martin Salzmann, and Philipp Weisser (2021). “Negative verb clusters in Mari and Udmurt and why they require postsyntactic top-down word-formation”. In: *Natural Language and Linguistic Theory* 39, 457–503.
- Givón, Talmy (1983). “Topic Continuity in Discourse: An Introduction”. In: *Topic Continuity in Discourse. A Quantitative Cross-Language Study*. Ed. by Talmy Givón. Amsterdam: John Benjamins, 1–41.

- Grenoble, Lenore (1998). *Deixis and Information Packaging in Russian Discourse*. Amsterdam: John Benjamins.
- Grubic, Mira (2014). *Lakka and the Ghosts*. URL: http://totemfieldstoryboards.org/stories/lakka_and_the_ghosts/ (visited on 06/22/2022).
- (2022). “Some remarks on the Storyboard task”. In: *Semantic Fieldwork Methods* 4.1. Special issue (part 1): Collecting semantic data: A sample of individual practices, Article 2.
- Guseva, Elina and Philipp Weisser (2018). “Postsyntactic reordering in the Mari nominal domain. Evidence from Suspended Affixation”. In: *Natural Language and Linguistic Theory* 36, 1089–1127.
- Gyuris, Beáta (2012). “The information structure of Hungarian”. In: *The Expression of Information Structure*. Ed. by Manfred Krifka and Renate Musan. Berlin and Boston: De Gruyter Mouton, 159–186.
- Halász, Ignác (1883). “Orosz-lapp nyelvtani vázlat”. In: *Nyelvtudományi Közlemények* 17, 1–45.
- Halliday, M. A. K. (1967). *Intonation and Grammar in British English*. Janua Linguarum. Series Practica 48. Berlin and Boston: De Gruyter Mouton.
- Hammer, Luan and Jeremy Bradley (2022). “Mari morpheme order revisited: a corpus-based analysis”. In: *Nordlyd* 46.1. Ed. by Lene Antonsen, Sjur Nørstebø Moshagen, and Øystein A Vangsnes, 59–74. DOI: 10.7557/12.6373.
- Herring, Susan C. (1990). “Information structure as a consequence of word order type”. In: *Annual Meeting of the Berkeley Linguistics Society* 16.1, 163–174.
- Herring, Susan C. and John C. Paolillo (1995). “Focus Position in SOV languages”. In: *Word order in discourse*. Ed. by Pamela Downing and Michael Noonan. Amsterdam: John Benjamins, 163–198.
- Hesselbäck, André (2005). *Tatar and Chuvash Code-copies in Mari*. Uppsala: Uppsala Universitet.
- Hirvonen, Johannes and Jeremy Bradley (to appear). “Negative Concord in Mari”. In: *Strict Negative Concord in Slavic and Finno-Ugric. Structure, Licensing, and Locality Conditions for Negative Expressions*. Ed. by Gréte Dalmi, Jacek Witkoś, and Piotr Ceglowski. Vol. 148. Studies in Generative Grammar.

- Holmberg, Anders (2010). "Null Subject Parameters". In: *Parametric Variation. Null Subjects in Minimalist Theory*. Ed. by Theresa Biberauer et al. Cambridge: Cambridge University Press, 88–124.
- Hyun-Oak Kim, Alan (1988). "Preverbal Focusing and Type XXIII Languages". In: *Studies in Syntactic Typology*. Ed. by Edith Moravcsik, Michael Hammond, and Jessica Wirth. Amsterdam: John Benjamins, 149–171.
- П'ин (1986) = Ильин, М. И. (1986). "Интонационная интерференция в русской речи мари". In: *Вопросы марийского языка. Грамматика и лексикология*. Ed. by Н.И. Исанбаев. [Intonational interference in Russian speech of Mari]. Йошкар-Ола: МарНИИ, 85–97.
- Janda, Gwen Eva (2015). "Northern Mansi Possessive Suffixes in Non-Possessive Function". In: *ESUKA – JEFUL* 6.2, 243–258.
- Janda, Gwen Eva, Johanna Laakso, and Helle Metslang (2022). "Person marking". In: *The Oxford Guide to the Uralic Languages*. Ed. by Marianne Bakró-Nagy, Johanna Laakso, and Elena Skribnik. Oxford: Oxford University Press, 894–903.
- Juutinen, Markus (2019). "Koltansaamen myönteiset ja kielteiset fokuspartikkelit". In: *Suomalais-Ugrilaisen Seuran Aikakauskirja* 97, 75–107.
- Kangasmaa-Minn, Eeva (1998). "Mari". In: *The Uralic Languages*. Ed. by Daniel Abondolo. Routledge, 219–248.
- Kazyro (2006) = Казыро, Алексей (2006). "Выражение ремы подлежащим в простом повествовательном предложении марийского языка". In: *Финно-Угроведение* 1. [Expression of the rheme by the subject in the simple narrative sentence in Mari], 91–100.
- Kert (1971) = Керт, Георгий Мартынович (1971). *Саамский язык (кильдинский диалект): Фонетика, морфология, синтаксис*. [The Saami language (Kildin dialect): Phonetics, Morphology, Syntax.] Ленинград: Наука.
- Khomchenkova (2022) = Хомченкова, Ирина (2022). "Дискурсивные функции посессивного показателя третьего лица единственного числа в горномарийском языке". In: *Урало-алтайские исследования* 45.2. [Discursive functions of the 3rd person singular possessive suffix in Hill Mari], 148–168.

- Klumpp, Gerson and Elena Skribnik (2022). “Information structuring”. In: *The Oxford Guide to the Uralic Languages*. Ed. by Marianne Bakró-Nagy, Johanna Laakso, and Elena Skribnik. Oxford: Oxford University Press, 1018–1036.
- Körtvély, Erika (2010). “On the Functions of Possessive Suffixes in Tundra Nenets: Possession and Semantic-Pragmatic Definiteness”. In: *Finnisch-Ugrische Mitteilungen* 32/33, 321–344.
- Krifka, Manfred (2008). “Basic Notions of Information Structure”. In: *Acta Linguistica Hungarica* 55.3-4, 243–276.
- Krifka, Manfred and Renate Musan (2012). “Information Structure: Overview and linguistic issues”. In: *The Expression of Information Structure*. Ed. by Manfred Krifka and Renate Musan. Berlin and Boston: De Gruyter Mouton, 1–43.
- Lambrecht, Knud (1994). *Information Structure and Sentence Form. Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.
- Lehiste, Ilse et al. (2005). *Meadow Mari Prosody*. Vol. 2. Linguistica Uralica Supplementary Series. Tallinn: Teaduste Akadeemia Kirjastus.
- Leinonen, Marja (2002). “Influence of Russian on the syntax of Komi”. In: *Finnisch-Ugrische Forschungen* 57.1–3, 195–358.
- Littell, Patrick (2010a). *Animal Party (Agent Focus)*. URL: https://totemfieldstoryboards.org/stories/animal_party/v1_0/ (visited on 06/22/2022).
- (2010b). *Animal Party (Theme Focus)*. URL: https://totemfieldstoryboards.org/stories/animal_party/ (visited on 06/22/2022).
- Luutonen, Jorma (1997). *The Variation of Morpheme Order in Mari Declension*. SUST 226. Helsinki: Suomalais-Ugrilainen Seura.
- Markus, Elena (2022). “Diskursusepartiklid isuri keele Soikkola murdes”. In: *Keel ja kirjandus* 65.1–2, 74–94.
- Matthewson, Lisa (2004). “On the Methodology of Semantic Fieldwork”. In: *IJAL* 70.4, 369–415.
- Maytinskaya (1982) = Майтинская, Клара Евгеньевна (1982). *Служебные слова в финно-угорских языках*. [Function words in the Finno-Ugric languages]. Москва: Наука.
- McCoy, Svetlana (2003). “Unifying the Meaning of Multifunctional Particles. The Case of the Russian *že*”. In: *U. Penn Working Papers in Linguistics* 9.1, 123–135.

- McFadden, Thomas (2004). "The position of Morphological Case in the Derivation. A Study on the Syntax-Morphology Interface". PhD Dissertation. University of Pennsylvania.
- Mitchell, Erika (2006). "The morpho-syntax of negation and the positions of NegP in the Finno-Ugric languages". In: *Lingua* 116, 228–244.
- Moseley, Christopher, ed. (2010). *Atlas of the World's Languages in Danger*. URL: <https://web.archive.org/web/20220531155026/http://www.unesco.org/tools/fileretrieve/43fdd320.pdf>.
- Neeleman, Ad and Reiko Vermeulen (2013). "The Syntactic Expression of Information Structure". In: *The Syntax of Topic, Focus, and Contrast: An Interface-based Approach*. Ed. by Ad Neeleman and Reiko Vermeulen. Berlin and Boston: De Gruyter Mouton, 1–38.
- Nelson, Diane, Elena Vedernikova, and Jeremy Bradley (to appear). "The Syntax of Meadow Mari: Some issues around TAM and phrase structure". In: *Uralic Syntax*. Ed. by Anne Tamm and Anne Vainikka. Cambridge: Cambridge University Press.
- Nikolaeva, Irina (2003). "Possessive affixes in the pragmatic structuring of the utterance: Evidence from Uralic". In: *International Symposium on Deictic Systems and Quantification in Languages Spoken in Europe and North and Central Asia: Collection of papers*. Ed. by Pirkko Suihkonen and Bernard Comrie. Izhevsk and Leipzig, 130–145.
- Öztürk, Balkız (2013). "Postverbal Constituents in SOV Languages". In: *Theoretical Approaches to Disharmonic Word Order*. Ed. by Theresa Biberauer and Michelle Sheehan. OUP, 270–305.
- Padučeva, Elena (1988). "La particule že: sémantique, syntaxe et prosodie". In: *Les particules énonciatives en russe contemporain* 3, 11–44.
- Panov, Vladimir (2020). "The marking of uncontroversial information in Europe: presenting the enimitive". In: *Acta Linguistica Hafniensia* 52.1, 1–44.
- Pengitov (1961) = Пенгитов, Н. Т. (1961). "Порядок слов в предложении в марийском языке". In: *Труды МарНИИЯЛИ* Выпуск XV. [Sentential word order in Mari], 113–124.
- Pleshak (2019) = Плешак, Полина (2019). "Морфосинтаксис именной группы в финно-угорских языках Поволжья". [Morphosyntax of the noun phrase

- in the Finno-Ugric languages of the Volga region. Master's Thesis. Москва: МГУ.
- Pomozi, Péter (1997). *Satzwertige Partizipien auf -mE/-mAš/δâmE im Tscheremissischen: Problematik und Klassifikation*. Szombathely: Savariae.
- Pregla, Andreas (2023). "On the Homogeneity of Word Order Variability in OV Languages". PhD Dissertation. Universität Potsdam.
- Räsänen, Martti (1920). *Die tschuwassischen Lehnwörter im Tscheremissischen*. SUST 48. Helsinki: Suomalais-ugrilainen seura.
- (1923). *Die tatarischen Lehnwörter im Tscheremissischen*. SUST 50. Helsinki: Suomalais-ugrilainen seura.
- Reinhart, Tanya (1981). "Pragmatics and Linguistics: An Analysis of Sentence Topics". In: *Philosophica* 27.1, 53–94.
- Riese, Timothy et al. (2019). *Mari (марий йылме): An Essential Grammar for International Learners*. [Draft version]. Vienna. URL: grammar.mari-language.com.
- Rießler, Michael (2007). "Grammatical borrowing in Kildin Saami". In: *Grammatical Borrowing in Cross-Linguistic Perspective*. Ed. by Yaron Matras and Jeanette Sakel. Berlin and New York: De Gruyter Mouton, 229–244.
- Roberts, Craige (2019). "Topics". In: *Semantics – Sentence and Information Structure*. Ed. by Paul Portner, Claudia Maienborn, and Klaus von Heusinger. Berlin and Boston: De Gruyter Mouton, 381–412.
- Rooth, Mats Edward (1985). "Association With Focus". PhD Dissertation. University of Massachusetts.
- Rozhanskiy, Fedor (2013). "Morphophonological Nature of Mari Accentuation as Viewed From the Uralic Perspective". In: *Linguistica Uralica* 49.3, 184–207.
- Saarinen, Sirkka (1997). "Language contacts in the Volga region. Loan suffixes and calques in Mari and Udmurt". In: *Language in time and space – Studies in honour of Wolfgang Viereck on the occasion of his 60th birthday*. Ed. by Heinrich Ramisch and Kenneth Wynne. Stuttgart: Franz Steiner Verlag, 388–396.
- (2008). "Turvaako kielilaki kielen aseman: suomalais-ugrilaisten kielten uhanalaisuus". In: *Murros: suomalais-ugrilaiset kielet ja kulttuurit globalisaation*

- paineissa*. Ed. by Sirkka Saarinen and E. Herrala. *Uralica Helsingiensia* 3. Helsinki: Helsingin yliopisto and Suomalais-Ugrilainen Seura, 38–46.
- Saarinen, Sirkka (2015). “Negation in Mari”. In: *Typological Studies in Language* 108. Ed. by Matti Miestamo and Beáta Wagner-Nagy, 325–352.
- (2022). “Mari”. In: *The Oxford Guide to the Uralic Languages*. Ed. by Marianne Bakró-Nagy, Johanna Laakso, and Elena Skribnik. Oxford: Oxford University Press, 432–470.
- Serdobol’skaya and Toldova (2012) = Сердобольская, Н.В. and С.Ю. Толдова (2012). “Тексты на луговом марийском языке (староторъяльский говор)”. In: *Финно-угорские языки. Фрагменты грамматического описания. Формальный и функциональный подходы*. Ed. by Н.В. Сердобольская et al. [Texts in Meadow Mari (Starotorjal dialect)]. Москва: Рукописные памятники Древней Руси, 709–745.
- Siegl, Florian (2021). “Die waldenzischen Possessivsuffixe aus taxonomischer und diskursorientierter Perspektive”. In: *Ural-Altaische Jahrbücher Neue Folge* 28, 44–116.
- Simonenko and Leont’ev (2012) = Симоненко, Александра and Алексей Леонтьев (2012). “Морфосинтаксис именного комплекса в финно-пермских языках: Анализ в рамках программы минимализма”. In: *Финно-угорские языки. Фрагменты грамматического описания. Формальный и функциональный подходы*. Ed. by Н.В. Сердобольская et al. [Morphosyntax of the nominal complex in the Finno-Permic languages: An analysis in the framework of the Minimalist Program]. Москва: Рукописные памятники Древней Руси, 259–337.
- Simonenko, Alexandra (2014). “Microvariation in Finno-Ugric possessive markers”. In: *Proceedings of the 43th annual meeting of the North East Linguistic Society (NELS 43)*. Ed. by H.-L. Huang, E. Poole, and A. Rysling. Vol. 2, 127–140.
- (2017). “Towards a semantic typology of specific determiners”. In: *Proceedings of the 21th Amsterdam Colloquium*. Ed. by Alexandre Cremers, Thom van Gessel, and Floris Roelofsen. Amsterdam: Institute for Logic, Language and Computation, 425–434.

- Simonenko, Alexandra (2020). "Existential Possession in Meadow Mari". In: *Approaches to Predicative Possession. The View from Slavic and Finno-Ugric*. Ed. by Gréte Dalmi, Jacek Witkoś, and Piotr Ceglowski. London: Bloomsbury Academic, 162–185.
- Skopeteas, Stavros and Gisbert Fanselow (2010). "Focus types and argument asymmetries". In: *Comparative and Contrastive Studies of Information Structure*. Ed. by Carsten Breul and Edward Göbbel. John Benjamins, 169–197.
- Skopeteas, Stavros, Ines Fiedler, et al. (2006). *Questionnaire on Information Structure: Reference Manual*. Interdisciplinary Studies on Information Structure 4. Universitätsverlag Potsdam.
- Surányi, Balázs et al. (to appear). "Information Structure in Finno-Ugric languages". In: *Uralic Syntax*. Ed. by Anne Tamm and Anne Vainikka. Cambridge: Cambridge University Press.
- Tánczos, Orsolya (2010). "Szórendi variációk és lehetséges okaik az udmurtban". In: *Nyelvtudományi Közlemények* 107. [Word order variation and its possible causes in Udmurt], 218–228.
- TFS Working Group (2011). *Bake-Off*. URL: https://totemfieldstoryboards.org/stories/bake_off/ (visited on 06/22/2022).
- Timofeyeva (1961) = Тимофеева, В. Т. (1961). *Современный марийский язык. Синтаксис сложного предложения*. [The contemporary Mari language. Syntax of the complex clause]. Йошкар-Ола: Марийское книжное издательство.
- Trosterud, Trond et al. (2023). *Corpus of Literary Mari*. URL: https://gtweb.uit.no/u_korp/?mode=mhr (visited on 09/13/2023).
- Truckenbrodt, Hubert (1995). "Phonological Phrases: Their Relation to Syntax, Focus, and Prominence". PhD Thesis. Cambridge, MA: MIT.
- Tuzharov (1987) = Тужаров, Г. М. (1987). *Грамматические категории имени существительного в марийском языке*. Вопросы советского финно-угроведения. [Grammatical categories of the noun in Mari]. Йошкар-Ола: Марийское книжное издательство.
- Valova and Slyusar' (2017) = Валова, Едвокия Алексеевна and Наталия Анатольевна Слюсарь (2017). "Исследование синтаксических свойств энклитики же: Корпусный и экспериментальный подход". In: *Вопросы языко-*

- знания 2. [A study on the syntactic properties of the enclitic *že*: A corpus-based and experimental approach], 33–48.
- Valtonen, Taarna, Merja Fofonoff, and Markus Juutinen (2018). “Anarâs- já nuort- tâlâškielâ partikkâl še”. In: *Dutkansearvvi diedalaš áigečála* 2, 129–158.
- Van der Wal, Jenneke (2016). “Diagnosing focus”. In: *Studies in Language* 40.2, 259–301.
- Vasikova (1982) = Васикова, Лидия (1982). “Структурные схемы предложений в марийском языке”. In: *Советское финно-угроведение* XVIII. [Sentence types in Mari], 273–278.
- Vilkuna, Maria (1998). “Constituent Order in European Uralic”. In: *Constituent Order in the Languages of Europe*. Ed. by Anne Siewierska. Berlin: de Gruyter, 173–233.
- (2022). “Word order”. In: *The Oxford Guide to the Uralic Languages*. Ed. by Marianne Bakró-Nagy, Johanna Laakso, and Elena Skribnik. Oxford: Oxford University Press, 950–960.
- Volkova, Anna (2014). “Licensing Reflexivity. Unity and variation among selected Uralic languages”. Doctoral Dissertation. Universiteit Utrecht.
- (2017a). “Reflexivity in Meadow Mari: Binding and Agree”. In: *Studia Linguistica* 71.1–2, 178–204.
- (2017b). “Syntactic Structure of Participial Clauses in Meadow Mari”. HSE Moscow Working Papers BRP 54/LNG/2017.
- Volkova, Anna et al. (2023). *Spoken corpus of Meadow Mari (as spoken in the village of Saryj Toryal, Novyj Torjal district, Mari El Republic, Russia)*. URL: http://lingconlab.ru/spoken_meadow_mari/ (visited on 08/31/2023).
- Wagner-Nagy, Beáta (2018). *A Grammar of Nganasan*. Leiden and Boston: Brill.
- Yegorova (1981) = Егорова, А. А. (1981). “Факторы, влияющие на порядок слов марийского литературного текста (на материале диалогической речи)”. In: *Советское финно-угроведение* XVII. [Factors influencing the word order in literary Mari texts: On the material of dialogue speech], 271–282.
- Yurayong, Chingduang (2020). “Postposed Demonstratives in Finnic and North Russian Dialects”. Doctoral Dissertation. University of Helsinki.
- Zakirova (2019) = Закирова, Айгуль (2019). “Частицы эмфатической идентичности в поволжском языковом союзе”. Master’s Thesis. Москва: HSE.

- Zamyatin, Konstantin (2016). "Russia's Minority Education and the European Language Charter". In: *Linguistic genocide or superdiversity. New and old language diversities*. Ed. by Janne Saarikivi and Reetta Toivanen. Clevedon: Multilingual Matters, 251–284.
- Zayzon, Réka (2015). "Observations on Non-Possessive Usages of Personal Markers (Possessive Suffixes) in Nganasan". In: *ESUKA – JEFUL* 6.2, 259–278.
- Zhornik and Rozhanskiy (2021) = Жорник, Дарья and Федор Рожанский (2021). "Длительность горномарийских гласных как просодический фактор". In: *Linguistica Uralica* 3. [The length of Hill Mari vowels as prosodic factor], 187–212.
- Zimmermann, Malte and Caroline Féry (2009). "Introduction". In: *Information Structure*. Ed. by Malte Zimmermann and Caroline Féry. Oxford: Oxford University Press, 1–12.
- Zubova, Iuliia et al. (2022). "Partikkel *ved'* udmurdi, komi ja ersa keeles". In: *Keel ja Kirjandus* LXV.1–2, 110–129. doi: 10.54013/kk770a7.
- Perepis'* (2020–2021). *Владение языками и использование языков населением*. [Language proficiency and language usage]. Росстат. URL: https://rosstat.gov.ru/storage/mediabank/Tom5_tab4_VPN-2020.xlsx (visited on 06/12/2023).

Deutschsprachiges Abstract

Diese Arbeit ist die erste allgemeine Beschreibung der morphosyntaktischen Markierung von Fokus, Topik und Kontrastivem Topik in der uralischen Sprache Wiesenmarisch. Auf Basis von elizitierten Daten und aktuellen Sprachdaten aus Korpora werden in dieser Arbeit die wichtigsten Arten beschrieben, wie diese informationsstrukturellen Kategorien morphosyntaktisch markiert werden.

Informationsfokus auf einzelnen Konstituenten kann im Marischen markiert werden, indem die betreffende Konstituente dem finiten Verb direkt vorangeht. Eine Ausnahme zur üblichen präverbalen Position von Fokus ist Präsentationsfokus, der postverbal auftreten kann. In der gesprochenen Sprache ist jedoch prosodische *in situ*-Markierung von Informationsfokus die üblichere Strategie. Diese Art von Fokusmarkierung ist unterspezifiziert und kann Fokus in Antworten auf Konstituentenfragen, korrektiven Fokus und Identifikationsfokus markieren. Identifikationsfokus kann außerdem durch einen Pseudocleft markiert werden. Das Wiesenmarische ist in dieser Hinsicht eine recht typische streng rechtshäuptige SOV-Sprache, da Fokus in diesem Sprachtyp üblicherweise direkt präverbal auftritt. Da das Wiesenmarische aber auch über die Möglichkeit verfügt, Fokus *in situ* zu markieren, kann es nicht als *strict position language* nach Büring (2016) klassifiziert werden.

Topiks sind in der linken Peripherie des marischen Satzes aufzufinden. Subjekttopiks werden in einer Position weit links im Satz platziert, wie auch andere Satzglieder, wenn sie Topiks in einer thematischen Progression sind. Konstituenten werden auch eher als Antezedenten für Nullsubjekte in einem folgenden Satz interpretiert, wenn sie die linkeste Konstituente in einem Satz sind. Satztopiks können des Weiteren über linksversetzte Konstituenten gewechselt werden. Diese sind üblicherweise prosodisch und morphosyntaktisch vom Hauptsatz getrennt. Linksversetzte Topiks können mit einer Topikmarkerphrase *manmašte* oder mit dem diskursiv gebrauchten 3SG Possessivsuffix *že* markiert werden. Im postverbalen Feld wiederum können Konstituenten aus verschiedenen Gründen erscheinen – als nachträgliche Einfälle oder Vervollständigungen des vorigen Satzes sowie als bekannte Topiks. In dieser Position erscheinen Konstituenten auch epiphänomenal als Resultat von Verbverschiebung nach links. Weiters können

prosodisch schwere Konstituenten in das postverbale Feld extraponiert werden. Das Wiesenmarische verhält sich im Allgemeinen also auch recht typisch für eine SOV-Sprache, was die Position der Topiks und Nicht-Topiks betrifft.

Neben der rein syntaktischen Markierung von Fokus und Topiks verfügt das Wiesenmarische über ein Possessivsuffix der dritten Person Singular, *žE*, das neben der Besitzmarkierung auch in diskursiven Funktionen gebraucht werden kann. Aus der bisherigen Forschung ist bekannt, dass das Suffix eine kontrastierende Funktion hat. In dieser Arbeit wird gezeigt, dass *žE* auch einen Wechsel des Satztopiks ohne Kontrast markieren kann. Zusätzlich wird argumentiert, dass einige Funktionen von *žE* durch strukturellen Einfluss des homophonen russischen Diskurspartikels *že* erklärt werden können.