

The information structure of Meadow Mari

© 2024

Johannes Hirvonen

Ludwig-Maximilians-Universität in Munich, Munich, Germany;
johannes.hirvonen@campus.lmu.de

Abstract: This paper provides a first descriptive overview of morphosyntactic strategies of marking information structure in contemporary Meadow Mari, based on both elicited and corpus data. I describe syntactic strategies of marking focus, topics, and contrastive topics in the language. Focus is variably realized *in situ*, in immediately preverbal or postverbal position, or via pseudocleft. The strategies differ in terms of their “markedness”, in the sense that more marked strategies encode subsets of the types of foci which the less marked strategies encode. Nevertheless, no type of focus requires a certain strategy. Topics are placed in the left periphery of the clause. In addition to these syntactic strategies of information structure marking, the suffix -ŽE is described as a morpheme denoting the selection of an individual from a superset and contrasting it with other individuals — essentially as a marker of contrastive topics. With its broad scope, the paper lays the groundwork for comparative research on information structure in Meadow Mari and other Uralic languages and for more detailed studies on information structure in Meadow Mari itself.

Keywords: contrast, focus, information structure, Mari, possession, topic, Uralic

Acknowledgements: This paper is an updated and condensed version of the author’s Master’s thesis (2023). The thesis and this paper have benefitted greatly from discussions with and comments by many people, including (in alphabetical order) Jeremy Bradley, Daniel Buring, Ekaterina Georgieva, Andreas Pregla, Ksenia Shagal, Elena Skribnik, Kriszta Szendrői, Elena Vedernikova, and an anonymous reviewer — thank you.

For citation: Hirvonen J. The information structure of Meadow Mari. *Voprosy Jazykoznanija*, 2024, 5: 97–115.

DOI: 10.31857/0373-658X.2024.5.97-115

Информационная структура в луговом марийском языке

Йоханнес Хирвонен

Мюнхенский университет Людвига-Максимилиана, Мюнхен, Германия;
johannes.hirvonen@campus.lmu.de

Аннотация: В статье представлен обзор морфосинтаксических стратегий маркирования информационной структуры в современном луговом марийском языке, основанный на элицированных и корпусных данных. Мы описываем синтаксические стратегии маркирования фокуса, топиков и контрастивных топиков в языке. Фокус может быть выражен *in situ*, в первой предглагольной или заглагольной позиции или с помощью псевдоклефта. Данные стратегии различаются «маркированностью», в том смысле, что более маркированные стратегии выражают такие подмножества типов фокуса, которые выражают и менее маркированные стратегии. Тем не менее ни один из типов фокуса не предполагает особого средства выражения. Топики находятся на левой периферии клаузы. Помимо указанных стратегий маркирования информационной структуры, суффикс -ŽE описывается как морфема, маркирующая выбор индивида из объемлющего множества и отношение контраста с другими индивидами — по сути маркером контрастивного топика. Учитывая широкую перспективу, статья закладывает основу для сравнительного изучения информационной

структуры в луговом марийском и других уральских языках, а также для более детального исследования информационной структуры в самом луговом марийском.

Ключевые слова: коммуникативная структура, контраст, марийские языки, посессивность, топик, уральские языки, фокус

Для цитирования: Hirvonen J. The information structure of Meadow Mari. *Voprosy Jazykoznanija*, 2024, 5: 97–115.

DOI: 10.31857/0373-658X.2024.5.97-115

1. Introduction

Although it has previously been recognized that information structure is an important factor governing variation in word order in Meadow Mari, e.g., in [Chkhaidze 1941: 29–31; Pengitov 1961: 122; Vilkuna 1998; Saarinen 2022: 463], information structure marking in the language has not been systematically described so far. This paper provides a first descriptive overview of information structure marking in contemporary Meadow Mari, based on both elicited and corpus data. At the center of this investigation are types of morphosyntactic marking which interact with the information structural notions of focus, topic, and contrastive topic.

For focus marking, Meadow Mari employs four different strategies: (1) prosodic in-situ marking, (2) immediately preverbal constituent focus and clause-initial verbal focus, (3) postverbal focus, and (4) focussing via pseudocleft. The focus marking strategies differ in terms of their “markedness”, in the sense that more marked strategies encode subsets of the types of foci which the less marked strategies encode: constituent focus is mostly marked in situ; the immediately preverbal and postverbal positions have restrictions as to when non-object constituents may appear there (e.g. in correction contexts); and finally, pseudoclefts as the structurally most complex type of focus marking are employed only for identificational focus. Furthermore, focussed verbs can appear in a left-peripheral position. These facts show that Meadow Mari does not have a designated focus position, and that no type of focus *requires* a certain strategy. Topics are located in the left periphery. This is true for both aboutness topics and contrastive topics. The most important morphological means of information structure marking is the morpheme *-žE*.¹ This suffix has several discursive functions, which include the selection of an individual from a superset and triggering a contrastive interpretation. In this function *-žE* is used for marking contrastive topics and the shift to new aboutness topics.

The paper is structured as follows: After a short description of the data sources in Section 2, I discuss focus, (contrastive) topics, and the functions of the morpheme *-žE* in Sections 3, 4, and 5, respectively. In Section 3, I describe focus marking via the four available strategies (in-situ focus, immediately preverbal, immediately postverbal placement, pseudocleft) and investigate their functions. Section 4 contains a short overview of the placement of topics in the left periphery of the clause. In Section 5, I examine the discursive functions of *-žE* in more detail. Section 6 summarizes the most important findings.

2. Data

The variety investigated in this study is contemporary (written and spoken) Meadow Mari. Meadow Mari is a Uralic language with 318,495 speakers, according to the 2020 All-Russian Population Census², spoken mainly in the titular republic Mari El in the Russian Federation.

¹ *žE* in capital letters is used as a shorthand for various allomorphs of the suffix, e.g. *-že*, *-žö*, *-še*, etc. (see Section 5).

² Available online at: <https://rosstat.gov.ru/vpn/2020> (last accessed 04.07.2024).

The data for this study comes from different sources: First, I draw from already existing work on Meadow Mari, including data gathered for my Master's thesis on the same topic [Hirvonen 2023]. Another source of data were corpora of Meadow Mari. These were used mainly for the investigation of topic marking and the functions of the morpheme *-ŽE*. The largest corpus of Meadow Mari which allows to display context is the Corpus of Contemporary Written Literary Meadow Mari [Arkhangelskiy 2019]. This corpus has a size of 5.53 million tokens and mostly consists of contemporary newspaper texts. Context can be shown up to four sentences before and after the target sentence. A similar corpus is the Corpus of Social Media in Meadow Mari which consists of 3.59 million words of written text from the social network “VKontakte” [Ibid.].³ In addition to these corpora of written Meadow Mari, there are two more important sources of contemporary spoken Meadow Mari: a small printed corpus of data from the village Saryj Torjal collected in the year 2000 [Serdobolskaya, Toldova 2012], and the larger Spoken Corpus of Meadow Mari, consisting of data from 2000–2004 and 2018 gathered in the same village [Volkova et al. 2024].⁴ The section on topics in this paper relies mostly on the smaller printed corpus. Thirdly, for data on all types of information structure marking investigated in this paper, I worked with a native speaker consultant. My consultant is a female speaker of Meadow Mari in the age group 40–50 and grew up in the central Meadow Mari (Morki / Sernur) dialect area. Most data in this study is elicited.

3. Focus

This section deals with strategies of focus marking in Meadow Mari. Focus is understood here as indicating the presence of alternatives relevant for the interpretation of the marked linguistic expression, see, e.g., [Krifka 2008]. In the descriptively oriented approach taken here, I investigate the marking strategies in different contexts known to trigger focus marking. Relevant contexts include the expression answering a *wh*-question or alternative question (1), and correction contexts (2). The expression marked with focus is formatted with SMALL CAPITAL LETTERS:

- (1) — *What do you put in your pasta sauce? / Do you put tarragon or thyme in your pasta sauce?*
— I put *THYME* in my pasta sauce. [Büring 2009: 178–179]
- (2) *I put THYME in my pasta sauce, not tarragon.* [Büring 2009: 179]

The focus marked expressions can be of different sizes, e.g. whole clauses in the case of clausal focus, constituents, or single constituents, as in the previous examples. Another special type investigated here, *verum focus*, targets the truth value of a sentence. Presentational focus, as in (3), has been explained as indicating a covert question around which discourse is structured, such as *What was there?*, see, e.g., [Krifka 2008; Büring 2009; van der Wal 2016; 2021].

- (3) *Once upon a time, there was a PRINCESS.* [Krifka 2008: 251]

Finally, the last distinction I make here is the distinction between information focus and identificational focus (sometimes also called exhaustive or contrastive focus) [É. Kiss 1998]. The latter identifies the subset of elements for which the predicate can hold, and thus excludes the other alternatives.

³ These two corpora are available online at: <https://meadow-mari.web-corpora.net/> (last accessed 04.07.2024).

⁴ Available online at: http://lingconlab.ru/spoken_meadow_mari/ (last accessed 04.07.2024). I want to thank one of the creators of the corpus, Aigul Zakirova, here for sending me the whole corpus data. The data in this corpus includes part of the data from Serdobolskaya and Toldova [2012], but I cited the source where I initially found the respective example.

It has long been recognized that the position of foci, and to some extent topics and backgrounded constituents, is associated with the word order type, viz. the headedness of a language. In head-final languages, constituent foci are typically found in the immediately preverbal position, see, e.g., [Dezső 1978; Kim 1988; Herring 1990], among others, and the postverbal position is often occupied by some type of backgrounded material, although there is much variation [Pregla 2024]. Meadow Mari exhibits a number of characteristics typical of a rigid head-final language: nouns follow their modifiers, compounds are head-final, the language employs postpositions, modal verbs follow lexical verbs, and clausal complements generally precede the main clause. The neutral word order is SOV with the dative object typically preceding the accusative object. Time and place adverbials are usually placed left in the clause; they may occur before or after subjects or topical constituents. Irrespective of their position relative to other parts of the sentence, time adverbials precede place adverbials [Chkhaidze 1941: 29–31; Pregla 2024: 23]. In the Uralic context, Meadow Mari word order has been characterized as “quite flexible” SOV as regards possible word order variation [Vilkuna 2022: 950], but as rigid in terms of argument reordering for focus purposes [Pregla 2024: 92]. A clause showing the word order in a neutral context is given in (4).⁵

- (4) *Tače kolhoz pasu-što vič rveze traktor dene mlandə-m udarno*
 today kolkhoz field-INE five lad tractor with land-ACC enthusiastically
kural-ə.
 plough-3PL
 ‘Today on the kolkhoz field, five lads enthusiastically plough the land with a tractor.’
 [Chkhaidze 1941: 30]

Previous research indicates that focussed constituents in Meadow Mari, including both subjects and objects, can either occupy the preverbal position, see, e.g., [Chkhaidze 1941: 29–31; Pengitov 1961: 122–123; Pregla 2024: 92–95], or the same position as in the unmarked clause, i.e. stay in situ [Nelson et al. (to appear); Saarinen 2022: 463; Vilkuna 1998: 195]. Additionally, two studies have found that “contrastively” focussed constituents may occupy a clause-final position as well [Vilkuna 1998: 195; Pregla 2024: 176–177]. In this section, I will discuss focus, paying attention to the interpretational effect of different morphosyntactic marking strategies. There are four different strategies employed to express focus: (1) in-situ focus, (2) immediately preverbal focus and clause-initial verbal focus, (3) postverbal focus, (4) usage of a special pseudocleft construction. Taking these four constructions as a starting point, this section discusses the focus marking strategies and makes an attempt to delimit the discourse contexts in which they occur.

3.1. In-situ focus marking

In Meadow Mari, focus can be marked prosodically. In the case of constituent focus, the marking involves a rise on the stressed vowel of the focussed expression, see the discussion in [Hirvonen 2023: 51–54]. In-situ focus via prosodic marking is a versatile strategy: it can be employed for both clausal focus and constituent focus, in answers to *wh*-questions, and in correction contexts. Furthermore, this marking strategy is underspecified for the type of focus and can mark both information and identificational focus. My consultant preferred in-situ focus in most contexts, but accepted preverbal answer and correction foci if presented with them.

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

⁵ Abbreviations used in the glosses are given at the end of the paper.

- (5) *Mo tǎšte? — 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.’ (elicited)
- (6) *Mo tǎšte 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.’ (elicited)

For constituent focus as answers to *wh*-questions, my consultant preferred prosodic in-situ marking (and answering the question only with the focussed constituent, or together with the verb).

- (7) *Kō tače una-lan kočkəš-əm kond-en? — MASKA tače (una-lan kočkəš-əm) kond-en.*
 who today guest-DAT food-ACC bring-PST2.3SG bear today guest-DAT food-ACC bring-PST2.3SG
 ‘Who brought the guests food today? — The BEAR brought the guests food today.’ (elicited)
- (8) *“Mo-m maska kond-en?” ur jod-o. “Maska KOČKƏŠ-ƏM kond-en,” krab vašešt-əš.*
 what-ACC bear bring-PST2.3SG squirrel ask-PST1.3SG bear food-ACC bring-PST2.3SG crab answer-PST1.3SG
 “‘What did the bear bring?’” asked the squirrel. “The bear brought the FOOD,” answered the crab.’ (elicited)

Focus in correction contexts is realized just like other types of information focus. Both subjects and objects are preferably placed in the same position as in the unmarked clause:

- (9) *Məj təj deč=at šukə-rak kogəl’ə-m əšt-en kert-am. — Uke,*
 1SG 2SG from=ADD much-CMPR pie-ACC do-CVB can-1SG no
MƏJ šukə-rak kogəl’ə-m əšt-en kert-am.
 1SG much-CMPR pie-ACC do-CVB can-1SG
 ‘I can bake more pies than you. — No, I can bake more pies.’ (elicited)
- (10) [‘In the garden, is the woman hitting a girl?’]
Uke, sad-əšte üdəramaş RVEZƏ-M šel-eš.
 no garden-INE girl boy-ACC hit-3SG
 ‘No, in the garden, the girl is hitting a BOY.’ (elicited)

Note, furthermore, that prosodic marking can express identificational focus as well, where focus identifies the exhaustive subset of which the predicate phrase holds [É. Kiss 1998] (for another strategy see Section 3.3):

- (11) [In a situation where only Mikhail is holding a rifle:]
MIXAIL pəčal-əm kuč-a.
 Mikhail rifle-ACC hold-3SG
 ‘MIKHAIL is holding the rifle (and not anyone else).’ (elicited)

3.2. Ex-situ focus marking

3.2.1. Immediately preverbal and clause-initial verbal focus

The second strategy of constituent focus marking is placing a focussed constituent in an immediately preverbal position, illustrated with elicited data from Pregla [2024]. The preverbal position is a marked position for non-object constituents as immediately preverbal subjects are

only acceptable when focussed, and not in neutral contexts (12). As mentioned above, my consultant consistently used in-situ focus for subjects but accepted preverbal subject foci if presented with them.⁶

- (12) *Ikeče jal-əšte joča-lan jomak-vlak-əm KOVA / *kova ojlə-š.*
 recently village-INE child-DAT tale-PL-ACC grandmother grandmother tell-PST1.3SG
 ‘GRANDMA / *Grandma recently told stories to the children in the village.’ [Pregla 2024: 94]

Importantly, Pregla [2024] has shown that what matters for the focal interpretation of a constituent is not the absolute position in a clause, but the relative position immediately preceding the verb. The focus-verb string can either appear in the clause-final position or clause medially, with the preceding constituent in focus:

- (13) *Poškudo TEDGEČE ERDENE puə-š joča-lan pört-əštö pərəs-əm.*
 neighbor yesterday morning give-PST1.3SG child-DAT house-INE cat-ACC
 ‘The neighbor gave a cat to the child in the house YESTERDAY MORNING.’ [Pregla 2024: 176]

In a related construction, clause-initial placement of a finite verb also allows focussing the verb itself [Hirvonen 2023: 50–51]. This placement marks types of focus associated with finite verbs, e.g. verum focus targeting the truth value of the sentence (14)–(15).⁷ This construction is commonly associated with correction contexts. Note, however, that for verbal focus a prosodic realization is possible as well, cf. [Hirvonen 2023: 50].

- (14) *Tenij māj Budapest-əške š-əm kaj. — KAJ-əš-əč taj tenij*
 this.year 1SG Budapest-ILL NEG.PST1-1SG go.CNG go-PST1-2SG 2SG this.year
Budapest-əške!
 Budapest-ILL

‘I did not go to Budapest this year. — You DID GO to Budapest this year!’ (elicited)

- (15) *Vrač joča-lan em-əm jüktə-š. — Əš JÜKTÖ*
 doctor child-DAT medicine-ACC give.to.drink-PST1.3SG NEG.PST1.3SG give.to.drink.CNG
vrač joča-lan em-əm!
 doctor child-DAT medicine-ACC

‘The doctor gave the child medicine to drink. — The doctor DID NOT GIVE the child medicine to drink!’ (elicited)

This phenomenon can best be illustrated with data from verb clusters involving displacement of one part of the cluster only, cf. [Georgieva et al. 2021: 463]. The following examples involve a verb cluster consisting of the modal verb *kert-* ‘can’, the negation auxiliary, and the lexical verb *əšte-* ‘do, make’. The order of these three verbs in a neutral context is lexical verb > negation > modal (16).⁸ As in example (15), the negation auxiliary is inflected and selects the invariant connegative form of the modal verb, which in turn selects the converbial form of the lexical

⁶ A reviewer raises the question of whether there are any restrictions on the usage of preverbal focus. My consultant did not categorically refuse a preverbal placement of constituents under focus in *wh*-question contexts, but consistently preferred and chose in-situ placement. I did not discover any strict constraints on the usage of preverbal constituent focus. The reasoning made here, that the preverbal position is a more marked position, thus comes from the preference for in-situ focus in *wh*-contexts and the general observation that the preverbal position is not available for non-focussed subjects, as shown in (12).

⁷ In example (15), the negation auxiliary occupies the clause-initial position together with the lexical verb in the connegative form. See the discussion and references in the remainder of this subsection for more detailed information on this phenomenon.

⁸ The order negation auxiliary > finite verb is unexpected in a head-final language, but found across Uralic languages [Yurayong et al. 2024: 736]. The relevant part for the argument made here is that the negation auxiliary and the verb in the connegative form cannot be separated from each other, but can be separated from a third verbal form. See Georgieva et al. [2021] for an in-depth analysis of this phenomenon.

verb. The negation auxiliary and the verb which it selects form a cluster in which nothing can intervene,⁹ but for focussing purposes this cluster can occupy the clause-initial position while leaving the lexical verb behind in a clause-final position (17):

- (16) *Məj tidə-m niguze əšt-en o-m kert.*
 1SG this-ACC no.way do-CVB NEG-1SG can.CNG

‘I cannot do this in any way.’ (elicited)

- (17) *O-M KERT məj tidə-m niguze əšt-en.*
 NEG-1SG can.CNG 1SG this-ACC no.way do-CVB

‘I CANNOT do this in any way.’ (elicited)

In sum, the position immediately preceding the verb is associated with constituent focus marking, irrespective of where in the clause the verb is placed. Clause-initial placement of the verb is associated with verbal or verum focus. What unites these constructions is that the immediately preverbal position can be employed for constituent focus, and that the verb receives a focussed interpretation if the verb itself is not preceded by any other constituent. In these constructions the number of postverbal elements is in principle unlimited and their ordering is the same as in the neutral context. Furthermore, the pre-focal elements do not receive any special information structural interpretation, cf. [Pregla 2024].

3.2.2. Postverbal focus

Immediately postverbal constituent focus is a more specialized strategy and is available mainly in two contexts — presentational focus and correction focus. My consultant dispreferred postverbal constituent focus in answers to *wh*-questions. A subtype of information focus, presentational focus indicates a covert question in the discourse. This strategy is typically employed in Meadow Mari to introduce a referent which is then continued as an aboutness topic (see Section 4 on topics) in the following clause, cf. *klad* ‘shed, barn’ in example (18), and a person called *Zina* in example (19). The new topics are indicated by boldfacing:

- (18) *Memnan pört vokane ə'le təgaj KLAT. Vot klat... pusta klat,*
 2SG.GEN house besides be.PST1.3SG such shed see shed empty shed
nigō tušto il-en og-əl. Tide klat ümbake küz-en
 nobody there live-CNG NEG.3SG-BE.CNG this shed onto climb-CVB
šinč-ən-am no.
 sit.down-PST1-1SG PTCL

‘Next to our house there was this SHED. Well, a shed... an empty shed, nobody lived there. **This shed**, I climbed up and sat down on it.’ [Serdobolskaya, Toldova 2012: 712]

- (19) [“I had a hard time learning to knit.”]
Məj pər'la tunem-ən-na ZINA dene, Morkinskij rajon-et, tudo peš
 1SG together learn-PST2-1PL Z. with M. district-PX2SG 3SG very
čot pid-eš ə'le. Tudə-n den tunem-əm məj tunem-aš... pid-ən-am,
 much knit-3SG AUX 3SG-GEN with learn-PST1.1SG 1SG learn-INF knit-PST2-1SG
tudo məj-əm tunəkt-en.
 3SG 1SG-ACC teach-PST2.3SG

‘I learned together with **ZINA** from Morkinskiy district, **she** used to knit a lot. **With her** I learned, to learn... I knit, she taught me.’ [Volkova et al. 2024: Biografiya Z. I. E. (Part 2)]

⁹ Jeremy Bradley (p.c.) remarks that this is not entirely as clitics like the additive clitic *=at* can indeed intervene, e.g. *o-m=at kert* [NEG-1SG=ADD can.CNG]. Georgieva et al. [2021] discuss the same phenomenon for Udmurt. In any case, the argument made here is not affected by this additional complication.

Additionally, constituents in an immediately postverbal position can receive a focussed interpretation in correction contexts [Pregla 2024]:¹⁰

- (20) (*Uke*,) *erdene joča-lan pört-əštö pərəs-əm pu-əš AŞNƏŞE PÖRJED!*
 no in.morning child-DAT house-INE cat-ACC give-PST1.3SG caring person
 ‘(No,) A CARING PERSON gave a cat to our child in the house in the morning [and not the neighbor]!’ [Pregla 2024: 186]
- (21) *Aşnəşe pörjeŋ joča-lan pört-əštö pərəs-əm pu-əš ERDENE!*
 caring person child-DAT house-INE cat-ACC give-PST1.3SG in.morning
 ‘A caring person gave a cat to our child in the house IN THE MORNING [and not in the evening]!’ [Pregla 2024: 187]

Postverbal focussing appears to be more marked than in-situ focus, as it is available mainly for presentational and correction focus. Furthermore, the strategy is positionally more restricted as only the focussed constituent is allowed to appear after the verb. The more constituents appear after the verb, the more the sentences are degraded. In the following clause, *pərəsəm* ‘the cat’ occurs in the immediately postverbal position and receives a focussed interpretation. If the verb is further left (and the number of constituents between the verb and the focussed constituent increases), the sentence is degraded (22). This is a clear difference to preverbal focus, where the number of postverbal elements is in principle unlimited (see Pregla [2024] for further discussion).

- (22) *Poškudo (ʔʔpu-əš) erdene (ʔʔpu-əš) joča-lan (ʔpu-əš) pört-əštö*
 neighbor give-PST1.3SG yesterday give-PST1.3SG child-DAT give-PST1.3SG house-INE
 (*pu-əš*) *PƏRƏS-ƏM.*
 give-PST1.3SG cat-ACC
 ‘The neighbor gave A CAT to the child in the morning.’ [Pregla 2024: 187–188]

3.3. Pseudocleft

In addition to the strategies of focus marking underspecified for the type of focus, there is a more marginal strategy to mark exhaustive focus, namely a pseudocleft. This construction involves a clause with a *wh*-phrase in subject position, followed by the determiner *tide* ‘this’ and the focussed constituent. Pseudoclefts are only felicitous if 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 of a pseudocleft is felicitous (23). However, 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 (24) where the predicate “holding a bottle” is true of all people:¹¹

- (23) [In a situation where only Mikhail is holding a rifle:]
Kö pəçal-əm kuč-a, tide MIXAIL.
 who rifle-ACC hold-3SG this Mikhail
 ‘It is MIKHAIL who is holding the rifle.’ (elicited)

¹⁰ In my previous work, I reported judgments about postverbal focus being dispreferred in correction contexts [Hirvonen 2023: 49]. The factors influencing the different judgments are not clear to me, but in light of the available evidence for postverbal focus in correction contexts in Vilkuina [1998] and Pregla [2024], I reject my earlier claim for it being dispreferred in this context.

¹¹ Note that these examples were not produced spontaneously, but elicited. My consultant did not consider the construction very natural, but their judgments about its felicity were very clear. For the parallel Russian construction see, for example, [Burukina et al. 2024: 31–32].

- (24) [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

‘It is JəVAN who is holding the bottle.’ (elicited)

3.4. Summary of focus marking strategies

In sum, there are four different strategies of constituent focus marking: prosodic in-situ marking, immediately preverbal constituent focus and clause-initial verbal focus, immediately postverbal focus, and focussing via pseudocleft. Prosodic in-situ marking is a versatile strategy for focus marking on any constituent, and, furthermore, underspecified for the type of focus as it can encode both information and identificational focus. Preverbal constituent focus is more marked for subject focus, for example, but does not categorically mark a different type of focus. Postverbal focus is commonly employed for presentational focus and in correction contexts. Finally, the pseudocleft is felicitous only for identificational focus. There are no strategies specific to clausal focus.

Importantly, no focus marking strategy discussed in this section *is required* by any type of focus. For example, correctional focus can be marked with the constituent in a postverbal position, but both the preverbal position and in-situ marking are available marking strategies as well. The focus marking strategies are characterized by increasing “markedness”, in the sense that more marked strategies encode subsets of the types of foci which the less marked strategies encode. Constituent focus is mostly marked in situ; the immediately preverbal and postverbal positions have pragmatic restrictions as to when non-object constituents may appear there; and finally, pseudoclefts as the most complex type of focus marking are employed only for identificational focus. The distribution conforms to the markedness scale of structural complexity of focus marking by Skopeteas and Fanselow [2010], which predicts that more marked types of focus require more marked strategies. In the broader typological picture, the data shows that Meadow Mari behaves like a rather typical SOV language, since it allows preverbal focus. But at the same time, it is not a strict-position focus marking language, as an alternative in-situ marking strategy exists, cf. [Büring 2009].

4. Topics

Information is not added to the discourse in an unstructured way. Rather, speakers attempt to organize this information and give hints about how it is to be classified. Reinhart [1981] argues that the discourse context is not an unordered set of propositions, rather it is structured into subsets of propositions, which are stored under defining entries, for example, propositions about individuals — this is the notion of topic. In example (25a), the new information provided by the sentence is stored as information about Aristotle Onassis, and in example (25b) it is stored as information about Jacqueline Kennedy, the respective topic. Topical constituents are marked by boldfacing:

- (25) [Krifka 2008: 265]

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

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. In relation to the discourse context, topics can serve different purposes. For example, some topics continue reference to previously introduced, given referents, while others are newly

introduced into the discourse or mark the shift to another topic. Givón [1983: 9] characterizes topics like the latter as “newly-introduced, newly-changed or newly-returned” topics, and like the former ones as “continuous topics”. Similar classifications have been made by many scholars, see, e.g., [Erteshik-Shir 2007; Frascarelli 2018; Frascarelli, Hinterhölzl 2007; van der Wal 2021], 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. After the discussion about topics in this section, I will argue in Section 5 that the Meadow Mari suffix *-ŽE* can mark a topic shift, but not a topic continuation.

Another notion which is commonly distinguished is that of contrastive topic. I follow the definition by Büring [2003; 2016], which states that contrastive topics trigger a conventional implicature that there is an identifiable alternative question including the referent of the contrastive topic expression.¹² Typical contexts that evoke sequences with contrastive topic and focus are multiple *wh*-questions, single *wh*-questions containing plurals, or implicational topics (26). This third type, like presentational foci, indicates additional questions such as *Where was the cook?* Contrastive topics are boldfaced in the examples.

(26) [Büring 2016: 68, 70]

- a. (*Which guest brought what?* —) **Fred** brought the BEANS.
- b. (*Where do your siblings live?* —) My **sister** lives in STOCKHOLM.
- c. (*Where was the gardener at the time of the murder?* —) The **gardener** was in the HOUSE.

In Meadow Mari, aboutness topics and contrastive topics are located in the left periphery of the clause, while expressions referring to given individuals can appear postverbally. There does not seem to be a designated structural position for topics, however, as they can be preceded or followed by various adverbs, as in example (29) below. In addition to topics which are part of the clause, Meadow Mari has left-dislocated aboutness topics as well. In the former case, the topics necessarily have a corresponding gap in the clause (e.g. the position for objects in the unmarked order), other than left dislocated topics. For aboutness topics, two main diagnostics can be employed: Most importantly, non-subjects as topics in a topic progression are preposed and located in the left periphery, e.g. objects in an OSV order (if preverbal subject focus is not the reason for reordering). In this case, they can still be preceded by adverbs. The second diagnostic, the classic aboutness-test, see, e.g., [Erteshik-Shir 2007: 19–20], reveals that aboutness topics are preferred in a position left of the verb and dispreferred in a postverbal position.¹³

In a clause uttered in a neutral context, subjects occupy a left position in the clause, and the subject functions as sentence topic as well. But if a non-subject becomes the topic of a clause in the sense that the clause is “about” the non-subject, the constituent can be placed in the sentence-initial position to express the progression, that is, the shift to this new topic. These topics may receive a contrastive interpretation, but not obligatorily. In example (27), the first sentence tells about what happened after a charity concert and introduces ‘200 000 rubles’. The following sentence describes what happened to that money, and the accusative-marked

¹² Note that Büring [2003; 2016] considers contrastive topics as independent of topics, in contrast to, e.g., Krifka [2008], who considers them a combination of topic and focus. Regardless of that fact, I discuss contrastive topics in the section about topics.

¹³ Based on data from Bradley and Hirvonen [2022], I previously suggested another diagnostic for the topichood of left-peripheral constituents—that placing a constituent in the clause-initial position increases its acceptability as antecedent for a null subject in the following clause [Hirvonen 2023: 64–66], based on the hypothesis that in consistent null-subject languages like Meadow Mari, null subjects are resolved with respect to the aboutness topic, see, e.g., [Holmberg 2010; Frascarelli 2018]. While there is indeed a small effect in the data, and topichood thus could play a role for null subject resolution, null subjects and preposed aboutness topics regularly occur together in the same clause. Therefore, null subject resolution and topichood are likely two different phenomena. I will not further discuss the availability of this diagnostic in this article.

phrase ‘this money’ occupies the sentence-initial position before the subject of the clause, in an OSXV order. Similarly, in example (28), belletristic books are continued as topic ‘such books’ in an OSXV order:

- (27) *Tunam, koncert ertarə-me dene kok šüdü tüžem tenge*
 then concert carry.out-PTCP.PASS with two hundred thousand ruble
pogən-en. Tide šijvundə-m murəzo onkologij čer dene
 be.gathered-PST2.3SG this money-ACC singer oncology illness with
čerlanə-še joča-vlak-əm eml-aš kusar-en.
 be.sick-PTCP.ACT child-PL-ACC cure-INF transfer-PST2.3SG
 ‘In the course of the concert 200 000 rubles were gathered. **This money**, the singer donated to the treatment of children with cancer.’ [Arkhangelskiy 2019, Journal *Kidšer* (26.03.2016)]

- (28) *Səlnəmut-an kniga-m avtor-vlak ške küšeš aļe sponsor polšəmo dene*
 literature-ADJ book-ACC author-PL own at.expense or sponsor help with
vele lukt-ət. A tagaj kniga-m pisatel'-vlak en ondak užal-aš
 only publish-3PL but such.kind.of book-ACC writer-PL SUP before sell-INF
tərš-at <...>.
 try-3PL
 ‘Belletristic books, the authors only publish on their own expense or with the support of sponsors. But **such books**, the writers first of all try to sell <...>.’ [Arkhangelskiy 2019, Newspaper *Marij El* (27.05.2005)]

The classic aboutness-test can also be applied. Example (29) shows that subjects as aboutness topics are preferably placed left in the clause, and the acceptability of a sentence degrades if they appear further right. Such subjects were not accepted in a postverbal position by my consultant. Only expressions referring to given individuals can be located there. This example also shows that the aboutness topic can be preceded by adverbs, cf. [Hirvonen 2023: 67–68].

- (29) [“Have you heard anything about Serge?”]
*(Serge), pialeš, (Serge) u paša-m (ʔSerge) mu-ən, (*Serge).*
 Serge fortunately Serge new work-ACC Serge find-PST2.3SG Serge
 ‘Fortunately, **Serge** has found a new job.’ (elicited)

Mere givenness does not explain the left-peripheral placement of a given constituent alone, as given constituents can also be located elsewhere in the clause. This can be observed in the first clause of example (28) above, where the topic precedes the given subject (*pisatel'-vlak* ‘the writers’), and in the following newspaper example (30) about the capital of Mari El, Yoshkar-Ola. Here, parts of the city are introduced in the first sentence and referred to by a phrase headed by the proximal demonstrative *tə* ‘this’ in the next sentence. The phrase is placed after the subject in the clause:

- (30) *Pətartəš žap-əšte Joškar-Ola-šte u mikrorajon-vlak pišə-n kušk-ət:*
 last time-INE Yoshkar-Ola-INE new microdistrict-PL quick-INS grow-3PL
Festival'naj, Mirnəj, Molodežnəj, Sportivnəj. Jeŋ-vlak tə kundem-əšte
 F. M. M. S. person-PL this area-INE
pačer-əm nal-ət, posnak samərak-vlak.
 apartment-ACC buy-3PL especially young-PL
 ‘In recent times, new microdistricts in Yoshkar-Ola are growing fast: Festival'nəj, Mirnəj, Molodežnəj, Sportivnəj. People are buying apartments in these areas, especially the young ones.’ [Arkhangelskiy 2019, Newspaper *Marij El* (20.03.2017)]

Contrastive topics are located in the left periphery of the clause as well. This is true for subjects (31) and objects as contrastive topics (32):

- (31) [“Where do your siblings live?”]

Iza-m *MOSKVA-ŠTE*, *a* *šūžar-em* *JOŠKAR-OLA-ŠTE*.
 older.brother-PX1SG MOSCOW-INE but little.sister-PX1SG Joškar-Ola-INE

‘My older brother (lives) in MOSCOW, and my little sister (lives) in YOSHKAR-OLA.’ (elicited)

- (32) *A* *Irina* *kok* *tarelka-m* *vele* *mušk-ən* *šuktə-š*. *Kumšo* *tarelka-žə-m*
 but Irina two plate-ACC only wash-CVB manage-PST1.3SG third plate-ŽE-ACC
tudo *mušk-ən* *əš* *šukto*.
 3SG wash-CVB NEG.PST.3SG manage.CNG

‘But Irina only managed to wash two plates. The third plate she did not manage to wash.’ (elicited)

One indication for a topical interpretation of left-peripheral non-subjects is the restricted availability of non-subject negative indefinite pronouns in this position. They cannot receive a topical interpretation because they do not refer to any individual. This can be observed in example (33) — an object negative indefinite pronoun cannot appear in a topical position even if the structure favors the clause-initial placement of the object in contrast to the topic in the preceding clause:¹⁴

- (33) *Araka* *manmašte*, *tudə-m* *iktaž-kö* *jü-ən* *mo?* — *Araka-m*
 vodka about 3SG-ACC INDEF-who drink-PST2.3SG Q vodka-ACC
iza-m *jü-ən*, *a* *{*ñimo-m}* *šoləm*
 older.brother-PX1SG drink-PST2.3SG and nothing-ACC younger.brother.PX1SG
{ñimo-m} *jü-ən* *og-əl*.
 nothing-ACC drink-CVB NEG.3SG-be.CNG

‘What about the vodka, did anybody drink it? — The vodka my older brother drank, but my younger brother drank nothing.’ (elicited)

Apart from topics in the left periphery of the clause, Meadow Mari allows for left-dislocated topics outside of the core clause as well. Example (34) shows such a topic (which is additionally marked with -ŽE, see Section 5). It is not part of the clause as the topic is in the nominative case like a subject, but the finite verb does not agree with it. Instead, the verb shows 3PL morphology agreeing with a 3PL null subject, which is typical for generic statements, cf. [Bradley, Hirvonen 2022]. These topics can optionally be introduced by the phrase *manmašte* ‘as regards, concerning’ (35). Optionally, a resumptive pronoun within the core clause (e.g. *tudo* ‘it, this’) can refer back to this topic.

- (34) *Marij* *televidenij-žə* *tidə-m=ak* *vele* *ončəkt-at* *kuze* *ubiratl-at*.
 Mari television-ŽE this-ACC=EMPH only show-3PL how harvest-3PL

‘Mari television, they only show such stuff, how they harvest crops.’ [Serdobolskaya, Toldova 2012: 731]

- (35) [“Before Miľa returns from work, she’s really busying herself ⟨...⟩.”]

Paša *manmašte*, *Sergej* *ñimənar-at* *jara* *ok* *šinče* ⟨...⟩.
 work about Sergej no.amount.of=ADD idle NEG.3SG sit.CNG

‘Speaking about work, Sergej does not sit around idly either ⟨...⟩.’ [Arkhangelskiy 2019, Newspaper *Marij El* (01.09.2007)]

After the discussion about positions for topics in this section, and foci in Section 3, a short remark about the other positions in the Meadow Mari clause is in order. The aboutness topics and contrastive topics discussed so far are located left in the clause. Section 3 on focussed

¹⁴ Note, however, that non-subject negative indefinite pronouns can be found in the initial position in corpus data, but the exact contexts have yet to be investigated. As contrastive topic in example (33), the negative indefinite was not accepted in any case.

constituents showed that if constituent focus is realized by positioning, then the constituent is placed in the immediately preverbal or postverbal position. Besides the left periphery as a place for topics, and the immediately preverbal and immediately postverbal position for focussed constituents, the other parts of the clause do not appear to receive any special interpretation. Pregla [2024: 92–95] noted this for constituents preceding focus. Regarding the postverbal field, Hirvonen [2023: 73–84] showed that it is a heterogeneous place in the Meadow Mari clause, and that constituents can end up there for various reasons, including: (1) as afterthoughts and completions of a clause, containing expressions referring to given individuals;¹⁵ (2) constituents can epiphenomenally appear to the right of the verb when the verb itself occupies a left-peripheral position for verbal focussing purposes; (3) prosodically heavy constituents can be extraposed to a right-peripheral position; (4) genre-specific reasons. Furthermore, presentational and corrective constituent focus can appear postverbally (see Section 3.2.2). Therefore, constituents can appear in the postverbal field for various unrelated reasons, and do not necessarily receive a certain interpretation. Reason (2) was described already in Section 3.2.1, and for (1), (3), and (4) I refer the reader to the discussion in Hirvonen [2023: 73–84].

Summarizing, the place for topics in the Meadow Mari clause is the left periphery. Aboutness topics and contrastive topics are located in that part of the clause, and this is most clearly visible when non-subjects appear as topics. Furthermore, Mari employs left-dislocated topics for topic shift. The postverbal field can be filled for various unrelated reasons, and the remaining positions in the clause do not appear to serve any specific functions related to information structure marking.

5. Functions of the morpheme -ŽE

Besides syntactic means of marking information structure, Meadow Mari also has at its disposal some morphemes interacting with information structure, including focus particles (*vele*, *gəna* ‘only’), the additive clitic (=at ‘also, even’), and the morpheme -ŽE, which will be described in this section in more detail. In contemporary Meadow Mari, this morpheme is used in two main distinct discursive functions: it can mark the selection of an individual from a superset, including contrastive topics (36) and a topic shift, and it marks uncontroversial information (37), cf. the category of “enimitive”, introduced by Panov [2020]. This section is mainly concerned with the former usage. The relevant parts of the sentence are highlighted by boldfacing in this section.

- (36) *No pūrtūs-še peš saj, a uslovij-že saj ogəl il-aš.*
 well nature-ŽE very good but condition-ŽE good NEG live-INF
 ‘Well, **nature** is very pretty (there), but **conditions** aren’t good to live.’ [Serdobolskaya, Toldova 2012: 728]

- (37) *Tače təlāt nimogaj 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-ŽE chicken-PX1PL-ACC eat-PST2.3SG
 ‘I cannot bring you any eggs today. The **fox** has eaten our chicken, **you see**.’ (elicited)

-ŽE is used as a shorthand for the suffix’s various allomorphs that appear as the result of processes like voicing assimilation (e.g. -že vs. -še), avoidance of consonant clusters (e.g. -əže vs. -že) or vowel

¹⁵ In Hirvonen [2023] I treated familiar topics separately from afterthoughts and completions. However, it has not been systematically investigated yet whether there are any prosodic or syntactic differences that justify treating familiar topics in Meadow Mari as a separate descriptive category, instead of treating them simply as given elements (thanks to Andreas Pregla for discussion).

harmony (e.g. *-že, -žo, -žō*) [Alhoniemi 1993: 72; Riese et al. 2022: 46–48]. A homophonous suffix also serves as the possessive suffix of the third person singular, but due to differences not only in meaning but also in distributional properties (and for ease of exposition), it is reasonable to treat these functions separately. This is also reflected in the glossing. *-ŽE* as discussed here has the following distributional properties differing from the 3SG possessive suffix usage, cf. also [Georgieva 2022; 2024]:

- 1) Possessive suffixes agree with the possessor in person and number, but *-ŽE* is invariant. This is visible most clearly in cases of ellipsis where the antecedent involves a non-3SG possessor, but the remnant bears the suffix *-ŽE* and thus does not agree with the possessor. The remnant can still receive a non-3SG interpretation:
- (38) [“You have one heavy and one light bag.”]
Nele sumka-t ūstembalne, a kuštalgo / kuštalgo-žo polka-što.
 heavy bag-PX2SG on.table but light light-ŽE shelf-INE
 ‘Your heavy bag is on the table, and **your light one** is on the shelf.’ [Georgieva 2022: 4]
- 2) The suffix can be stacked onto other possessive suffixes, in which case *-ŽE* follows the possessive suffix and does not agree with any constituent, cf. [Georgieva 2022; Tuzharov 1987: 67; Simonenko 2014]:
- (39) [“What food do your pets like to eat?”]
Pij-em(-že) pəzə-m kočk-aš jōrat-a, pəras-em(-že) kol-əm kočk-aš jōrat-a.
 dog-PX1SG-ŽE meat-ACC eat-INF love-3SG cat-PX1SG-ŽE fish-ACC eat-INF love-3SG
 ‘My dog likes to eat meat, (and) my cat likes to eat fish.’ (elicited)
- 3) In the enimitive function it can mark other constituents as well. In example (40), *-ŽE* attaches to a whole PP:
- (40) [*Araka deč posna*]-žē veselitl-aš ok lij mo?
 vodka from separate-ŽE have.fun-INF NEG.3SG be.CNG Q
 ‘Can’t you have a good time **without alcohol**?’ [Arkhangelskiy 2019, VKontake]

The suffix has been described before as a marker of definiteness, e.g., in [Alhoniemi 1993: 75–76], identifiability [Nikolaeva 2003], specificity [Georgieva 2022; Simonenko 2017], contrast [Riese et al. 2022: §9.1.5; Tuzharov 1987: 67–68; Nikolaeva 2003; Simonenko 2014], contrastive topics [Georgieva 2024; Hirvonen 2023; Yurayong 2020: 115], topic shift [Hirvonen 2023], as a marker establishing reference to previously mentioned individuals in the discourse [Riese et al. 2022: §9.1.5; Saarinen 2022: 445], or some other kind of emphasis [Riese et al. 2022: §9.1.5; Tuzharov 1987: 67]. Georgieva [2022: 6] notes that some Meadow Mari speakers also allow usage of the suffix in anaphoric contexts. For closely related Hill Mari, Khomchenkova [2022; 2023] has argued that the suffix marks topics and has a contrastive use with topics, but less with focus.

Here I show that the main syntactic and interpretational properties of discursive, non-enimitive *-ŽE* are as follows: (1) it attaches to the very right of any nominal constituent; (2) the marked constituents are mostly left-peripheral; (3a) it expresses the selection from a superset and a contrast, but (3b) its usage is optional if the context already provides a superset and a contrast; (4a) it can mark a topic shift; and (4b) marking a shift with *-ŽE* appears to have an adversative function, or expresses unexpectedness. In contemporary usage, the suffix, furthermore, seems to mirror usage of the Russian discourse particles *že* (же) (including the enimitive) and *-to* (-то), cf. [Hirvonen 2023: 109–115], but these functions will not be discussed here.

Property (1), the attachment to the very right of a nominal constituent, is visible most clearly in cases where the suffix follows other possessive suffixes, e.g. in (39). In most cases where the suffix appears on another phrase, it receives an enimitive interpretation, as in, e.g., (40). Property (2), the left-peripheral placement of the marked constituent, is indicative of the fact that *-ŽE* often marks constituents which receive a topical interpretation, since these constituents are

placed in the left periphery (see Section 4). This can be seen in most examples in this section as well. Of special interest are sentences with topical objects in sentence-initial position preceding subjects, cf. example (32) repeated below, as they show this very clearly. This example also illustrates that both of the contrasted expressions need not be marked at the same time, cf. [Hirvonen 2023: 112–113; Georgieva 2024: 13]:

- (41) *A Irina kok tarelka-m vele mušk-an šuktā-š. Kumšo tarelka-žə-m*
 but Irina two plate-ACC only wash-CVB manage-PST1.3SG third plate-ŽE-ACC
tudo mušk-an əš šukto.
 3SG wash-CVB NEG.PST.3SG manage.CNG
 ‘But Irina only managed to wash two plates. **The third plate** she did not manage to wash.’
 (= (32))

Most commonly, the suffix marks a contrast between referents of a superset which is contextually available (property 3a), cf. [Simonenko 2014; 2017; Georgieva 2022]. But importantly, neither superset nor contrasted referent have to be explicitly introduced in the context. Property (3a) can be shown in contexts which typically elicit sequences of contrastive topic and focus, namely if a superset is introduced in the *wh*-question setting the context, and the answer provides a pair-list reading, e.g. (39). When the context provides the superset, its usage is optional, however (property 3b) — in example (39) my consultant did not express a preference for using or omitting the suffix.

Summarizing so far, the suffix marks that, from an available superset, one alternative is chosen and contrasted with another. This happens mostly with topics, and marking is optional if the context already provides the superset. The pragmatic effect of actually using the morpheme can be illustrated best with cases in which the context neither provides nor excludes a superset. In such a context, the suffix implies the presence of a contrast, although the contrast does not have to be known to the speaker (42). This is a typical property of contrastive topic markers [van der Wal 2021: 36]:

- (42) *Maša kočkəš-əm jamdəl-en mo? — Šür-žə-m tudo šolt-en.*
 Maša food-ACC prepare-PST2.3SG Q soup-ŽE-ACC 3SG cook-PST2.3SG
Ves kočkəš-əm tudo jamdəl-en aʔe uke — o-m pale.
 other food-ACC 3SG prepare-PST2.3SG or NEGEXIST NEG-1SG 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 [by some one else] as well.)’ (elicited)

The suffix also marks the shift to a topic which is newly introduced, newly changed, or newly returned to (property 4a). Here, a superset could be understood as the set of discourse participants or possible topics, of which then one is chosen as an aboutness topic for the clause (cf. [Khomchenkova 2022; 2023] on Hill Mari). In the following dialogue example, speaker A introduces *pamaš* ‘springs’ into the discourse in their turn. Speaker B shifts the aboutness topic in the sentence to this previously introduced constituent. But then, speaker A again rapidly shifts the topic to *snege* ‘wild strawberries’. In both cases, the topic is marked with -ŽE:

- (43) A. *Pamaš šuko ulo tušto, srazu pamaš gəč vüd-əm numal-ət.*
 spring many exist there at.once spring from water-ACC carry-3PL
 ‘There’s lots of springs, people just go right away to get some water from the springs.’
 B. *Memnan pamaš-əže nimat kod-ən og-əl <...>.*
 1PL.GEN spring-ŽE nothing stay-CVB NEG.3SG-be.CNG
 ‘At our place, **springs**, there is nothing left <...>.’
 A. *Snegə-že vet čot šuko ulo, šoldəra möre gaj.*
 wild.strawberry-ŽE see very much EXIST large garden.strawberry like
 ‘(And) **wild strawberries**, you know, there’s lots of, like large strawberries from the garden.’ [Serdobolskaya, Toldova 2012: 729]

Again, its usage appears to be optional in this function, as most topic shifts are not marked with -ŽE. But if the suffix is used, it may mark adversativity or unexpectedness in topic shift contexts (property 4b). In the beginning of the narration which the following example is taken from, a shed (*saraj* ~ *levaš* ‘shed, barn’) 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, which, in turn, stresses the unexpectedness of the events:

- (44) [“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 vūd saraj-əm mušk-ən
but once beat.down-CVB flow-PTCP.PASS spring water shed-ACC wash-CVB

nanğaj-en. Tiddeč vara levaš-əm ves vere čon-en-ət.
take.away-PST2.3SG from.this then shed-ACC other to.place build-PST2-3PL

Revolucij žap-əšte kerməč jamdələmaš čarn-en-ət. Saraj-že šuko žap
revolution time-INE brick production stop-PST2.3PL shed-ŽE much time

eše šinč-en.
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. **The shed, however**, was still standing for a long time.’ [Serdobolskaya, Tol-dova 2012: 724]

Usage of the suffix is not felicitous if the context does not allow the construal of a superset, and thus violates property (3a), cf. [Georgieva 2022: 5; Hirvonen 2023: 101–106]. Consider examples (45)–(46). Above I argued that -ŽE can mark a topic shift, and Section 4 discussed that these topics are typically placed in the left periphery. Conversely, expressions referring to known individuals, including topics, can be encountered in the postverbal position. In example (45), -ŽE marks a constituent in the postverbal position and the context does not provide or even allow the construal of a superset. My consultant did not accept marking this kind of expression with -ŽE. However, if the context provides a superset and/or a contrast or allows its construal superset and/or a contrast (e.g. a cat and a dog), such marking is felicitous (46):

- (45) [“What about your lost cat — did you look for it?”]

*Da, i mu-ən lukt-ən-am, pəras-em-əm / *pəras-em-žə-m.*
yes and find-CVB lead.out-PST2-1SG cat-PX1SG-ACC cat-PX1SG-ŽE-ACC

‘Yes, and I did find (him), **my cat**.’ (elicited)

- (46) [“What about your lost cat and dog — did you look for them?”]

Da, i mu-ən lukt-ən-am, pəras-em-əm / pəras-em-žə-m.
yes and find-CVB lead.out-PST2-1SG cat-PX1SG-ACC cat-PX1SG-ŽE-ACC

A pij-em-əm eše mu-ən om-əl.
but dog-PX1SG-ACC yet find-CVB NEG.1SG-be

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

Summarizing, the morpheme -ŽE has two main functions: It marks uncontroversial information, serving as an “enimitive” marker, cf. [Panov 2020], or selects an individual from a superset while implying a contrast to the other members of the set. In the latter function, the suffix is commonly used with topical constituents, including contrastive topics and a topic shift, or in contexts of unexpectedness and adversativity. Importantly, the suffix is optional in this function in the sense that the semantic well-formedness of the sentence does not depend on the morpheme.

6. Summary

This paper provided an overview of the morphosyntactic strategies of marking topic, focus, and contrastive topic in Meadow Mari. In this section the most important findings are summarized. Constituent focus can variably be marked by prosodic in-situ marking, immediately preverbal or postverbal positioning, and marginally, a pseudocleft construction. Importantly, no type of focus *requires* a certain strategy, and any type of focus can be marked prosodically, but the marking strategies increase in “markedness” — postverbal focus, for example, is available only in a subset of in-situ focus contexts. Aboutness topics and contrastive topics are placed in the left periphery of the clause. Furthermore, left-dislocated topics can be employed for a topic shift. Finally, the morpheme -ŽE marks the selection of an individual from a superset while implying a contrast to the other members. The morpheme is commonly used with contrastive topics and for topic shift.

As regards focus, Meadow Mari behaves like a rather typical Eurasian SOV language, since it allows preverbal focus. But at the same time, it is not a strict-position focus marking language, as an alternative in-situ marking strategy exists, cf. [Büring 2009]. Of further special interest in a comparative Uralic context is the suffix -ŽE, since a number of other Uralic (and Altaic) languages also employ possessive suffixes in discursive functions, see, e.g., [Nikolaeva 2003; Simonenko 2014; Gerland 2014; Yurayong 2020: 115–119], among many others. This overview of information structure in Meadow Mari contributes to the growing comparative body of work on information structure marking in the Uralic languages, see [Surányi et al. (to appear)], and lays the foundation for more detailed studies on topics in Meadow Mari information structure.

ABBREVIATIONS

1, 2, 3 — the 1st, 2nd, 3rd person
 ACC — accusative case
 ADD — additive clitic
 ADJ — adjectivizing suffix
 CMPR — comparative
 CNG — connegative verb form
 CVB — converb
 EMPH — emphatic clitic
 EXIST — existential copula
 ILL — illative case
 INDEF — indefinite prefix
 INE — inessive case
 INF — infinitive
 INS — instrumental

NEG — negation auxiliary or constituent negation
 NEGEXIST — negative existential predicate
 PL — plural
 PST1 — first past tense
 PST2 — second past tense
 PTCP.ACT — agent relativizing participle
 PTCP.PASS — passive participle
 PTCL — particle
 PX — possessive suffix
 Q — polar question marker
 SG — singular
 SUP — superlative
 ŽE — Mari discourse marker -ŽE

REFERENCES

- Alhoniemi 1993 — Alhoniemi A. *Grammatik des Tscheremisschen (Mari)*. Hamburg: Buske, 1993.
 Arkhangelskiy 2019 — Arkhangelskiy T. Corpora of social media in minority Uralic languages. *Proc. of the 5th Workshop on Computational Linguistics for Uralic Languages, Tartu, Estonia, January 7–8, 2019*.
 Pirinen T. A., Kaalep H.-J., Tyers F. M. (eds.). Association for Computational Linguistics, 2019, 125–140.
 Bradley, Hirvonen 2022 — Bradley J., Hirvonen J. Null subjects in Mari. *Null subjects in Slavic and Finno-Ugric: Licensing, structure and typology*. Dalmi G., Tsedryk E., Ceglowski P. (eds.). Boston; Berlin: De Gruyter, 2022, 281–306.
 Büring 2003 — Büring D. On D-trees, beans, and B-accents. *Linguistics and Philosophy*, 2003, 26(5): 511–545.

- Büring 2009 — Büring D. Towards a typology of focus realization. *Information structure*. Zimmermann M., Féry C. (eds.). Oxford: Oxford Univ. Press, 2009, 177–205.
- Büring 2016 — Büring D. (Contrastive) topic. *Oxford handbook of information structure*. Féry C., Ishihara Sh. (eds.). Oxford: Oxford Univ. Press, 2016, 64–85.
- Burukina et al. 2024 — Burukina I., Borise L., den Dikken M. A ‘big DP’ analysis of Russian copular constructions with *ëto*. *Glossa*, 2024, 9(1): 1–37.
- Chkhaidze 1941 — Chkhaidze M. P. *Sintaksis lugovo-vostochnogo mariiskogo yazyka* [Syntax of Meadow Mari]. Moscow: Uchpedgiz, 1941.
- Dezső 1978 — Dezső L. Towards a typology of theme and rheme: SOV languages. *Wortstellung und Bedeutung*. Vol. 1. Conte M.-E., Giacalone Ramat A., Ramat P. (eds.). Tübingen: Niemeyer, 1978, 3–11.
- É. Kiss 1998 — É. Kiss K. Identificational focus versus information focus. *Language*, 1998, 74: 245–273.
- Erteshik-Shir 2007 — Erteshik-Shir N. *Information structure: The syntax-discourse interface*. Oxford: Oxford Univ. Press, 2007.
- Frascarelli 2018 — Frascarelli M. The interpretation of *pro* in consistent and partial null-subject languages: A comparative interface analysis. *Null subjects in Generative Grammar*. Cognola F., Casalicchio J. (eds.). Oxford: Oxford Univ. Press, 2018, 211–239.
- Frascarelli, Hinterhölzl 2007 — Frascarelli M., Hinterhölzl R. Types of topics in German and Italian. *On information structure, meaning and form: Generalizations across languages*. Schwabe K., Winkler S. (eds.). Amsterdam; Philadelphia: John Benjamins, 2007, 87–116.
- Georgieva 2022 — Georgieva E. Fake possessives, partitives and definites sharing the same D. *Talk at the 45th GLOW Colloquium, Queen Mary Univ. of London, April 28, 2022*.
- Georgieva 2024 — Georgieva E. On sluicing in Meadow Mari and Udmurt. *Talk at Jezik & Linguistics Colloquium, Univ. of Nova Gorica, May 30, 2024*.
- Georgieva et al. 2021 — Georgieva E., Salzmann M., Weisser Ph. Negative verb clusters in Mari and Udmurt and why they require postsyntactic top-down word-formation. *Natural Language and Linguistic Theory*, 2021, 39: 457–503.
- Gerland 2014 — Gerland D. Definitely not possessed? Possessives with non-possessive function. *Frames and concept types: Applications in language and philosophy*. Gamerschlag Th., Gerland D., Osswald R., Petersen W. (eds.). Dordrecht: Springer, 2014, 269–292.
- Givón 1983 — Givón T. Topic continuity in discourse: An introduction. *Topic continuity in discourse: A quantitative cross-language study*. Givón T. (ed.). Amsterdam: John Benjamins, 1983, 1–41.
- Herring 1990 — Herring S. C. Information structure as a consequence of word order type. *Annual Meeting of the Berkeley Linguistics Society*, 1990, 16(1): 163–174.
- Hirvonen 2023 — Hirvonen J. *Focus and topic in Meadow Mari*. Master’s thesis. Vienna: Univ. of Vienna, 2023. <https://theses.univie.ac.at/detail/69244/>.
- Holmberg 2010 — Holmberg A. Null subject parameters. *Parametric variation: Null subjects in Minimalist theory*. Biberauer Th., Holmberg A., Roberts I., Sheehan M. (eds.). Cambridge: Cambridge Univ. Press, 2010, 88–124.
- Khomchenkova 2022 — Khomchenkova I. A. Diskursivnye funktsii posessivnogo pokazatelya tret’ego litsa edinstvennogo chisla v gornomariiskom yazyke [Discourse functions of the 3rd person singular possessive marker in Hill Mari]. *Ural-Altaic Studies*, 2022, 45(2): 148–168.
- Khomchenkova 2023 — Khomchenkova I. A. Diskursivnye funktsii posessivnogo pokazatelya 3-go litsa edinstvennogo chisla [Discourse functions of the 3rd person singular possessive marker]. *Elementy gornomariiskogo yazyka v tipologicheskoy osveshchenii*. Kashkin E. V. (ed.). Moscow: Buki Vedi, 2023, 614–630.
- Kim 1988 — Kim A. H.-O. Preverbal focusing and type XXIII languages. *Studies in syntactic typology*. Hammond M., Moravcsik E. A., Wirth J. (eds.). Amsterdam: John Benjamins, 1988, 149–171.
- Krifka 2008 — Krifka M. Basic notions of information structure. *Acta Linguistica Hungarica*, 2008, 55(3–4): 243–276.
- Nelson et al. (to appear) — Nelson D., Vedernikova E., Bradley J. The syntax of Meadow Mari: Some issues around TAM and phrase structure. *Uralic syntax*. Tamm A., Vainikka A. (eds.). Cambridge: Cambridge Univ. Press. To appear.
- Nikolaeva 2003 — Nikolaeva I. Possessive affixes in the pragmatic structuring of the utterance: Evidence from Uralic. *International Symposium on Deictic Systems and Quantification in Languages Spoken in Europe and North and Central Asia: Coll. of papers*. Suihkonen P., Comrie B. (eds.). Izhevsk; Leipzig: Udmurt State Univ.; Max Planck Institute for Evolutionary Anthology, 2003, 130–145.

- Panov 2020 — Panov V. The marking of uncontroversial information in Europe: Presenting the enimitive. *Acta Linguistica Hafniensia*, 2020, 52(1): 1–44.
- Pengitov 1961 — Pengitov N. T. Poryadok slov v prostom rasprostranennom predlozhenii [Word order in a simple expanded sentence in Mari]. *Trudy MarNIIYaLI*, 1961, 15: 113–124.
- Pregla 2024 — Pregla A. *Word order variability in OV languages: A study on scrambling, verb movement, and postverbal elements with a focus on Uralic languages*. PhD thesis. Potsdam: Univ. of Potsdam, 2024. <https://doi.org/10.25932/publishup-64363>
- Reinhart 1981 — Reinhart T. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica*, 1981, 27(1): 53–94.
- Riese et al. 2022 — Riese T., Bradley J., Yefremova T. *Mari (марий йылым): An essential grammar for international learners*. Vienna: University of Vienna, 2022. grammar.mari-language.com.
- Saarinén 2022 — Saarinen S. Mari. *The Oxford guide to the Uralic Languages*. Bakró-Nagy M., Laakso J., Skribnik E. (eds.). Oxford: Oxford Univ. Press, 2022, 432–470.
- Serdobolskaya, Toldova 2012 — Serdobolskaya N. V., Toldova S. Yu. Teksty na lugovom mariiskom yazyke (starotor’yal’skii govor) [Texts in Meadow Mari (Starotoryal dialect)]. *Finno-ugorskie yazyki: fragmenty grammaticheskogo opisanija. Formal’nyi i funktsional’nyi podkhody*. Kuznetsova A. I. (ed.). Moscow: Rukopisnye pamyatniki Drevnei Rusi, 2012, 709–745.
- Simonenko 2014 — Simonenko A. Microvariation in Finno-Ugric possessive markers. *NELS 43: Proc. of the 43th annual meeting of the North East Linguistic Society*. Vol. 2. Huang H.-L., Poole E., Rysling A. (eds.). Amherst: GLSA, 2014, 127–140.
- Simonenko 2017 — Simonenko A. Towards a semantic typology of specific determiners. *Proc. of the 21st Amsterdam Colloquium*. Cremers A., van Gessel Th., Roelofsen F. Amsterdam: Institute for Logic, Language and Computation, 2017, 425–434.
- Skopeteas, Fanselow 2010 — Skopeteas S., Fanselow G. Focus types and argument asymmetries. *Comparative and contrastive studies of information structure*. Breul C., Göbbel E. (eds.). Amsterdam; Philadelphia: John Benjamins, 2010, 169–197.
- Surányi et al. (to appear) — Surányi B., Brattico P., Asztalos E., Sahkai H. Information structure in Finno-Ugric languages. *Uralic syntax*. Tamm A., Vainikka A. (eds.). Cambridge: Cambridge Univ. Press. To appear.
- Tuzharov 1987 — Tuzharov G. M. *Grammaticheskie kategorii imeni sushchestvitel’nogo v mariiskom yazyke* [Grammatical categories of the noun in Mari]. Yoshkar-Ola: Mari Book Publ., 1987.
- van der Wal 2016 — van der Wal J. Diagnosing focus. *Studies in Language*, 2016, 40(2): 259–301.
- van der Wal 2021 — van der Wal J. *The BaSIS basics of information structure*. Aug. 2021. Manuscript.
- Vilkuna 1998 — Vilkuna M. Constituent order in European Uralic. *Constituent order in the Languages of Europe*. Siewierska A. (ed.). Berlin: de Gruyter, 1998, 173–233.
- Vilkuna 2022 — Vilkuna M. Word order. *The Oxford guide to the Uralic languages*. Bakró-Nagy M., Laakso J., Skribnik E. (eds.). Oxford: Oxford Univ. Press, 2022, 950–960.
- Volkova et al. 2024 — Volkova A., Zakirova A., Voronov M., Dolgodvorova M., Klyucheva Z., Kokoreva S., Makarchuk I., Khomchenkova I., Arkhangel’skiy T., Sokur E. *Spoken corpus of Meadow Mari (as spoken in the village of Staryj Torjal, Novyj Torjal district, Mari El Republic, Russia)*. Moscow: HSE Univ., 2024. Electronic resource: http://lingconlab.ru/spoken_meadow_mari/ (accessed on 01.07.2024).
- Yurayong 2020 — Yurayong Ch. *Postposed demonstratives in Finnic and North Russian Dialects*. Doctoral diss. Helsinki: Univ. of Helsinki, 2020.
- Yurayong et al. 2024 — Yurayong Ch., Niva H., Miestamo M., Kittilä S., Sandman E. Miscellanea typologica uralica ex oriente et occidente. *Itämeren kieliapajilta Volgan verkoille: Pühendusteos Riho Grünthalile 22. mail 2024*. Björklöf S., Jantunen S., Junttila S., Kuokkala J., Laitinen S., Pasanen A. (eds.). Helsinki: Suomalais-Ugrilainen Seura, 2024, 729–760.