JEREMY BRADLEY
VIENNA

Mari function words on the border of syntax and morphology

This contribution examines the prospective morphological status of several elements in Mari that are conventionally classified as function words (postpositions, particles, conjunctions) but that have long been known to play essential roles in Mari grammar, complementing the (uncontroversial) case suffixes and enclitics. The paper asks the question if conventions of orthography are disincentivizing the classification of strongly morphologized elements as such. Different factors examined as metrics of morphologization include: the orthographic realization of word forms (esp. in preliterary texts and transcriptions of oral texts), restrictions on the placement of elements, possessive suffixation of elements, voice assimilation and vowel harmonic processes across orthographic word boundaries, prosodic factors.

1. Introduction

1.1. Mission statement

Thomas (Talmy) Givón's famous aphorism that "today's morphology is yesterday's syntax" (Givón 1971: 25) is plainly evident when considering the morphologization of postpostions in Uralic languages. Finnic offers ample data in this domain (see Grünthal 2003), where we can see the same basic building blocks acting as postpositions in some languages and suffixes in others, e.g., Finnish *päin* 'towards; from' (< *pää* 'head') ~ Veps elative *-späi*, with collectors of language data (generally cooperating with the speakers when

transcribing the texts) oftentimes clearly unsure if the element should be written as a separate element or not (ibid.: 117–118).

This paper looks at a category of function words (including, but not restricted to, postpositions) in Mari[1] undergoing similar processes and implies that their morphologization might be further along than conventional descriptions, and especially the norms of orthography, communicate. These elements, henceforth called *D&G words* (their categorization as *words* is tenuous, but as they tend to be orthographically realized as such, I will continue using this term), are frequently used postpositions, conjunctions, and particles that are written with an initial <д> /d/ (historically /ð/) or <г> /g/ (historically /ɣ/), such as *дене* /dene/ 'with', *да* /da/ 'and', *гына* /gə̂na/ 'only'. Their orthographic realization suggests a violation of Mari phonotactics: Mari does not allow for these sounds in the word-initial position (Ivanov 2000: 90). But, this is simply the product of an arbitrary convention of Mari orthography: the words in question cannot occur in sentence-initial position, obligatorily have a host word, and are subject to orthographically unmarked voice assimilation (/d/ ~ /t/, /g/ ~ /k/): *йолташем дене* /joltašem **d**ene/ 'with my friend', but *йолташет дене* /joltašet **t**ene/ 'with your friend' (cf. Riese et al. 2022: 31–32). Undoubtedly these words have been cliticized, where a clitic is understood as "a bound word-form – i.e., a word-form that is prosodically dependent on a host" (Haspelmath & Sims 2010: 322) , "a surface element part-way between a word and an affix in properties" (Dixon 2010: 221). Here an affix is understood as a form that is "added to the root or stem, the whole functioning as one grammatical word (and generally also as one phonological word [...])" (ibid.). More debatable is if these forms have proceeded further down the path of grammaticalization and can, at least in some cases, be considered de facto paradigmatic.

This raises the question if it is also the product of an arbitrary convention that these elements are considered words in the first place, rather than bound forms? This paper will address the classification of the assorted D&G words in Mari, addressing the following questions (the findings are summarized in Section 4):

- Do all function words starting in orthographic <д> or <г> fall into this category, or can some of them be found in sentence initial position without a host word? This will be examined based on recently published corpora of literary Mari (See Sections 2 & 4).

---

1. The content of this paper pertains to both literary standards of Mari (Meadow-Eastern & Hill) and all four Mari dialect groups (Meadow, Eastern, Hill, Northwestern). Data labelled simply as "Mari" is Meadow/Eastern; Hill Mari data is indicated as such.

- Does prosody offer any data on the status of these forms? (see Section 3.1)
- Numerous D&G words are borrowings (e.g., *∂a* /da/ 'and' < Chuvash, *гына* /gə̂na/ 'only' < Tatar, Moisio & Saarinen 2008: s.v. *δa*, *γə̂na*) – can morpheme-like parameters assigned to these words in Mari such as voice assimilation (see Section 3.2) and vowel-harmonic adaptation (see Section 3.3) already be assigned to these words in the source languages?
- Does a survey of the dialectal text collections of Mari containing transcripts from the late 19th and early 20th century give any indication of dialectal and short-scale diachronic differences are regards these parameters? (see Section 3.4)
- Do D&G words show idiosyncratic possessive suffixation when compared to relational nouns? (see Section 3.5)
- How do postpositional D&G words differ in morphosyntax from other elements frequently assigned to the category of postpositions – relational nouns such as *йымалне* /jə̂malne/ 'under' (< *йымал* /jə̂mal/ 'bottom' + non-productive locative marker -*не* /-ne/)? What are differences in syntax (what degrees of freedom do the respective words enjoy as regards their placement) and in person marking? (see Section 3.6)
- Some of the postpositions in question have been classified as complementary to the case system especially in Finnish sources (Alhoniemi 1985: 62). To what extent is this based on a "top-down view" (i.e., like Finnish, Mari has an illative and inessive case, but what is the functional counterpart to the Finnish elative case?), and to what extent is this classification suggested by the data itself? (see Section 3.7)

1.2. The rather syntactic morphology of Mari

As is typical for the (Uralic and Turkic) languages of the Volga-Kama Region, Mari possesses rich and highly concatenative morphology in which morpheme boundaries can be unambiguously drawn. The degrees of freedom afforded by Mari morphology are, however, atypical even in the Volga-Kama Region. Note different permissible arrangements of case suffixes (Cx), possessive suffixes (Px), and number suffixes (Nx) in examples taken from the Corpus of Literary Mari[2].

---

2. ‹http://corpus.mari-language.com›

Mari

(1) a.   йолташ-влаклан
          joltaš-em-βlak-lan
          friend-1SG-PL-DAT
          'to my friends' (Px-Nx-Cx) (Corpus of Literary Mari)

    b.   пире-влакетлан
          pire-βlak-et-lan
          wolf-PL-2SG-DAT
          'to your wolves' (Nx-Px-Cx) (Corpus of Literary Mari)

    c.   йоча-влакланже
          joča-βlak-lan-že
          child-PL-DAT-3SG
          'to his/her/their.SG children' (Nx-Cx-Px) (Corpus of Literary Mari)

This situation led Erkki Itkonen (1960), and especially Jorma Luutonen, to analyse the "relative analyticity of the Mari declensional forms" (Luutonen 1997: 41), and to liken forms traditionally classified as postpositional with forms traditionally classified as morphological (ibid.: 47–49). Even in orthographic norms, these trends can be observed in the hyphenated realization of the plural suffixes *-влак* and *-шамыч* (Riese & Bradley & Yefremova 2022: 54). Metrics that can be used to distinguish morphology from syntax will be discussed in Section 3, but it is notable that by many of these (voice assimilation, vowel harmony) the evidence for classifying the aforementioned plural suffixes as morphological is comparatively weak (cf. also Luutonen 1997: 49–50).

2.   An inventory of D&G words

Given the inadmissibility of word-initial /d/ and /g/ in Mari outside of D&G words and onomatopoeia, the <д> and <г> sections of Mari dictionaries (e.g., Galkin et al. 1990–2005; Riese & Bradley & Guseva 2014) are fairly sparsely populated and dominated by Russian borrowings. During the compilation of the Mari Web Project's Mari grammar for international learners (Riese & Bradley & Yefremova 2022), the words listed alphabetically below with initial <д> and <г>, falling into the established categories, were identified on the basis of the aforementioned dictionaries and the University of Turku's electronic Mari word list (Luutonen et al. 2007). Etymological data provided here is taken from the *Tscheremissisches Wörterbuch* (Moisio & Saarinen 2008).

It should be noted that some elements are polyfunctional and straddle the boundaries between word classes (thus word classes were provided here, for polyfunctional elements all admissible word classes), and that compound elements (e.g., *дег посна* /deč̌ posna/ 'without', literally 'from separate') were excluded.

- *гай* /gaj/ 'like, as; it seems, … might, like' (postposition, particle)
- *гала* /gala/ 'indeed, really, again; not only' (particle)
- *гана* /gana/ 'time(s)' (postposition)
- *годсек* /godsek/ 'since, beginning with' (postposition), < stem *год-* (not used independently)[3]
- *годым* /godə̂m/ 'during, at the time' (postposition), < stem *год-* (not used independently)
- *гоҕ(ын)*[4] /goč̌(ə̂n)/ 'over, across, through; during, throughout; using, via' (postposition)
- *гутлан* /gutlan/ 'about, around (a time)' (postposition), < *кут* 'length'
- *гутлаште* /gutlašte/ 'around, in the area; about, around, during, at (a time)' (postposition), < *кут* 'length'
- *гын* /gə̂n/ 'if; as for …, … for one (emphatic, interrogative, contrastive)' (conjunction, particle)
- *гына* /gə̂na/ 'only, just; barely; rather, fairly; (emphatic particle)' (particle), < Tatar
- *гынат* /gə̂nat/ 'even if; some, any; even' (conjunction, particle), < гын 'if' + clitic -*am* 'also, and'
- *гыҕ(ын)* /gə̂č̌(ə̂n)/ 'from, out of' (postposition)
- *да* /da/ 'and; yes; so, well, indeed; yes?, right?, isn't it?; if only; (emphatic particle)' (conjunction, particle), < Chuvash and Russian (two borrowings with fuzzy division in modern language)
- *даже* /daže/ 'even' (conjunction), < Russian
- *даҥыт* /daŋə̂t/ 'to, up to' (postposition), < *таҥ* 'friend; together' < Chuvash
- *дек(е)* /dek(e)/ 'to' (postposition), < pronominal stem *дер-*

---

3.   The stem *год-* /god-/ has been etymologically connected to Russian *год* /god/ 'year' by Mikhail Zhivlov (2014: 137), while more conservative etymological resources on Mari do not make such a connection (cf. Bereczki et al. 2013: 78, Moisio & Saarinen 2009 s.v. *γоδ-*). Irrespective of its origins, these forms are found in all Mari varieties and are thus not a recent addition to Mari.

4.   A number of postpositions have short and long forms; the difference between the forms in usage remains evasive (cf. Riese & Bradley & Yefremova 2022: 182).

- *ден* /den/ 'and; with; by, at' (conjunction, postposition), < postposition *дене* 'with; by, at'
- *дене*[5] /dene/ 'with; by, at' (postposition), < pronominal stem *дер-*
- *деран* /deran/ 'at, by' (postposition), < pronominal stem *дер-*
- *дегʹ(ын)* /deč(ə̂n)/ 'from' (postposition), < pronominal stem *дер-*
- *докан* /dokan/ 'probably, perhaps, it seems; (emphatic particle)' (particle)
- *дык* /də̂k/ 'if; well, so (emphatic particle)' (conjunction, particle)
- *дыр* /də̂r/ 'probably, perhaps, it seems' (particle), < Tatar

3.   Metrics of morphologization

This section will consider individual factors that could be considered as arguments for or against the morphologization of individual function words, and review points of discussion pertaining to individual function words. An overview of the data reviewed here is given in Table 3 in Section 4 below.

3.1.   Prosody

A critical feature in determining the morphologization of function words cannot be properly analysed based on written texts alone: do the relevant function words maintain their prosodic independence, or do they form units with their host words? The (esp. pre-literary) orthographic realization of function words in connection with their host words (see Section 3.1) implies, but does not conclusively prove, a lacking independence of these words. This has been noted in the literature as well, where a wide range of postpositions are said to form stress units with their base words and are said to not receive word stress (Ivanov 2000: 130).

Empirical research with Mari native speakers would be desirable here but might be especially frustrated by current[6] geopolitical circumstances: while it is possible to access Mari native speakers outside of Russia (notably in Estonia), it might not be possible to access the optimal informants for research of this type. It would be especially interesting to measure the spoken realization of said function words by speakers who are not familiar with the

---

5.   The relationship between *ден* 'and' and *дене* 'with' is somewhat complex, in that the shorter form is tendentially used as a conjunction and the longer from tendentially as a postposition, but this is by no means an ironclad rule; counterexamples can be encountered in both directions.

6.   These lines were written in June 2023.

Mari literary norms, but as the Mari community in Estonia largely has its roots in academic exchange the sample of accessible Mari native speakers is likely to be biased towards the literate.

3.2. Vowel harmony

While vowel harmony can be found throughout Mari varieties, its exact nature differs greatly between these. While Hill Mari shows both radical and suffixal palato-velar vowel harmony (Krasnova et al. 2017: 44–46), literary Meadow Mari shows no radical vowel harmony, but an alternation between unstressed final *-e*, *-o*, and *-ö* in suffixes (Riese & Bradley & Yefremova 2022: 33–34). Eastern varieties of Mari show an additional form of vowel harmony not rendered orthographically: labial harmony (both within word stems and affecting suffixes), where after an /ö/ or /ü/ in the first syllable the sound /e/ in subsequent syllables is labialized and becomes /ö/: *пӧлем* /pölem/ 'room' becomes /pölöm/, *мӧрем* /möröm/ 'my strawberry' becomes /möröm/, *ӧпкелем* /öpkelem/ 'I repent' becomes /öpkölöm/ (Ivanov 2000: 69). It has been noted that this process also transcends word boundaries and affects monosyllabic *дек* 'to', *ден* '(here) with', *деч* 'from': *пӧрт дек* /pört tek/ 'to the house' becomes /pört tök/, *кӱтӱ ден* /kütü den/ 'with the herd' becomes /kütü dön/, *шӱр деч* /šür dеč/ 'from the soup' becomes /šür döč/ (ibid.). Assimilation of this type affecting these D&G words can be found in transcriptions from the Mari text collections and dialect dictionaries, e.g., Meadow Mari *pört-tö·k* 'zu einem Haus' (Moisio & Saarinen 2008: s.v. *ðer-*) – but not consistently, cf. Eastern Mari *šär-γü·i̯ ðen* '[mit] Kieselstein[n]' (ibid.: s.v. *kə̂škem*); on occasion, versions with /ö/ can be encountered where seemingly not motivated by vowel harmony, e.g., Meadow Mari *sola-ðȍ·ṇo* 'bei dem Dorf' (ibid.: s.v. *ðer-*).

The two D&G words borrowed into Meadow-Eastern Mari (but not found in Hill Mari or Northwestern Mari) from Tatar, *гына* 'only, just' and *дыр* 'probably, perhaps' (already discussed in Section 3.4 above), show orthographically realized vowel harmonic adaptation to their base words in literary Tatar: *һава гына* 'only the weather', but *бер генә* 'only one' (Ersen-Rasch 2009: 53, 58); *кайдадыр* 'somewhere' < *кайда* 'where', but *кемдер* 'somebody' < *кем* 'who' (Asylgarayev et al. 2007: s.v. *дыр*).

Notably, Hill Mari palato-velar vowel harmony is not realized in the transcription of D&G words in text collections (e.g., Ramstedt 1902). However, there is a rather large corpus of recently grammaticalized morphology in Hill Mari that does not show vowel harmony: numerous highly frequent suffixes including the plural marker *-влä* and the derivational suffix *-лык* (Krasnova et al. 2017: 46), as well as the endings of the negative forms of the second past tense (Alhoniemi 1985: 119).

3.3.   Voice assimilation

The allomorphy of suffixes with an initial orthographic <д> or <г> is well-described in reference materials (e.g., Alhoniemi 1985: 34; Riese & Bradley & Yefremova 2022: 31): though orthography shows a voiced consonant, it is in pronunciation only voiced after vowels and (voiced) sonorant consonants; after (in Mari in the syllable-final position obligatorily voiceless) obstruents, a voiceless <t> or <k> is realized. This can be exemplified by the possessive suffix second person plural *-да* (*ялда* /jalda/ 'your village', but *пöртда* /pörtta/ 'your house', Alhoniemi 1985: 34) or the comitative ending *-ге* (*вуйге-почге* /vujge-poček/ 'completely (lit. with head and tail)', Riese & Bradley & Yefremova 2022: 31).

As detailed in Section 1, this allomorphy transcends orthographic word boundaries (Ivanov 2000: 101). This is oftentimes obvious in transcription-based text collections, even in cases where the D&G words are rendered as separate words separated from their host words with spaces rather than hyphens, e.g. (emphasis mine) the realization of *дене* 'with' here: *mijen kajenât nur dene, pasū dene, olăk tene,* [...] 'Sie begaben sich über Feld und Acker und die Wiese, [...]' (Paasonen & Siro 1939: 7).

It should be noted that the two D&G words borrowed into Meadow-Eastern Mari from Tatar, *гына* 'only, just' and *дыр* 'probably, perhaps', show the same voice assimilation pattern in Tatar, where the alternation is also orthographically realized in the literary standard (Ersen-Rasch 2009: XVIII, 44. 106; Asylgarayev et al. 2007: s.v. *гына, дыр*): *һава гына* 'only the weather', but *ун минут кына* 'only ten minutes' (Ersen-Rasch 2009: 53, 84). In the case of *дыр*, there is also no orthographic separation from the host word in literary Tatar: *кайдадыр* 'somewhere' < *кайда* 'where', but *ихластыр* 'honestly' < *ихлас* 'sincerity' (Asylgarayev et al. 2007: s.v. *дыр*).

3.4.   Orthographic realization

Though the spelling of D&G words as separate words is mostly conventionalized today (exceptions will be discussed below), as codified for both Meadow Mari (Ivanov et al. 2011) and Hill Mari (Vasikova 1994), there was considerably less regularity and certainty around this matter in the past. This can be illustrated by comparing different Mari-language bible translations separated by over a century (Tipo-Litografiya I. V. Yermolayevoy 1906; Raamatunkäännösinstituutti 2007), where the older version shows some D&G words connected to their host words with hyphens.

Matthew 2:1	*гыч* 'from, out of'
English (KJV):	[...] there came wise men from the east to Jerusalem,
1906:	[...] кэчы̌ лэкмаш могы̌р-гы̌чы̌н [...]
	(lit. 'from the sunrise side')
2007:	[...] эрвел гыч [...] (lit. 'from the morning side')

Mark 1:36	*дене* 'with'
English (KJV):	And Simon and they that were with him followed after him.
1906:	[...] туды̌н-дены̌ пы̌рльä (lit. 'with him together')
2007:	[...] тудын дене пырля

Matthew 5:20	*деч* 'from'
English (KJV):	[...] exceed the righteousness of the [...] Pharisees [...]
1906:	[...] правдалы̌кы̌шт-деч утла [...]
	(lit. 'from their righteousness more')
2007:	[...] фарисей-влак деч утларак [...]
	(lit. 'from the Pharisees more')

Other D&G words are not connected to their host words in the 1906 translation, e.g., *гы̌н* 'if' (modern *гын*) and *да* 'and'. Also, voice assimilation (see Section 3.4) is not orthographically indicated (as illustrated by Matthew 5:20 above). Furthermore, even for those D&G words encountered in combination with hyphens, the convention is not consistent: they are also encountered separated from their complements by a space, or not separated from them at all.

Matthew 2:6	*деч* 'from'
English (KJV):	[...] out of thee shall come [...].
1906:	[...] ты̌й дэчэт [...] (lit. 'from you')
2007:	[...] тый дечет [...]

Luke 9:5	*гыч* 'from, out of'
English (KJV):	[...] out of that city [...].
1906:	[...] туды̌ олагы̌ч [...]
2007:	[...] тудо ола гыч [...]

Alho Alhoniemi and Sirkka Saarinen describe their hardships as editors working with texts from the early 20th century as follows:

"Ausser im Hinblick auf die Lautgestalt der Wörter gibt es in den Märchen starke Inkonsequenzen bei der Getrennt- oder Zusammenschreibung einiger Wörter. Viele kleine Wörter, Partikeln wie *ta*, *δa* ,und', *γ�în̂*, *kͣ͂n* ,wenn', desgleichen einige Postpositionen, z.B. *ten*, *δen* ,bei, und', *tek*, *δek* ,zu' werden von Jevsevjev entweder unmittelbar nach dem Auslautbuchstaben des vorangehenden Wortes oder mit einem Bindestrich geschrieben; dann schreibt er sie wieder ganz getrennt. Uneinheitlich ist auch die Bezeichnung einiger suffixaler Elemente: die Endung *-lan* des Dativs z.B., das Suffix *-la* des Modals und z.B. das Possessivsuffix der 3. Sg. können ebenfalls entweder mit einem Bindestrich oder als selbständige Wörter geschrieben sein. In solchen Schwankungsfällen muss schon allein zur Erleichterung des Lesens vereinheitlicht werden, was dann gemäss dem Brauch der heutigen Schriftsprache geschieht. Wegen ihrer grossen Anzahl konnten diese Korrekturen nicht durch Fussnoten gekennzeichnet werden. Anzumerken ist Jedoch, dass das Morphem *γ�î͂ĉ̌*, *kͣ͂ĉ̌* ,aus' das der heutigen Schriftsprache nach als getrennte Postposition geschrieben wird, in Jevsevjevs Texten so konsequent mit dem vorangehenden Wort verbunden wird – nur zweimal begegnet kein Bindestrich –, dass es stets mit diesem Wort zusammengeschrieben wird." (Alhoniemi & Saarinen 1983: 23–24.)

Uncertainty in the written realization of D&G words is also self-evident in the works of early field researchers who worked directly with Mari native speakers and were not bound by norms of the then-incipient orthographies.

Original: Kuγo jükan toleš-kͧń, [...]
Modern: Кугу йӱкан толеш гын, [...]
Translation: 'Wenn ein Grobstimmiger kommt, [...]' (Genetz 1889: 25, 111)

Original (Hill):   iβanam roen-šͳ̣ndät-tä petskäškͧ opten-šͳ̣ndät, [...][7]
Modern (Hill):   Иваным роэн шӹндӓт да пецкӓшкӹ оптен шӹндӓт, [...]
Translation: 'Iwan hauen sie in Stücke und legen diese in eine Tonne, [...]'
                         (Ramstedt 1902: 177)

---

7.   Note also the hyphen connecting the two elements of aspectual converb constructions, cf. (Bradley 2016).

Original: n̥v tudo kā·t∫ɛ o·ɫno tȳl'a·∫ to·lɛ∫ y·dᵘrdɛ:k'ɛ.
Modern: Ну тудо каче олно тӱлаш толеш ӱдыр деке.
Translation: 'Der Bursche kommt zu dem Mädchen den Brautpreis bezah-
              len' (Lewy 1925–1926: I: 6, II: 7)

Original: [...] ši βakʼš-tè·kɘ̀ naŋgai̯a·t, [...]
Modern: [...] ший вакш деке нангаят, [...]
Translation: '[...] in eine silberne Mühle geschafft, [...]'
              (Wichmann 1931: 70)

Original: kɔnam apᵩšat molot ner-ɣɘts βn̥r̥ ᵈjoɣa, [...]
Modern: кунам апшат молот нер гыч вӱр йога, [...]
Translation: 'Wenn aus der Spitze des Hammers Blut fließt, [...]'
              (Beke 1931: 26)

At the same time, a number of cliticized elements in modern orthography
directly joined to their host words are also oftentimes separated from their
host words by hyphens in the early text collections, e.g., *n̥δn̥r-at* (modern
ӱдырат) 'auch das Mädchen' (Beke 1931: 25).

　　A systematic survey of early Mari language documentation and litera-
ture would be in order before excessively bold statements can and should be
made. The vast body of dialectal texts (described in Bradley 2016: 119–129) has
currently not been digitized in a publicly accessible, transparent, and search-
able manner, disallowing a time-efficient analysis of the written realization of
individual items. It is also not transparent to which extent scholars and writ-
ers were influenced by one another, and by orthographic norms of other lan-
guages (e.g., Finnish, where the elative case ending *-stA* is orthographically
realized in connection with its base word). However, the manner in which
D&G words were previously more commonly joined to their host words im-
plies that orthographic norms might be holding back, possibly even revers-
ing, morphologization processes, especially in respect to *гыч* 'from, out of ',
filling a conspicuous gap in the Mari nominal paradigm – which, according to
conventional descriptions, lacks an elative (separative, '(out) from') case to
go with its illative (directive, 'into') and inessive (locative, 'in').

　　In modern literary Mari, the joining of D&G words to their host words
is mostly restricted to individual lexicalizations (e.g., *эрдене* 'in the morning'
< *эр* 'morning' + *дене* 'with', *кастене* [sic] 'in the evening' < *кас* 'evening' +
дене 'with', *тыгодым* 'at this time' < *ты* 'this' + *годым* 'during'). An ex-
ception to this trend is the usage of *гыч* 'from, out of ' connected directly to
verbal stems, in combination with the verbs *кояш*[I] 'to appear, to be seen' and
шокташ[II] 'to sound, to be heard', illustrated in (2). This structure is quite

common: the Corpus of Literary Mari[8] and Timofey Arkhangelskiy's Meadow Mari Corpora[9] (see Hammer & Bradley 2022 for an overview of recently published corpus infrastructures for Mari) contains dozens of examples of this structure[10], with various different verbal stems occurring in (direct) combination with *гыҥ*.

Mari

(2)   [A] молан вара тый декет толгыч огыт кой?

a	*mo-lan*	*βara*	*tⱥj*	*dek-et*	*tol-gⱥč́*
but	what-DAT	then	2SG	to-PX2SG	come-out_of

ogⱥt   *koj?*
NEG.3PL   appear.CNG

'Why can one never see who comes to visit you?'
(Riese & Bradley & Guseva 2014: s.v. *кояш* (*-ям*))[11]

3.5.   Possessive suffixation

Postpositions and relational nouns alike take possessive suffixes in Mari (though for some postpositions, e.g. *годым* 'during', possessive suffix usage is rare, while for some such as *даҥыт* 'to, up to', no examples using possessive suffixes can be found in the corpus at all). Notably, however, there are some grave differences between the regular manner possessive suffixes attach to relational nouns and some postpositions (see Table 1):

> Some postpositions behave in a highly idiosyncratic manner when coupled with personal pronouns, demanding the usage of the nominative in some persons and the genitive in others, and demanding the usage of possessive suffixes in some persons but not allowing it in others. (Riese & Bradley & Yefremova 2022: 157.)

---

8.   ‹http://corpus.mari-language.com›

9.   ‹http://meadow-mari.web-corpora.net/index_en.html›

10.   These can be found in the Corpus of Literary Mari by searching for a word that ends in *гыҥ* but is not *гыҥ* (to exclude the free-standing postposition), followed by a word with the base form *кояш* or *шокташ*.

11.   All interlinear glosses in this contribution are my own; transcriptions were created using the COPIUS Transcription & orthography toolset at ‹http://www.copius.eu/ortho›.

**Table 1: PX attached to relational nouns, postpositions (Riese & Bradley &
Yefremova 2022: 157)**

	ончылно 'in front of'	*дек(е)* 'to'
1SG	ончылнем	мый декем
2SG	ончылнет	тый декет
3SG	ончылныжо	тудын деке
1PL	ончылнына	мемнан деке
2PL	ончылныда	тендан деке
3PL	ончылнышт	нунын деке

The obligatory usage of an overt pronoun in combination with these elements
is in line with the inadmissibility of these forms without a host word. More
surprising is the difference between the first and second person singular to
the other persons: in these persons, the pronoun is in the nominative and the
postposition takes a possessive suffix. In other persons, the pronoun is in the
genitive and there is no possessive suffix. No obvious explanation suggests
itself for this paradigm-internal discrepancy, but following the general prin-
ciple that highly frequent forms allow for more irregularity, irregular connec-
tion of possessive suffixes to a postposition can be seen as an argument for its
grammaticalization. The relevant data for this point can be found in the Mari
Web Project's Mari grammar for international learners (Riese & Bradley &
Yefremova 2022: 160–181); critically, only for D&G postpositions was idiosyn-
cratic possessive suffixation systematically observed (cf. ibid.: 157 about the
non-obligatory occurrence of idiosyncratic possessive suffixation on other
postpositions). As some postpositions do not conventionally take possessive
suffixes at all, this point of investigation is moot for them – as it is moot for
conjunctions and particles.

The connection of possessive suffixes to adnominal phrases deserves
investigation as well. When a possessive suffix is attached to a phrase consist-
ing of a noun and a relational noun, the possessive suffix can be attached to
either the noun (N-Px RN) or the relational noun (N RN-Px), the latter case
described as possessive suffix climbing e.g. by Kata Kubínyi (2015).

Mari

(3) a. суртем воктене   b.   сурт воктенем
     surt-em   *βoktene*   *surt*   *βokten-em*
     farmstead-PX1SG beside   farmstead beside-PX1SG
     'beside my farmstead' (Corpus of Literary Mari)

The same alternation (N-Px PP ~ N PP-Px) does also occur among the D&G postpositions (cf. Luutonen 1997: 36) – for those postpositions that conventionally take possessive suffixes, cases of both arrangements can be found.

Mari

(4) a. нерем деке   b. нер декем
     ner-em   deke   *ner   dek-em*
     nose-PX1SG   to   nose   to-PX1SG
     '(up) to my nose' (Corpus of Literary Mari)

Thus, in this domain there is no salient evidence for a difference in the degree of morphologization between relational nouns and D&G postpositions[12].

3.6.   Word order

Throughout the Uralic language family and beyond, a distinction is made between *relational nouns* and *adpositions*. Though these categories are functionally similar – both relational nouns and adpositions couple with a complement noun phrase and express spatial, temporal, and other relations – they differ in that relational nouns preserve a clear nominal character and in their morphosyntactic behaviour. Diachronically, especially body part nouns frequently are grammaticalized as relational nouns; relational nouns are frequently grammaticalized as postpositions; postpositions are frequently grammaticalized as case suffixes (cf. Grünthal 2022: 961).

While none of the relevant sources for Mari discussed in Section 2 distinguishes between relational nouns and postpositions (they are all classified as postpositions), a dividing line between these categories is salient in Mari, especially as regards syntactic distribution. This can be illustrated through the comparison of a relational noun *шеҥгелне* /šeŋgelne/ 'behind' (< noun *шеҥгел* /seŋgel/ 'back, rear, hind') and a postposition *дек(е)* /dek(e)/ 'to' (< pronominal stem *дер-* /der/): the first can occur with an overt preposed nominal compliment (5a), a pronominal complement marked only by a possessive suffix (6a), or with no complement noun phrase at all (7). In this usage relational nouns are oftentimes classified as adverbs and are in some sources even treated as separate items, though this seems questionable: it is clearly typical for one and the same word form to be used in these different syntactic positions (in combination with a host word as a postposition, independently

---

12. Also personal correspondence with Bogáta Timár, who is currently studying possessive suffix climbing in Mari.

as an adverb) and thus not necessary to postulate these as homonymic forms. For postpostions like *дек(е)* 'to' on the other hand, only usage with a preposed noun phrase is valid (5b); even pronominal complements must be expressed overtly (6b).

      Mari

(5) a. Йÿр шенгелне илет.
      jür   šeŋgelne il-et.
      rain   behind   live-2SG
      'You live behind the rain.' (Corpus of Literary Mari)

      b. Эчук деке каена.
      Ećuk   deke kaj-ena.
      Echuk   to   go-1PL
      'We're going to Echuk's.' (Corpus of Literary Mari)

(6) a. Шенгелнем ала-мо шырт-шорт шоктыш.
      šeŋgeln-em   ala-mo   šə̂rt-šort   šoktə̂-š.
      behind-PX1SG INDEF-what *šə̂rt-šort* sound-PST1.3SG
      'Something went *šə̂rt-šort* (crackled, crunched) behind me.'
      (Corpus of Literary Mari)

      b. Мый декем ит лишем!
      mə̂j   dek-em   it   lišem!
      1SG   to-PX1SG NEG.IMP.2SG approach.CNG
      'Don't get close to me!' (Corpus of Literary Mari)

(7)   Шенгелне — Москва!
      šeŋgelne —   Moskβa!
      behind   Moscow
      'Back there is Moscow!' (Corpus of Literary Mari)

Elements classified as "particles" in the reference materials are rather heterogenous as well. These include elements such as *можыг* 'maybe' that essentially form clauses of their own (see 8), and *гына* 'only', which immediately follows the element it modifies (see 9 for an illustration of how different placement of the element can radically alter its meaning) and cannot occur in sentence-initial position.

Mari

(8)   Можыч, кофем йӱына?
možâč̌, kofe-m   jü-ǝna?
maybe   coffee-ACC   drink-1PL
'Maybe let's drink some coffee?'
(Riese & Bradley & Yefremova 2022: 334)

(9) a.   Мый тыйым гына йӧратем, тыйым гына!
mâj   tâj-ǝm   gǝna jörat-em, tâj-ǝm   gǝna!
1SG   2SG-ACC   only love-1SG   2SG-ACC   only
'I love only you, only you!' (Corpus of Literary Mari)

   b.   Мый гына тыйым конга гыч лукмо шокшо мелнала вучем[.]
mâj   gǝna tâj-ǝm   koŋga   gǝč̌   luk-mo
1SG   only 2SG-ACC   oven   out_of   take_out-PTCP
šokšo   melna-la   βuč̌-em.
hot   pancake-CMPR   wait-1SG
'Only I am waiting for you like a hot pancake from the oven'
(Corpus of Literary Mari)

Timofey Arkhangelskiy's Meadow Mari Corpora are highly suited for determining if a word can occur in sentence-initial position (the many OCR and sentence boundary issues currently still found in the Corpus of Literary Mari preclude its usage as a tool here): It can easily be determined if a word ever occurs in sentence-initial position. Most words under consideration were inadmissible in sentence-initial position according to the corpus, with a few exceptions:

- *∂a* 'and; yes; so, well, indeed': there are two distinct origins of a conjunction/particle *∂a* in Mari, Chuvash and Russian (Moisio & Saarinen 2008: s.v. *δa*). At first glance one could suspect that only *∂a* mirroring modern Russian usage as a particle is used sentence-initially, but numerous encountered examples clearly are examples of this word as a conjunction (mirroring Chuvash and earlier Russian usage). It can also be encountered, as a conjunction, following a comma:

Mari

(10) Кокшайскыш кол кучаш кайынем, да билет уке.
  Kokšajsk-ə̂š   kol   kuč-aš   kajə̂-ne-m,   da   bilet   uke.
  Kokshaysk-ILL fish catch-INF go-DES-1SG and ticket NEG.EX
  'I want to go to Kokshaysk to fish, but I haven't got a ticket.'
  (Riese et al. 2017: 307)

- *даже* 'even': This clear Russian borrowing is used in accordance with Russian syntax and generally occurs sentence-initially, as was corroborated by the corpus.
- *дык* 'if; well, so': The sentence-initial usage of this particle was already noted when compiling the Mari Web Project's Mari grammar for international learners (Riese & Bradley & Yefremova 2022: 340) and was encountered in the corpus as well.

3.7. Functional and frequency factors

Alho Alhoniemi describes certain postpositions as *[s]ijajärjestelmää funktionaalisesti täydentävät muodot* 'forms that functionally complete the case system' (Alhoniemi 1985: 61); he describes *гыч* 'out of' as an elative marker and *деч* 'from' as an ablative marker (ibid.: 62). Along similar lines, *дек(е)* 'to' could be described as an allative marker, *дене* 'with; by, at' as an adessive (as well as comitative/instrumental) marker, *гоч* 'over, across, through' as a prolative marker, etc. Frequency alone is a poor argument for grammaticalization (but it can be a catalyst), but it is nevertheless interesting (though given the complementary usage, perhaps not surprising) to observe how many of these postpositions exceed some case suffixes in their frequency. Table 2 illustrates this through data on the occurrences of inflected[13] and postposition-marked forms of two nouns, the animate *айдеме* 'person' and inanimate *пӧрт* 'house', in the Corpus of Literary Mari with its currently 57.38 million tokens – for the sake of simplicity, plural forms as well as forms taking possessive suffixes and enclitic particles are excluded; all searches were case-insensitive.

---

13. Paradigms of Mari nouns can be created with the Mari paradigm generator at ‹http://paradigm.mari-language.com›.

Table 2: Frequency of case-inflected and postposition-marked forms of two common nouns

	'house'	# (ab-solute)	'person'	# (ab-solute)
Nominative	пӧрт	–	айдеме	–
Genitive	пӧртын	1,113	айдемын	5,040
Dative	пӧртлан	369	айдемылан	1,804
Accusative	пӧртым	4,815	айдемым	3,189
Comparative	пӧртла	–[14]	айдемыла	–
Comitative	пӧртге	5	айдемыге	0
Inessive	пӧртыштӧ	4,869	айдемыште	11
Illative	пӧртыш(кӧ)	7,364	айдемыш(ке)	301
Lative	пӧртеш	598	айдемеш	52
гыч 'out of' (~Elative)	пӧрт гыч(ын)	1,783	айдеме гыч(ын)	19
деч 'from' (~Ablative)	пӧрт деч(ын)	151	айдеме деч(ын)	228
дене 'with; at' (~Adessive)	пӧрт ден(е)	200	айдеме ден(е)	422
гоч 'through' (~Prolative)	пӧрт гоч(ын)	70	айдеме гоч(ын)	5
…				

Looking at the totality of word forms tagged as postpositions in the Corpus of Literary Mari[15], we can see that the postpositions considered here do indeed belong to the most frequent postpositions in Mari. The top 10 Mari postpositions in the corpus by frequency are:

1. *дене* 'with; at'
2. *гыч* 'out of'
3. *деч* 'from'
4. *нерген* 'about'
5. *годым* 'during'
6. *дек(е)* 'to'
7. *марте* 'up to'
8. *верч* 'for (the sake of)'
9. *гоч* 'through'
10. *мучко* 'across'[16]

---

14. The homonymy of the comparative case ending with a plural marker *-ла* renders corpus results unusable for this grammatical case without the manual analysis of all results.

15. Found by searching for "grammatical analysis" "contains" "Adp.Po" – as of the moment of writing, searching for the part-of-speech "postposition" does not yield any results due to a technical problem in the corpus infrastructure.

16. I have excluded *посна* 'separate' from this listing as it is in my opinion an adverbial element classified as a postposition due to its usage in the compound postposition *деч посна* 'without (lit. from separate)'.

4.   Summary & prospects

Table 3 provides an overview of the data collected on the individual points investigated in this contribution, where + indicates the presence of a feature, – indicates the absence of (evidence for) a feature, ? indicates an unclear classification, and / indicates that a metric is not applicable for the form at hand. In summation it can be said that the case for the morphologization of *гыц* 'out of' as an elative marker, *дец* 'from' as an ablative marker, *дек(е)* as an allative marker, and *дене* 'with; by, at' as an adessive/comitative/instrumental marker is the strongest (though it must be noted that the evidence is biased by the availability of more metrics here, namely vowel harmony, as only an original /e/ is available for vowel-harmonic alternation), while the Russian conjunction *даже* 'even' slipped into this investigation by accident; it can and should be disregarded from any further surveys pertaining to the morphologization of function words.

As regards the joining of postpositions to host words in old written texts, the data provided here is quite perfunctory – as there currently are no adequately digitized versions of the materials needed to systematically survey this factor, I could here only provide positive data (i.e., I found forms of the relevant item connected to a host word orthographically, either directly or with a hyphen), but no negative data (i.e., I was not able to corroborate if esp. the less common function words listed here are in fact not orthographically joined to host words as I could not with a reasonable expenditure of time and energy find usage examples of them in the sources).

The restriction to D&G words in this survey might have seemed arbitrary but is somewhat supported not only by methodology but also by the data itself: their phonological characteristics introduce some metrics to evaluate their degree of morphologization that are not available for other function words, but also the critical mass of the most common postpositions in Mari, as detailed in Section 3.7, actually do belong to this category. Be that as it may, there are function words with other initial sounds that show some symptoms of grammaticalization, for example *нерген* 'about' (highly frequent usage; does not occur sentence-initially) or *веле* 'only' (quite similar in usage and distribution to *гына* but unlike that element also used in Hill Mari and Northwestern Mari, cf. Moisio & Saarinen 2008: s.v. *βele*, *γə̂na*) that would also deserve attention through the lens of morphologization.

Table 3: Summary of data on the morphologization of individual D&G words

	Joined to host word in old texts	Not used sentence-initially	Un-usual Px	Voice assim-ilation	Vowel har-mony
гай 'like, as'	+	+	+	+	/
гала 'indeed, really'	?	+	/	+	/
гана 'times'	+	+	/	+	/
годсек 'since, beginning with'	?	+	/	+	/
годым 'during, at the time'	?	+	/	+	/
гоҕ(ын) 'over, across, through, via'	+	+	/	+	/
гутлан 'about, around (a time)'	?	+	/	+	/
гутлаште 'around, about, during'	?	+	/	+	/
гын 'if; as for …, … for one'	+	+	/	+	/
гына 'only, just; barely; rather'	?	+	/	+	/
гынат 'even if; some, any; even'	+	+	/	+	/
гыҕ(ын) 'from, out of'	+	+	/	+	/
да 'and; yes; so, well, indeed'	+	–	/	+	/
даже 'even'	–	–	/	–	/
даҥыт 'to, up to'	?	+	/	+	/
дек(е) 'to'	+	+	+	+	+
ден 'and; with; by, at'	+	+	+	+	+
дене 'with; by, at'	+	+	+	+	–
деран 'at, by'	?	+	+	+	–
деҕ(ын) 'from'	+	+	+	+	+
докан 'probably, perhaps'	?	+	/	?	/
дык 'if; well, so'	?	–	/	+	/
дыр 'probably, perhaps'	?	+	/	+	/

Acknowledgements

Abbreviations

1	1st person	IMP	imperative
2	2nd person	INDEF	indefinite
3	3rd person	INF	infinitive
ACC	accusative	NEG	negation
CMPR	comparative (case)	PL	plural
CNG	connegative	PST1	1st past tense
DAT	dative	PTCP	participle
DES	desiderative	PX	possessive suffix
EX	existential	SG	singular
ILL	illative		

References

Alhoniemi, Alho. 1985. *Marin kielioppi* (Apuneuvoja suomalais-ugrilaisten kielten opintoja varten 10). Helsinki: Suomalais-Ugrilainen Seura. <http://www.sgr.fi/apuneuvoja/apuneuvojax.pdf>

Alhoniemi, Alho & Saarinen, Sirkka (eds.). 1983. *Timofej Jevsevjevs Folklore-Sammlungen aus dem Tscheremissischen I* (Mémoires de la Société Finno-Ougrienne 184). Helsinki: Suomalais-Ugrilainen Seura.

Asylgarayev et al. 2007 = Асылгараев, Ш. Н. & Ганиев, Ф. А. & Закиев, М. З. & Миннуллин, К. М. & Рамазанова, Д. Б. 2007. *Татарско-русский словарь*. Kazan: Институт языка, литературы и искусства им. Г. Ибрагимова.

Beke, Ödön. 1931. *Tscheremissische Texte zur Religion und Volkskunde*. Oslo: Oslo Etnografiske Museum.

Bereczki, Gábor & Agyagási, Klára & Winkler, Eberhard. 2013. *Etymologisches Wörterbuch des Tscheremissischen (Mari)* (Veröffentlichungen der Societas Uralo-Altaica 86). Wiesbaden: Harrassowitz.

Bradley, Jeremy. 2016. *Mari converb constructions: Productivity and Regional Variance*. Vienna: University of Vienna. (Doctoral dissertation.) <https://www.doi.org/10.25365/thesis.43606>

Dixon, Robert M. W. 2010. *Basic linguistic theory, Volume 1: Methodology*. Oxford: Oxford University Press.

Ersen-Rasch, Margarete I. 2009. *Tatarisch: Lehrbuch für Anfänger und Fort-geschrittene*. Wiesbaden: Harrassowitz.

Galkin et al. 1990–2005 = Галкин, И. Г. et al. (eds.). 1990–2005. *Словарь марийского языка* [I–X]. Yoshkar-Ola: Марийское книжное издательство/МарНИИ. ‹https://dict.fu-lab.ru›

Genetz, Arvid. 1889. *Ost-tscheremissische Sprachstudien: Sprachproben mit deutscher Übersetzung* (Journal de la Société Finno-Ougrienne 7). Helsinki: Suomalais-Ugrilainen Seura.

Givón, Thomas. 1971. Historical syntax & synchronic morphology: An archeologist's field trip. *Papers from the Seventh Regional Meeting* 7. 394–415.

Grünthal, Riho. 2003. *Finnic adpositions and cases in change* (Mémoires de la Société Finno-Ougrienne 244). Helsinki: Suomalais-Ugrilainen Seura. (Doctoral dissertation.)

Grünthal, Riho. 2022. Adpositions and adpositional phrases. In Bakró-Nagy, Marianne & Laakso, Johanna & Skribnik, Elena (eds.), *The Oxford guide to the Uralic languages*. Oxford: Oxford University Press. ‹https://www.doi.org/10.1093/oso/9780198767664.003.0050›

Hammer, Luan & Bradley, Jeremy. 2022. Mari morpheme order revisited: A corpus-based analysis. *Nordlyd* 46(1). 59–74. ‹https://www.doi.org/10.7557/12.6373›

Haspelmath, Martin & Sims, Andrea D. 2010. *Understanding morphology* (Understanding language series). [2nd edition.] London: Hodder Education.

Itkonen, Erkki. 1960. Tšeremissin kielen sanaliittojen suhteesta yhdyssanoihin ja taivutusmuotoihin. *Virittäjä* 64. 311–319.

Ivanov 2000 = Иванов, И. Г. 2000. *Кызытсе марий йылме: Фонетика*. Yoshkar-Ola: Марий книга савыктыш.

Ivanov et al. 2011 = Иванов, И. Г. et al. (eds.). 2011. *Марий орфографий мутер*. Yoshkar-Ola: МарНИИЯЛИ. ‹http://mari-lab.ru/index.php?title=Марийские_словари›

Krasnova, Nadezhda & Riese, Timothy & Bradley, Jeremy & Yefremova, Tatiana. 2017. *Reading Hill Mari through Meadow Mari*. 1.0. Vienna: University of Vienna. ‹http://rhm.mari-language.com/›

Kubínyi, Kata. 2015. Possessive clitic climbing as a pattern of agreement with the possessor in Permic and Mari postpositional phrases. (Presentation given at the XII International Congress for Finno-Ugric Studies in Oulu, August 17.)

Lewy, Ernst. 1925–1926. *Tscheremissische Texte* [I–II]. Hanover: Orient-Buchhandlung Heinz Lafaire.

Luutonen, Jorma. 1997. *The variation of morpheme order in Mari declension* (Mémoires de la Société Finno-Ougrienne 226). Helsinki: Suomalais-Ugrilainen Seura.

Luutonen, Jorma & Saarinen, Sirkka & Moisio, Arto. 2007. *Electronic word lists: Mari, Mordvin and Udmurt* (Lexica Societatis Fenno-Ugricae 31:1). Helsinki: Suomalais-Ugrilainen Seura.

Moisio, Arto & Saarinen, Sirkka. 2008. *Tscheremissisches Wörterbuch* (Lexica Societatis Fenno-Ugricae 32). Helsinki: Suomalais-Ugrilainen Seura.
‹https://doi.org/10.33341/sus.877›

Paasonen, Heikki & Siro, Paavo. 1939. *Tscheremissische Texte*. Gesamm. von H. Paasonen, Hrsg. von Paavo Siro (Mémoires de la Société Finno-Ougrienne 78). Helsinki: Suomalais-Ugrilainen Seura.
‹http://fennougrica.kansalliskirjasto.fi/handle/10024/89914›

Raamatunkäännösinstituutti. 2007. *У Сугынь*. Helsinki: Raamatunkäännös-instituutti.

Ramstedt, G. J. 1902. *Bergtscheremissische Sprachstudien* (Mémoires de la Société Finno-Ougrienne 17). Helsinki: Suomalais-Ugrilainen Seura.
‹http://fennougrica.kansalliskirjasto.fi/handle/10024/89840›

Riese, Timothy & Bradley, Jeremy & Guseva, Elina. 2014. *Mari-English dictionary*. Vienna: University of Vienna. ‹http://dict.mari-language.com/›

Riese, Timothy & Bradley, Jeremy & Yakimova, Emma & Krylova, Galina. 2017. *Оҥай марий йылме: A comprehensive introduction to the Mari language* [Release 3.2]. Vienna: University of Vienna.
‹http://omj.mari-language.com/›

Riese, Timothy & Bradley, Jeremy & Yefremova, Tatiana. 2022. *Mari (марий йылме): An essential grammar for international learners* [Release 1.0]. Vienna: University of Vienna.
‹http://grammar.mari-language.com/›

Tipo-Litografiya I. V. Yermolayevoy. 1906. *Святое Евангелие на марийском языке*. Kazan: Типо-Литографія И. В. Ермолаевой (бывш. Ключникова).

Vasikova 1994 = Васикова, Л. П. 1994. *Кырык марла орфографи лӹмдер*. Yoshkar-Ola: Мары Элӹн периодика.

Wichmann, Yrjö. 1931. *Volksdichtung und Volksbräuche der Tscheremissen* (Mémoires de la Société Finno-Ougrienne 59). Helsinki: Suomalais-Ugrilainen Seura.
‹http://fennougrica.kansalliskirjasto.fi/handle/10024/89893›

Zhivlov, Mikhail. 2014. Studies in Uralic vocalism III. *Вопросы языкового родства* 12. 113–148.

All online resources were last accessed on 21 December 2023.