

## **External Merge to Specifier of CP: Complementizers Projecting an Argument**

**Abstract:** The standard assumption that Spec,CP is always an A-bar position has been questioned for several languages where embedded C heads are involved in agreement and case-assignment; however, the idea that no XP can be introduced in Spec,CP by external merge has remained unchallenged. The paper presents novel control data from Mari (Uralic; nominative, SOV) and argues that, in this language, a particular type of C head is capable of thematically licensing an overt argument externally merged in Spec,CP: the complementizer *manən* used in infinitival complement clauses projects a dative Goal of communication. This behavior of *manən* follows from its dual nature: it is a semi-grammaticalized verb ‘say’ that retains some lexical characteristics. I further suggest that the dative Goal can be considered an overt realization of the ADDRESSEE discourse variable, in line with the recent work on the presence of SPEAKER, ADDRESSEE, and logophoric center in syntax.

**Keywords:** logophoric control, complementizer, CP, discourse arguments, A/A-bar distinction, grammaticalization.

### **1 Introduction**

The standard assumption that Spec,CP is always an A-bar position has been questioned for several languages, including Japanese, Korean, Mongolian,<sup>1</sup> i.a., where an embedded C head is involved in agreement and Case-assignment. This confirms that at least some C heads can have A-features in addition to A-bar features; see Wurmbrand (2019) for an overview of the problem and references therein. However, the idea that no XP can be introduced in Spec,CP by external merge, put forward by Chomsky (2000:102), has remained unchallenged. In this paper I argue that a C head can thematically license a DP in Spec,CP. Presenting previously

---

<sup>1</sup> See Tanaka (2002), Horn (2008), and Yoon (2007) on ECM in embedded CPs in Japanese and Korean and Fong (2019) on hyperraising in Mongolian.

undescribed data from Mari,<sup>2</sup> a Uralic language (head-final, SOV), I demonstrate that a particular type of embedded C head projects a Goal argument.

The discussion is centered on obligatory control sentences with a matrix speech act verb. In Mari speech act verbs serve as mandative predicates (also known as directives and verbs of order) when they embed an infinitival/subjunctive clause.

(1) a. Maša mə-lam            tol-aš            (manən)            kalas-en.<sup>3</sup>  
       Maša I-DAT.1SG            come-INF            COMP            tell-PST2  
       ‘Maša told me to come.’

b. Maša mə-lam            tə-lan-et            tol-aš            (manən)            kalas-en.  
       Maša I-DAT.1SG            you-DAT- POSS.2SG come-INF            COMP            tell-PST2  
       ‘Maša told me for you to come.’<sup>4</sup>

---

<sup>2</sup> Unless specified otherwise, the data presented in the paper come from the Morkinsko-Sernur dialect of Meadow Mari (Eastern Mari) spoken in the Mari El republic, Russia. Several examples come from the Kuznetsovo variety of Hill Mari (Western Mari) spoken in the Kuznetsovo village, Mari El. The data have been collected during my field work in 2019–2020. The double-dative construction under discussion is attested in both varieties and I have found no difference in the distribution.

<sup>3</sup> Glossing abbreviations: ACC = accusative, COMP = complementizer, CONJ = conjunction, CVB = converb, DAT = dative, EL = elative, GEN = genitive, IMP = imperative, IN = inessive, INF = infinitive, JUS = jussive, NEG = negation, NPST = non-past, PL = plural, POSS = possessive, PROG = progressive, PST = past (aorist), PST2 = past (perfective), SG = singular.

<sup>4</sup> Throughout the paper I accompany the double-dative examples with neutral translations; see Section 2 for a detailed discussion of interpretation of such sentences.

The question arises of how to analyze (1b), keeping in mind that in Mari it is usually prohibited to have two dative dependents in a single clause and that double-datives are not attested, for instance, in Hungarian (another Uralic language) or in Russian (a contact language). As I will demonstrate, the first ‘left’ dative DP (*məlam*) in (1b) is merely a matrix Goal of communication, in parallel to the only dative DP in (1a). The second ‘right’ dative DP (*təlanet*) is more interesting. It forms a constituent with the infinitival clause that excludes the matrix verb, however, it cannot be analyzed as an overt embedded subject (2).

(2) [ ... DP<sub>DAT1</sub> ... [XP DP<sub>DAT2i</sub> [X' [ PRO<sub>i</sub> infinitive ] X<sup>0</sup> ] ... SAY ]

I argue that the X head in (2) is the embedded complementizer head. The second dative DP (DP<sub>DAT2</sub>) is base-generated in Spec,CP and is thematically licensed by C manifested as the complementizer *manən*/its null allomorph; it controls the embedded PRO subject and denotes a participant that is simultaneously the ultimate Goal of communication and the mandee.

(3) [VP DP<sub>DAT1</sub> [V' [CP DP<sub>DAT2i</sub> [C' [FinP PRO<sub>i</sub> [Fin' [TP t<sub>i</sub> infinitive ] Fin<sup>0</sup> ]]] C<sup>0</sup> *manən* ]]] V<sup>0</sup> ]]

I further propose that the exceptional ability of *manən* to project an argument follows from its semi-grammaticalized status: it is derived from the verb of communication *manaš* ‘say, tell’ and retains some of its lexical properties, including the ability to introduce a Goal of communication; see Heine & Kuteva (2002) for a discussion of grammaticalization of ‘say’ into complementizers in the world’s languages.

The paper contributes to the ongoing discussion of the properties of CPs by expanding the range of A-type phenomena that an embedded C head can be involved in. Additionally, I

suggest that the second dative DP projected by the embedded C head (DP<sub>DAT2</sub> in (3) above) can be considered an overt realization of the ADDRESSEE discourse variable; see Speas (2004), Baker (2008), Sundaresan (2018), Spadine (2018, 2019), i.a., on the presence of (overt or covert) SPEAKER, ADDRESSEE, and logophoric center in syntax.

The paper proceeds as follows. Section 2 describes the relevant syntactic properties of the double-dative construction. Section 3 focuses on the complementizer *manən* and its behavior. Section 4 presents a formal analysis for the mandative sentences and dismisses alternative approaches. Section 5 concludes the paper by discussing directions for future research.

## 2 Double-dative constructions

### 2.1 Speech act verbs used as mandative predicates

Mari verbs of communication, such as *kalasaš* ‘say, tell’, *kutəraš* ‘say, speak’, and also *kəčkəraš* ‘shout’ or *seraš* ‘write’, etc., are interpreted either as plain speech act predicates or as mandatives depending on the type of the clausal complement. In (4) the verb *kalasaš* ‘say, tell’ embeds a finite indicative clause and the sentence receives a standard declarative reading.

- (4) Maša mə-lan-na [rveze-vlak kniga-m už-ən-ət manən] kalas-en.  
 Maša we-DAT-POSS.1PL boy-PL book-ACC see-PST2-3PL COMP tell-PST2  
 ‘Mary told us that the boys had seen the book.’

The structure of such sentences is outlined in (5). I adopt an analysis à la Chomsky (1981) whereby the verb combines first with a CP argument, the linguistic material (in Comp,VP) and then with a Goal DP (in Spec,VP). The Agent is projected by a separate head, *v* (Chomsky 1995). I further assume that the Goal is assigned inherent DAT and the Agent is probed by the T head and is assigned NOM under agreement. As noted in Section 1, Mari is a head-final language with the standard word order being SOV.

(5) [<sub>VP</sub> Agent [<sub>V'</sub> [<sub>VP</sub> Goal [<sub>V'</sub> [<sub>CP</sub> ... ] V<sup>0</sup> ] ] V<sup>0</sup> ]]

In contrast to (4), in (6a) and (6b) the same verb of communication *kalasaš* ‘tell’ selects a finite subjunctive clause and an infinitival clause, respectively, as its complement and the sentences must be interpreted as orders. The embedded non-finite clause can optionally<sup>5</sup> be accompanied by the complementizer *manən*, which will be discussed in detail in Section 3.

(6) a. Maša mə-lan-na [rveze-vlak kniga-m už-əšt manən]kalas-en.

Maša we-DAT-POSS.1PL boy-PL book-ACC see-JUS.PL COMP tell-PST2

‘Maša told us that the boys should see the book.’

b. Maša mə-lan-na [tol-aš (manən)]kalas-en.

Maša we-DAT-POSS.1PL come-INF COMP tell-PST2

‘Maša told us to come.’

Let us focus on the mandative construction in (6b). It contains a dative DP (DP<sub>DAT</sub>) that simultaneously refers to the Goal of communication (the one who receives the message and can pass it on) and the obligation holder (the one who should carry out the order). The DP<sub>DAT</sub> and the understood embedded subject must be co-indexed. However, as will be shown in Section 4, the coreference can be partial; as further demonstrated in (7), non-c-command

---

<sup>5</sup> So far I was unable to determine what regulates the distribution of *manən* and its null allomorph. Some of the speakers that I consulted prefer the overt complementizer and some suggest that its use is completely optional. The ‘covertness’ of *manən* does not appear to correlate with the availability of the second dative DP, and is thus not directly relevant here.

control and long-distance control are prohibited. Thus I assume that the embedded subject is an obligatorily controlled PRO and not a copy/trace or *pro*.

- (7) Mašak [təj [Petja-n<sub>m</sub> joltaš-əžə-vlak-lan]<sub>i</sub> [PRO<sub>i/\*k/\*m</sub> təšecə kaj-aš]  
 Maša you Petja-GEN friend-POSS.3SG-PL-DAT here.ELgo-INF  
 kalas-en-at manən] pal-a.  
 tell-PST2-2SG COMP know-NPST.3SG  
 ‘Maša knows that you told Petja’s friends to come.’

## 2.2 Double datives: Overview

The declarative/mandative ambiguity illustrated in the previous subsection is typical for speech act verbs in many languages, including English, Russian, Spanish, among others. What makes the Mari case interesting is that in control sentences with a mandative interpretation two non-coordinated dative nominal phrases can appear. This is illustrated in (8) for an embedded intransitive verb and in (9) for an embedded transitive verb.

- (8) Maša mə-lam (tə-lan-et) tol-aš manən kalas-en.  
 Maša I-DAT.1SG you-DAT-POSS.2SG come-INF COMP tell-PST2  
 ‘Maša told me (for you) to come.’
- (9) Təj mə-lan-na (Petja-lan) kapka-m ačal-aš manən kalas-əš-əč.  
 you we-DAT-POSS.1PL Petja-DAT gate-ACC fix-INF COMP tell-PST-2SG  
 ‘You told us (for Petja) to fix the gate.’

The construction is schematized in (10). Throughout the paper I refer to the first dative DP as  $DP_{DAT1}$  (*məlam* in (8) and *məlanna* in (9)) and to the second dative DP as  $DP_{DAT2}$  (*təlanet* in (8) and *Petjalan* in (9)); the  $DP_{DAT2}$  is always in bold. The  $DP_{DAT2}$  controls the embedded PRO; Section 2.3 provides support against a raising/ECM analysis. Examples in (11) demonstrate that control is obligatory and conforms to the c-command and locality restrictions.

(10) [ $DP_{DAT1}$   **$DP_{DAT2i}$**  [ $PRO_i$  infinitive] verb]

(11) Maša mə-lan-na<sub>k</sub>            [**Petja-n<sub>m</sub>**            **joltaš-əžə-vlak-lan**]<sub>i</sub>  
 Maša we-DAT-POSS.1PL Petja-GEN            friend-POSS.3SG-PL-DAT  
 [ $PRO_{i/*k/*m}$  təšəč            kaj-aš            manən]            kalas-en.  
    here.EL            go-INF            COMP            tell-PST2  
 ‘Maša told us for Petja’s friends to come.’

As captured by the prose translation in (8) and (9), the first dative DP ( $DP_{DAT1}$ ) refers to the immediate Goal of communication, i.e. the intermediary that receives the original message.<sup>6</sup> Typically for Goals of communication it is restricted to [+Sentient] referents (usually [+Human]); for example, an inanimate means of communication – a letter or a message –

---

<sup>6</sup> The intermediary may be responsible for controlling the task: in this case, (8) receives the reading ‘Mary told me to ensure that you will come’. However, this is not required, as (8) can also be interpreted as ‘Mary told me to tell you to come’, with a plausible continuation along the line ‘... but I did not tell you’ or ‘... but you did not come’. The semantic properties of an obligation holder and a Goal of communication are blended in the  $DP_{DAT2}$  and this participant is best described as the ultimate recipient of the order; see Section 2.3.

cannot be marked dative (12). A [+Animate] but non-human  $DP_{DAT1}$  is accepted only if the context allows animals to become appropriate addressees, for instance, in a fairy tale.

- (12) Maša serəš-əšte / \*serəš-lan mə-lam tol-aš manən kalas-en.  
 Maša letter-IN letter-DAT I-DAT.1SG come-INF comp tell-PST2  
 ‘In a letter, Maša told me to come.’

The  $DP_{DAT1}$  is structurally equivalent to the  $DP_{DAT}$  in the single-dative construction, which becomes evident in sentences with idiosyncratic case-marking. For instance, the verb *sörvalaš* ‘beg’ requires an accusative Goal; the  $DP_{ACC}$  can co-occur with an independent  $DP_{DAT2}$  (13) and such sentences receive interpretations parallel to those in (8) and (9).

- (13) Maša {jumə-m / \*jumə-lan} (mə-lan-na) tol-aš manən sörval-en.  
 Maša God-ACC God-DAT we-DAT-POSS.1PL come-INF COMP beg-PST2  
 ‘Maša begged God (for us) to come.’

The question arises about the structural position and licensing of the  $DP_{DAT2}$ . As mentioned in the introduction, double-dative constructions are not attested in Russian (a contact language) or in Hungarian (another Uralic language). In Mari too double datives are regularly prohibited within a single clause: consider example (14) showing that Mari allows dative Beneficiaries and Goals but those cannot co-occur in the same clause.

- (14) Məj Maša-lan (\*tə-lan-et) vurgem-əm nal-ən-am.  
 I Maša-DAT you-DAT-POSS.2SG clothes-ACC buy-PST2-1SG  
 ‘I bought Maša clothes.’ or ‘I bought clothes for Maša, on her behalf.’



Not available: ‘I bought you clothes for Maša, on her behalf.’ or ‘I bought Maša clothes for you.’

The next subsection addresses this question by considering the relevant properties of the  $DP_{DAT2}$  and showing that it belongs to the same immediate constituent as the non-finite clause but is not thematically licensed by the embedded predicate.

### 2.3 $DP_{DAT2}$ as a Goal/obligation holder

Let us take a closer look at the  $DP_{DAT2}$ . It forms a constituent with the embedded non-finite clause that excludes the  $DP_{DAT1}$  and the matrix predicate. For instance, the  $DP_{DAT2}$  and the infinitive cannot be separated by a matrix adverb (15), even though in Mari adjuncts scramble freely within a clause.

- (15) (Tače) təj (tače) mə-lan-na [Petja-lan (\*tače) kapka-m erla  
 today you today we-DAT-POSS.1PL Petja-DAT today gate-ACC tomorrow  
 ačal-aš manən] kalas-əš-əč.  
 fix-INF COMP tell-PST-2SG  
 ‘Today you told us for Petja to fix the gate tomorrow.’

Likewise, the  $DP_{DAT2}$  and the non-finite clause must be dislocated together under extraposition (16) and in fragment answers (17).

- (16) Təj m-lam (\*Petja-lan) kalas-əš-əč [(Petja-lan) kapka-m ačal-aš manən].  
 you I-DAT.1SG Petja-DAT tell-PST-2SG Petja-DAT gate-ACC fix-INF COMP  
 ‘You told me (for Petja) to fix the gate.’



c. #Maša mə-lan-na [ušk-al-lan šudə-m kočk-aš] kalas-en.

Maša we-DAT-POSS.1PL cow-DAT grass-ACC eat-INF tell-PST2

Intended: ‘Maša told us for the cow to eat grass.’

The DP<sub>DAT2</sub> is further restricted to referring to a conscious addressee that can, potentially, receive the message. Crucially, the same restriction applies to all Goals of communication. Consider the following scenario. The children are already asleep and they should sleep until the evening. The doctor has talked to me and asked to check on them. In this situation (19a), which contains a finite subjunctive clause, is perfectly acceptable as an order/recommendation. In contrast, (19b), where an embedded non-finite clause is accompanied by a dative DP, is infelicitous: since the children are already asleep, it does not make sense for me to talk to them and tell them to continue sleeping until the evening; the sentence would be appropriate only if the children were awake at that moment.

(19) a. Vrač mə-lan-em [joča-vlak kas marte mal-əšt manən] kalas-en.

doctor I-DAT-POSS.1SG child-PL evening until sleep-JUS.PL COMP tell-PST2

‘The doctor told me that the children should sleep until the evening.’

b. #Vrač mə-lan-em joča-vlak-lan kas marte mal-aš manən kalas-en.

doctor I-DAT-POSS.1SG child-PL-DAT evening until sleep-INF COMP tell-PST2

Intended: ‘The doctor told me for the children to sleep until the evening.’

The same restriction applies to the DP<sub>DAT</sub> in single-dative sentences: similarly to (19b), (20) is only felicitous if the doctor could address the children directly, i.e. if they were awake.

(20) Vrač joča-vlak-lan kas marte mal-aš manən kalas-en.

doctor child-PL-DAT evening until sleep-INF COMP tell-PST2

‘The doctor told the children to sleep until the evening.’

Taking these data into account, I argue that the DP<sub>DAT2</sub> is interpreted not just as an obligation holder, but as a combination Goal/mandee, that is, it is assigned a Goal role by a nearby head (as will be shown later in the paper, the embedded semi-grammaticalized complementizer).

Two other pieces of data provide evidence against analyzing the DP<sub>DAT2</sub> as an argument of the embedded predicate. First, double-dative sentences do not pass the idiom chunk test, typically used to distinguish control and raising. In (21a) the expression *šem pərəs koklaštəna kudal ertəš*, literally ‘the black cat ran between us’, is embedded under ‘tell’ as a combination of a non-finite clause and a DP<sub>DAT</sub>. It does not retain the idiomatic reading ‘we quarreled’ even though it is possible to say non-periphrastically ‘Maša told Peter for us to quarrel’ (21b).

(21) a. #Maša Petja-lan [šem **pərəs-lan** koklaštə-na kudal ert-aš manən] kalas-əš.

Maša Petja-DAT black cat-DAT between-POSS.1PL run-INF COMP tell-PST

‘Maša told Petja to tell the cat to run between us.’

Not available: ‘Maša told Petja for us to quarrel.’

b. Maša Petja-lan **mə-lan-na** vursedəl-aš kalas-əš.

Maša Petja-DAT we-DAT-POSS.1PL quarrel-INF tell-PST

‘Maša told Petja for us to quarrel.’

Second, unlike ordinary subjects (22a) and similarly to matrix dependents (22b), the  $DP_{DAT2}$  cannot reconstruct under the embedded negation (22c), which suggests that it is externally merged higher in the structure.<sup>7</sup>

(22) a. [Čəla rveze-vlak] em-əm jü-ən ogətəl.

all boy-PL medicine-ACC drink-CVB NEG.PST.3PL

‘All the boys did not drink the medicine.’ / ‘Not all boys drank the medicine.’

b. Petja [čəla rveze-vlak-lan]<sub>i</sub>[PRO<sub>i</sub> em-əm jü-aš ogəl manən] kalas-en.

Petja all boy-PL-DAT medicine-ACC drink-INFNEG COMP tell-PST2

Only: ‘Petja told all the boys that they should not drink the medicine.’

c. Maša mə-lan-na [[čəla rveze-vlak-lan] em-əm jü-aš ogəl

Maša we-DAT-POSS.1PL all boy-PL-DAT medicine-ACC drink-INF NEG

---

<sup>7</sup> It might be suggested that the  $DP_{DAT2}$  is base-generated as an argument of the embedded predicate and later undergoes A-movement to a position at the left periphery where it receives a second theta-role (Susi Wurmbrand and Idan Landau p.c.). For instance, Wurmbrand & Lohninger (2020) point out that DPs involved in cross-clausal A-dependencies in Japanese also do not reconstruct and tentatively attribute this to the presence of the second theta-role. At this point, I have no independent evidence for or against a multiple theta-role analysis and leave the issue for future research. Note, however, that no other examples that could involve this phenomenon have been found in Mari. Moreover, adopting such an approach would not undermine the crucial idea that the  $DP_{DAT2}$  receives the Goal role outside of the non-finite embedded TP, as argued in this paper.

manən] kalas-en.

COMP tell-PST2

Strongly preferred: ‘Maša told us that all the boys should not drink the medicine.’

The properties of the  $DP_{DAT2}$  discussed above are straightforwardly explained under the assumption that there is an intermediate head that takes a non-finite clause as its complement and introduces the  $DP_{DAT2}$ , assigning to it the Goal of communication role. In what follows I argue that this head is the C head manifested as a semi-grammaticalized complementizer *manən*. At the same time that the  $DP_{DAT2}$  is not selected by the embedded predicate posits a serious problem to an analysis whereby it is the embedded subject occupying the Spec,TP position (23).

(23) [ $DP_{DAT1}$  ... [ $V'$  [ $FinP/CP$  [ $TP$   $DP_{DAT2}$ ; infinitive]  $Fin/C^0$  ]  $V^0$ ] – rejected structure

### 3 Double-dative constructions: The role of *manən*

#### 3.1 $DP_{DAT2}$ is projected by the C head

In double-dative constructions under discussion the  $DP_{DAT1}$  is a matrix Goal and the  $DP_{DAT2}$  is simultaneously a Goal and a mandee related to the non-finite clause but it is not an argument of the embedded predicate. This is schematized in (24), based on the standard structure for verbs of communication in (5), where  $X^0$  connects the  $DP_{DAT2}$  and the embedded clause.

(24) [ $VP$   $DP_{DAT1}$  [ $V'$  [ $XP$   $DP_{DAT2}$ ; [ $X'$  [ $PRO_i$  infinitive ]  $X^0$ ] ]  $SAY$  ]]

I argue that the X head is a C head of a particular type: its exponent is the complementizer *manən* or its null allomorph.

(25) Maša mə-lan-na [CP **Petja-lan**<sub>i</sub>[<sub>FinP</sub> PRO<sub>i</sub> tol-aš] (manən)] kalas-en.

Maša we-DAT-POSS.1PL Petja-DAT come-INF COMP tell-PST2

‘Maša told us for Petja to come.’

I propose that the complementizer *manən* not only selects a non-finite FinP as its complement but also exceptionally projects an argument in Spec,CP – the DP<sub>DAT2</sub> – and assigns to it the Goal role. This analysis accounts for all properties of the DP<sub>DAT2</sub> listed above including the selectional restriction. Furthermore, it straightforwardly captures the correlation: only those predicates that can embed a non-finite complement clause with the complementizer *manən* allow double datives; for instance, evaluative adjectives, which do not embed non-finite clauses with *manən*, can never appear with two dative DPs (26).<sup>8</sup>

(26) Ač’a-ž-lan [(**\*mə-lan-na**) təšeč kaj-aš (**\*manən**)] nele.

father-POSS.3SG-DAT we-DAT-POSS.1PL here.EL go-INF COMP hard

‘For his/her father it is hard to leave.’

---

<sup>8</sup> Sentences with *nele* ‘hard, difficult, heavy’ show negative results for the standard raising diagnostics, including the selection test (i).

(i) #Kogəl’-lan kü-aš(-əžə) nele.

pie-DAT cook-INF-POSS.3SG hard

Intended (infelicitous): ‘It is difficult for the pie to cook.’

The ability to assign a thematic role is considered to be a property of lexical heads and Spec,CP has traditionally been described as an A-bar position suitable for internal merge of dislocated elements but not for external merge of brand-new participants (Chomsky 2000). I argue that the exceptional status of *manən* as a complementizer that can project an argument is due to its being a semi-grammaticalized<sup>9</sup> element derived from the verb *manaš* ‘say, tell’; see Savatkova (2002), Toldova & Serdobolskaya (2014) for a discussion of the history of *manən* and Serdobolskaya & Toldova (2011) on ‘say’-based functional items in Udmurt, another Uralic language. I propose that in modern Mari the following entries for *manən* coexist.

First, there is a lexical verb of the type  $\langle e, \langle e, t \rangle \rangle$  that selects a proposition and projects a Goal argument (27a). Second, there is a ‘pure’ fully grammaticalized complementizer that is used in indicative/subjunctive CP complements, adjunct purpose clauses, and subject clauses (27b). I assume that it combines with a fully saturated FinP ( $\langle s, t \rangle$ ) and can reasonably be compared to *that* in English or *što/štoby* in Russian; see Demirok et al. (2020) and references therein

---

<sup>9</sup> Complementizers that are traced back to ‘say’ verbs are attested in many language families, including, for instance, Indo-Iranian, Austronesian, and Nakh-Daghestanian languages (Hock 1982, Klamer 2000, Heine & Kuteva 2002, Daniel 2007). Throughout the paper I use the term „semi-grammaticalized” to indicate that, in modern Mari, *manən* has not yet become an exclusively functional item void of any lexical properties and that several entries of *manən* corresponding to different grammaticalization stages are used in parallel. In the literature, the term is also used to refer to those complementizers whose distribution is restricted to a particular type of embedded clauses; cf. for instance, semi-complementizers in Cantonese Yue and Hakka as discussed by Chappell (2008). That lexical roots can be merged into functional positions as a step on the grammaticalization path was also proposed, for example, for aspectual and modal verbs in Dutch and Afrikaans by Cavirani-Pots (2020).



discussing the semantics of such elements. Finally, there is a semi-functional complementizer that is of interest in this paper (27c). It appears in non-finite clausal complements of speech act verbs (i.e. in ‘speech act’ contexts); as will be discussed in more detail in Section 4, this *manən* combines with a property-type FinP ( $\langle e, \langle s, t \rangle \rangle$ ) and is capable of projecting a Goal of communication. Unlike the lexical *manən*, it does not encode a separate production of speech event and does not introduce a proposition, that is, the linguistic material. This explains the interpretation of (27c): the sentence is felicitous even if the message is not passed on to the ultimate Goal, *təlanet* ‘you’ (see also footnote 6).

- (27) a. Mə-lan-na            tengeče    [kol-əm    kuč-aš    kaj-et]            man-ən-at.  
           we-DAT-POSS.1PLyesterday fish-ACC catch-INF go-NPST.2SG say-PST2-2SG  
           ‘You said to us yesterday that you would go fishing.’
- b. %[Rveze-vlak    kniga-m    už-ən-ət            manən]    saj.  
           boy-PL            book-ACC    see-PST2-3PL    COMP            good  
           ‘That the boys have seen the book is good.’
- c. Maša    mə-lam            [tə-lan-et<sub>i</sub>    [PRO<sub>i</sub>    to1-aš]            manən]    kalas-en.  
           Maša    I-DAT.1SG            you-DAT-POSS.2SG come-INF    COMP            tell-PST2  
           ‘Maša told me for you to come.’

In the next section I will discuss properties of the complementizer *manən* in more detail.

### 3.2 The complementizer *manən*

That *manən* appears in embedded indicative, subjunctive, and non-finite<sup>10</sup> clauses has first been reported in Isanbaev (1961), Galkin (1964), i.a., where *manən* is described as a subordination marker, without a formal analysis. Morphologically *manən* is identical to the non-agreeing converb/PST.3SG form of the verb *manaš* ‘say, tell’, which can still occasionally be used as a lexical predicate (27a).

*Manən* as a complementizer is desemanticized. It is not confined to speech act contexts and also appears in subject clauses (27b) and in complement clauses embedded under mental and emotive predicates, such as ‘believe’, ‘know’, or ‘be afraid’ (28a). Furthermore, it can be used in adjunct purpose clauses where it clearly does not contribute any ‘speech act’ semantics (28b).

(28) a. Iza      üšan-a                      [šüžar-že      ok      šojəšt    manən].

brother believe-NPST.3SG sister-POSS.3SGNEG.3SG lie      COMP

‘The brother believes that his sister will not lie to him.’

b. [Rveze-vlak    pur-əšt      manən]    me      kapka-m    poč-ən-na.

boy-PL      enter-JUS.PL COMP      we      gate-ACC    open-PST2-1PL

‘We opened the gate so that the boys could enter.’

Although in the examples in (28) *manən* is clearly used as a functional item, it may still be tempting to analyze the double-dative construction with a matrix speech act/mandative verb as involving a lexical *manən*, that is, a converb that embeds a non-finite clause; see also Abe

---

<sup>10</sup> As for non-finite clauses, *manən* is allowed only in complements of verbs of information transfer, which are discussed in this paper, and in purpose adjuncts (likely under the influence of Russian). The latter can either be infinitival or subjunctive; in both cases they are fully saturated and contain an overt DP/*pro* subject.

(Koopman & Sportiche 1989) and Turkish (Özyıldız et al. 2018), i.a., for analyses of ‘say’-based complementizers as Vs. In Mari, converbs either head adjunct clauses or appear in verb serialization. Under a converb analysis, the example in (27c) is close in meaning to ‘Mary told me something, saying to you to come’. The following facts, however, provide evidence that *manən* in single- and double-dative sentences in Mari is a C head and not a fully lexical item. Let us first consider a converbial adjunct analysis; an approach along this line, whereby an agreeing ‘say’-based complementizer is considered to be a V whose extended projection is a clausal adjunct, has recently been put forward for Lubukusu (Bantu) by Major et al. (2021). In Mari, however, it finds no support. The embedded clauses with *manən* under consideration behave as complements of the main speech act verb. Thus, they cannot co-occur with an internal DP argument, such as ‘fact’ or ‘joke’ (29), as this would violate the Theta Criterion.

- (29) a. Me Petja-lan tidə məskara-m kalas-en-na.  
 we Petja-DAT this joke-ACC tell-PST2-1PL  
 ‘We told Petja this joke.’
- b. \*Me Petja-lan tidə məskara-m [**tud-lan** tol-aš man-ən] kalas-en-na.  
 we Petja-DAT this joke-ACC s/he-DAT come-INF say-CVB tell-PST2-1PL  
 Intended: ‘We told Petja this joke, saying to her/him to come.’

In addition to this, unlike clausal adjuncts, clausal complements allow sub-extraction (30).<sup>11</sup>

---

<sup>11</sup> Sub-extraction out of an adjunct clause is occasionally possible. To account for the contrast between such English examples as \**What<sub>i</sub> did John drive Mary crazy [before reading *t<sub>i</sub>]*?*

and *What<sub>i</sub> did John drive Mary crazy [whistling *t<sub>i</sub>]*?*

Truswell (2007) proposes that ‘if the event denoted by the adjunct occupies an event position in the argument structure encoded in the

- (30) a. Nuno mə-lan-na [kō-m šel-aš manən] kalas-en-ət?  
 they we-DAT-POSS.1PL who-ACC hit-INF COMP tell-PST2-3PL  
 ‘Who did they tell us to hit?’
- b. Kō-m<sub>i</sub> nuno mə-lan-na [t<sub>i</sub> šel-aš manən] kalas-en-ət?  
 who-ACC they we-DAT-POSS.1PL hit-INF COMP tell-PST2-3PL  
 ‘Who did they tell us to hit?’
- c. [Kō-m šel-ən] me kaj-əš-na?  
 who-ACC hit-CVB we go-PST-1PL  
 ‘Who did we leave having hit?’
- d. ??Kō-m<sub>i</sub> me [t<sub>i</sub> šel-ən] kaj-əš-na?  
 who-ACC we hit-CVB go-PST-1PL  
 Intended: ‘Who did we leave having hit?’

Note also that the morphological form of *manən* in speech act verb constructions is fixed. For instance, a negative *-de* form can be derived out of a converb (31a); however, *-de* forms are never used in the sentences under discussion (31b).<sup>12</sup>

---

matrix verb, then extraction of the complement from within that adjunct is possible’. While it is possible to imagine that in the Mari sentences under discussion the two speech-events (the main one and the one expressed by *manən*) are grouped together, the fact/joke examples in (29) remain problematic for such an approach. I am grateful to Michael Diercks for pointing this problem out to me.

<sup>12</sup> Compare this, for instance, to the behavior of the say-type complementizer *le* in Kipsigis, which, as argued by Driemel & Kouneli (2020), should be analyzed as a lexical verb: it is

- (31) a. Maša        salam-əm        kalas-əde /        man-de        pur-əš.  
           Maša        hello-ACC        tell-CVB.NEG tell-CVB.NEG enter-PST  
           ‘Maša entered without saying hello.’
- b. \*Maša    t-lat            [təšeč    kaj-aš /    kaj-Ø    man-de]        kalas-en.  
           Masa    you-DAT.2SG here.EL    go-INF    go-IMP    tell-CVB.NEG tell-PST2  
           Intended: ‘Maša told you not to leave.’ or ‘Maša did not tell you to leave.’

Another option for a converb in Mari is to be a part of a serial verb construction; thus, the single- and double-dative construction might be interpreted as V.CVB (*manən*) + V.FINITE serialization. Recent proposals to analyze sentences with a ‘say’-based complementizer in terms of serialization include, for instance, Major & Torrence (2020) on Avatime (Niger-Congo), who argue for VP chaining. Again, such an approach is challenged by Mari data.

First, although at first glance the sentences with a main speech act verb seem to match the general serialization pattern in Mari – [infinitive] + *manən* + V.FINITE versus V.CVB + V.FINITE – the two constructions have different interpretations. In the latter case, the converb contributes the main lexical meaning, while the finite verb (selected from a small closed set) is desemantized and often serves as an aspectual marker (32). In the sentences with *manən*

---

inflected in the subjunctive mood and can be marked for Aspect. On the other hand, the say-type complementizer *-li* in Lubukusu cannot take applicative morphology and is restricted when it comes to Tense, Aspect, etc.; Major et al. (2021) analyze it as a light verb and propose that its limited distribution is precisely due to its lexical contribution being rather weak. Still, *-li* regularly appears with an agreement marker, while the complementizer *manən* in Mari does not allow any morphological modification.

under discussion, on contrary, it is *manən* that is grammaticalized, with the finite verb as the main predicate.

- (32) Petja kapka-ž-əm                      tükəl-en      šənd-əš.  
 Petja gate-POSS.3SG-ACC close-CVB sit-PST  
 ‘Petja has closed his gate.’

Second, serialization in Mari requires adjacency in that the converbial constituent cannot be dislocated (33). As was shown in Section 2, examples (16) and (17), the restriction does not apply to the double-dative sentences under discussion: the *manən* constituent can be extraposed and isolated in fragment answers.<sup>13</sup>

---

<sup>13</sup> Note also that both in Lubukusu and in Avatime, mentioned above, the presence of two lexical verbs of communication is brought by the division of labor: the main speech act verb on its own cannot introduce the linguistic material (i.e. what is being said) and for that a second verb is needed (Major & Torrence 2020; Major et al. 2021). This is not the case in Mari: as shown below, ‘tell’ (and other verbs of communication) do not require the presence of *manən* when the linguistic material is present.

- (i) a. Petja Maša-lan      zabor      törlə-mə-m      keles-en.      [Hill Mari]  
 Petja Maša-DAT fence fix-NMZ-ACC tell-PST2  
 ‘Petja told Maša that the fence had been fixed.’
- b. Maša ävä-žə-län                      zabor törlə-mə gišän keles-en.      [Hill Mari]  
 Maša mother-POSS.3SG-DAT fence fix-NMZ about tell-PST2  
 ‘Maša told her mother about fixing the fence.’

(33) #Petja šənd-əš [kapka-ž-əm tükəl-en].

Petja sit-PST gate-POSS.3SG-ACC close-CVB

Not available: ‘Petja has closed his gate.’

Only: ‘Petja sat, having closed the gate.’ (converbial adjunct)

It is further important to mention that *manən* used as a complementizer cannot be substituted by a converb form of a synonymous speech act verb. This is not expected if *manən* is a converb of a lexical predicate, since the verbs *manaš*, *kalasaš*, *kutəraš* ‘tell, say’ have similar semantic and syntactic distribution.

(34) \*Təj mə-lan-na təšeč kaj-aš kutər-en kalas-əš-əč.

you we-DAT-POSS.1PL here.EL go-INF say-CVB tell-PST-2SG

Intended: ‘You told us to leave.’

Taking these data into account, I argue that *manən* is being grammaticalized as a functional element. Its grammaticalization has not been complete yet: in speech act contexts it retains some properties of the lexical verb *manaš*, such as the ability to combine with a non-finite clausal complement and to license the Goal argument.<sup>14</sup>

---

<sup>14</sup> The general question remains about the source of dative case in Mari. On the one hand, it can be analyzed as an inherent case assigned to all Goals of communication. On the other hand, at least in some varieties of Mari, overt dative subjects are available in infinitival purpose clauses with the complementizer *manən* (i), which might suggest that dative is a structural case assigned by embedded T; see a detailed discussion of Mari purpose clauses and dative subjects

## 4 Deriving single- and double-dative constructions

### 4.1 Logophoric control

When developing a formal analysis for the constructions discussed in the paper, I adopt Landau’s (2015) predicative versus logophoric control distinction. Analyzing obligatory control constructions, Landau divides the predicates that embed clauses with a controlled PRO into two groups – nonattitude versus attitude, – depending on whether they select a nonattitude or an attitude complement, respectively. The difference between the two is that a definite description can be substituted *salva veritate* inside a nonattitude context, but not inside an attitude context. Consider the examples in (35): if Ralph is the new boss, but Bill does not know that, the ‘new boss’ alternative in (35a) is felicitous, however, in (35b) it is not acceptable.

- (35) a. Bill started to talk to Ralph. – aspectual, nonattitude  
       ⇒ Bill started to talk to the new boss.
- b. Bill planned to meet Ralph soon. – desiderative, attitude  
       ⇏ Bill planned to meet the new boss soon.

Similarly to their translation equivalents in English, speech-act verbs used as mandatives in Mari fall into the attitude category. When it comes to the syntactic structure, Landau further

---

in Burukina (2021). Both approaches could explain why the DP<sub>DAT2</sub> in double-dative sentences is marked dative and more data are required to decide between them.

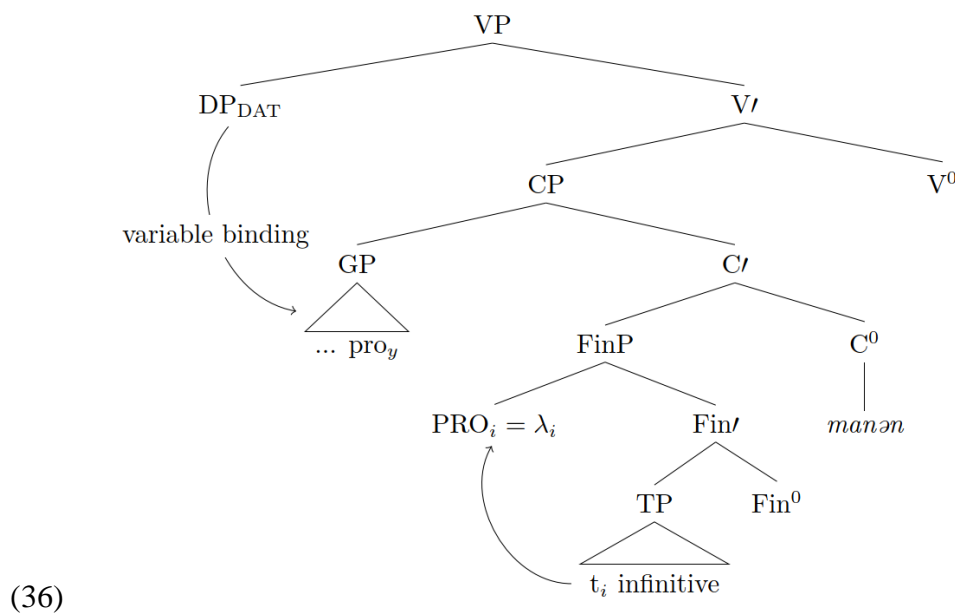
- (i) %[Kogəl’-lan kü-aš manən] me duxovkə-m čükt-əš-na.  
       pie-DAT cook-INF COMP we oven-ACC turn.on-PST-1PL

‘We turned on the oven for the pie to cook.’



proposes that attitude complements are structurally larger than nonattitude complements. The latter are FinPs predicated directly of the controller; hence, the term *predicative control*. The former, on the other hand, are propositional and must include AUTHOR (SPEAKER), ADDRESSEE, TIME, and WORLD coordinates that relate them to the context and allow for the attitude interpretation. (At least some of) these coordinates are syntactically projected within the CP layer added on top of the FinP; hence the term *logophoric control*.

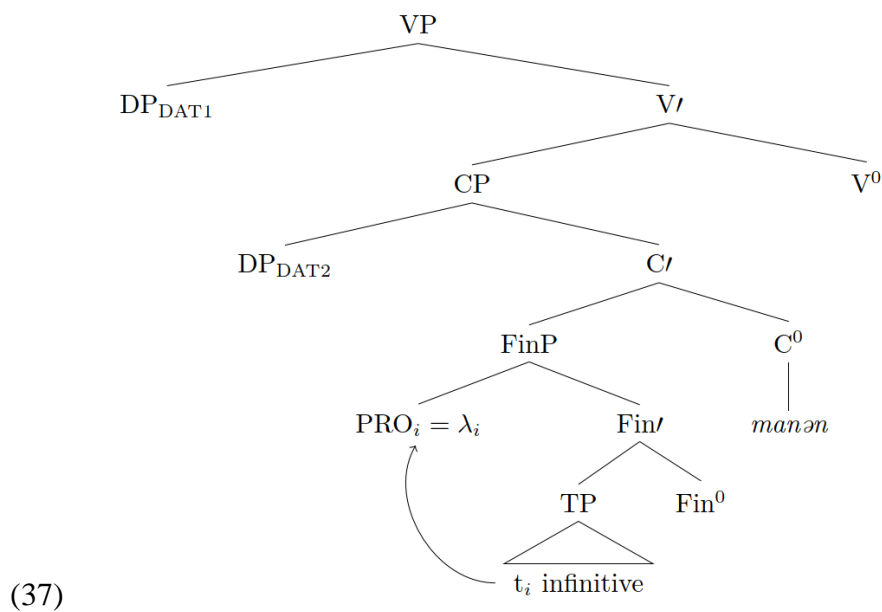
To derive the structure of the single- and double-dative sentences under consideration, I adapt Landau’s (2015) logophoric control analysis, elaborating the basic structure for sentences with a verb of communication given in (5). The following structure corresponds to sentences with a single DP<sub>DAT</sub> (36).



A crucial component of the structure in (36) is the concept generator phrase (GP), which can introduce the required AUTHOR (SPEAKER), ADDRESSEE, TIME, and WORLD coordinates, with their values being determined by elements in the matrix clause; see Baker (2008), Zanuttini (2008), Diercks (2013), i.a., on the structural presence of discourse-related operators. In

sentences with logophoric control, the AUTHOR and ADDRESSEE coordinates mediate the relation between the matrix controller and the embedded PRO. In object control sentences the ADDRESSEE coordinate is syntactically projected as *pro<sub>y</sub>* bound by the matrix object, typically a Goal of communication. *Pro<sub>y</sub>* further values the embedded PRO variable via predication, where *pro<sub>y</sub>* is the subject and the embedded FinP is turned into a predicate via the operator movement of PRO to Spec,FinP (see Williams 1980, Hendrick 1988, Clark 1990, i.a.).

I propose to derive the structural representation for double-dative sentences from the structure in (36). The only difference between the two is that in (37) the embedded C head projects the DP<sub>DAT2</sub> – that is, the ultimate Goal/obligation holder.



Being a referential expression, the DP<sub>DAT2</sub> in (37) cannot be bound by the matrix Goal. Aside from that, it essentially plays the role of *pro<sub>y</sub>* in (37). In this configuration *manən* selects a property-type FinP (<e,<s,t>>), just like other obligatory control C heads (Landau 2015), and,

similarly to verbs of speech, projects a Goal/Addressee argument, instead of forcing one of the coordinates to project as Spec,CP.<sup>15</sup>

#### 4.2 Additional support for the proposed analysis

The analysis presented in Section 4.1 makes several predictions regarding the distribution of double datives that are borne out. First, in control sentences with a single DP<sub>DAT</sub> a flexible binding relation is established between the DP<sub>DAT</sub> (the controller) and *pro*<sub>y</sub> in the embedded clause, and we expect partial coreference between the controller and *pro*<sub>y</sub> (and consequently PRO) to be allowed. In contrast, in sentences with two dative DPs the relation between the DP<sub>DAT2</sub> and RPO is that of predication, which only leaves a possibility for exhaustive control. This is observed in Mari: partial/split control is allowed in single-dative sentences (38a), but in double-dative sentences strict coreference between the controller and PRO is required (38b).<sup>16</sup>

---

<sup>15</sup> I assume that the Goal DP projected by *manən* serves as the subject for the complex predicate formed by *manən* and the FinP; cf. Pytkänen's (2008) analysis for object depictives of the type *Sue saw Peter tired* whereby a depictive secondary predicate is merged into the complement position of a transitive V head combining directly with the latter via Predicate Modification, as both are of the type <e,<s,t>>. Alternatively, the Komposition operation, put forward by Williams (2005) for resultatives, such as *John pounded the cutlet flat* and *Mo sang her throat hoarse*, can be applied. I leave the in-depth examination of the semantics of the double-dative construction for future research.

<sup>16</sup> Similarly to its translation equivalent in English, the modifier *pərl'a* 'together' must be linked to a semantically plural nominal phrase (i). For (38a) to be grammatical, the embedded PRO subject must be plural, even though the controller is singular.

(i) a. Me təšec̣ pərl'a kaj-en-na.

- (38) a. Maša<sub>k</sub> t-lat<sub>i</sub> [PRO<sub>i+(k)</sub> təšeč pərl'a kaj-aš manən] kalas-en.  
 Maša you-DAT.2SG here.EL together go-INF COMP tell-PST2  
 'Maša told you (sg) to leave together with her/someone.'
- b. \*Maša<sub>k</sub>mə-lam t-lat<sub>i</sub>[PRO<sub>i+(k)</sub> təšeč pərl'a kaj-aš manən] kalas-en.  
 Maša I-DAT.1SG you-DAT.2SG here.EL together go-INF COMP tell-PST2  
 Intended: 'Maša told me for you (sg) to leave together.'

Second, for the embedded FinP to be predicated of the embedded *pro*<sub>y</sub> or DP<sub>DAT2</sub> it must contain a PRO variable (36, 37) and double-dative sentences with fully saturated finite FinPs are expected to be ungrammatical. This is true for Mari: as shown in (39), double-datives are incompatible with finite embedded clauses even though, in principle, speech act verbs can select finite subjunctive/indicative clausal complements.

- (39) a. Maša t-lat (\*Petja-lan) [rveze-vlak kniga-m uš-əšt manən] kalas-en.  
 Maša you-DAT.2SG Petja-DAT boy-PL book-ACC see-JUS.PL COMP tell-PST2  
 'Maša told you that the boys should see the book.'
- b. Maša t-lat [\*Petja-lan / Petja tol-žo manən] kalas-en.

---

we here.EL together go-PST2-1PL

'We left together.'

- b. Məj təšeč (\*pərl'a) kaj-en-am.  
 I here.EL together go-PST2-1SG  
 'I left.'

Maša you-DAT.2SG      Petja-DAT    Petja    come-JUS COMP    tell-PST2

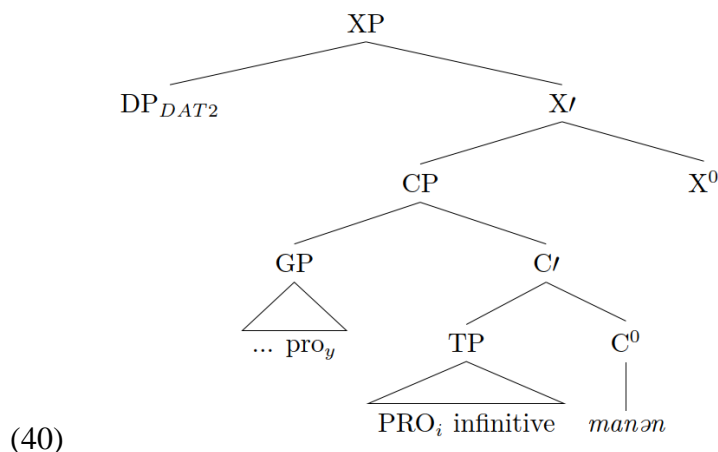
‘Maša told you that Petja should come.’

These properties of double-dative constructions are straightforwardly accounted for by the proposed analysis and posit a problem for alternative silent predicate approaches, as will be discussed in Section 4.3.

### 4.3 Against alternative silent predicate analyses

Instead of placing the  $DP_{DAT2}$  immediately in Spec,CP it might be suggested that it is one of the arguments of a silent dyadic predicate ( $X^0$ ), the second one being the embedded clause.

This is schematized in (40); the XP is later selected by the matrix speech act verb.



A plausible candidate for  $X^0$  would be a silent deontic modal, a counterpart of lexical modal verbs *kūlaš* ‘be necessary’ and *liaš* ‘be allowed’. The latter require a dative obligation/permission holder and an embedded clause; the Holder obligatorily controls the embedded PRO subject.

- (41) Ač’a-ž-lan<sub>i</sub>                      [PRO<sub>i</sub> təšeč      kaj-aš]                      kül-eš.  
 father-POSS.3SG-DAT                      here.EL    go-INF                      be.necessary-NPST.3SG

‘For his/her father it is necessary to leave.’

However, the silent modal analysis is challenged by the following empirical observations. First, as discussed in Section 2, the DP<sub>DAT2</sub> is a Goal and must comply with the [+Human] restriction; most importantly it must refer to someone conscious who can potentially receive the message. In contrast, the standard obligation holder restriction is merely [+Animate] (42).

- (42) a. Uškal-lan<sub>i</sub> [PRO<sub>i</sub> šudə-m kočk-aš] kül-eš.  
 cow-DAT grass-ACC eat-INF be.necessary-NPST.3SG  
 ‘For the cow it is necessary to eat grass.’
- b. \*Krovat’-vlak-lan pərd-əž vokten šog-aš kül-eš.  
 bed-PL-DAT wall near stand-INF be.necessary-NPST.3SG  
 Intended: ‘For the beds it is necessary to be near to the wall.’

Second, examples with a matrix deontic modal and an embedded non-finite clause with *manən* are ungrammatical (43).

- (43) \*Ač’a-ž-lan<sub>i</sub> [PRO<sub>i</sub> təšeč kaj-aš manən] kül-eš.  
 father-POSS.3SG-DAT here.EL go-INF COMP be.necessary-NPST.3SG  
 ‘For his/her father it is necessary to leave.’

Third, *kūlaš* and *liaš* are compatible with finite subjunctive clauses (44).

- (44) [Kogəl’o vaškerak kü-žo manən] mə-lan-na kül-eš.  
 pie quickly cook-JUS COMP we-DAT-POSS.1PL be.necessary-NPST.3SG

‘It is necessary for us for the pie to cook quickly.’

A reasonable assumption would be that a silent modal is present in sentences with a matrix speech act verb whenever a mandative interpretation appears, i.e. when a non-finite or a subjunctive clause is embedded. However, recall from Section 4.2 that double datives are incompatible with finite complements. Thus, to adopt (40), we would have to dissociate the presence of a modal and the availability of a mandative reading. We would further have to stipulate that a silent deontic modal is strikingly different in its distribution from its overt counterparts: unlike those, it should select non-finite clauses with *manən* and impose a typologically unusual [+Human] and [conscious] restriction on the Obligation Holder. Conversely, the analysis outlined in this paper does not require any such accommodation and straightforwardly captures all the relevant properties.

## **5 Implications and concluding remarks**

The paper attempted to demonstrate that in Mari in sentences with a matrix speech act verb and a non-finite clausal complement the embedded C head is capable of thematically licensing an argument projected in Spec,CP. This becomes possible due to the semi-grammaticalized status of the complementizer: it is a hybrid derived from a verb ‘say’ that already has a distribution of a functional item but still retains some properties of a lexical predicate. I further adopted Landau’s (2015) logophoric control analysis to provide a structural representation for the constructions under discussion that fully accounts for their properties.

Evidence that Spec,CP can be an A-position comes from many languages where cross-clausal A-dependencies are possible in embedded finite clauses over a CP boundary. Examples include ECM into finite clausal complements in Japanese (Tanaka 2002; Horn 2008) and hyperraising in Mongolian (Fong 2019), to name a few; see Wurmbrand (2019) for a general discussion. In all these cases Spec,CP serves as a landing site to which an embedded argument can move to

be further probed by a matrix head or the C head itself. The present paper expands this discussion by showing that Spec,CP as an A-position can be suitable not only for internal merge but also for external merge and theta-role assignment.

The proposed analysis opens up several directions for future research. First, it is worth looking at other languages where speech act verbs are being grammaticalized into complementizers. The question is whether the so-called semi-complementizers found, for instance, in Cantonese Yue and Hakka (Chappell 2008) exhibit mixed behavior similar to that of *manən* and can be analyzed as complementizers with ‘remnant’ lexical properties. Furthermore, as mentioned in Section 3, in many languages ‘say’-based complementizers are being re-analyzed as V heads; see Major & Torrence (2020), Driemel & Kouneli (2020), Major et al. (2021). As I showed, such an approach is untenable if applied to the Mari data; however, comparison of Mari with the languages of Africa raises an important question about describing an element as semi-grammaticalized (a C with some properties of V) or semi-lexicalized (a light V) and whether there is a fine line between the two notions.

Second, the assumption that there are several coexisting lexical entries for *manən* used in different environments – the semi-complementizer appears in so-called speech act contexts and the fully grammaticalized *manən* heads indicative and subjunctive clausal complements, subjects and adjuncts – leaves open the following question: How could this distribution be explained? Although I do not have an immediate answer, one way to address the problem would be to consider in more detail semantic properties of matrix predicates focusing on the attitude/nonattitude distinction. Proposals along this line have been made by Yoon (2007), Horn (2008), Wurmbrand (2019), Wurmbrand & Lohninger (2020) to account for the (im)possibility of ECM in Korean and Japanese. Thus, Yoon (2007) suggests that the exceptionally marked DP moved to an A-position at the very left periphery of the embedded clause enters a predication relation with the embedded clause itself. Horn (2008: 7/186) shares



this intuition and proposes that ECM is allowed when “the proposition expressed by an accusative-quotative complement [is] a property ascription on the referent of the accusative subject when evaluated with respect to the belief world of the agent of attitude.” The C head that behaves exceptionally in mandative sentences in Mari, i.e. in attitude contexts, appears to follow a similar pattern.

Third, as noted at the end of Section 4.1, in double-dative sentences the DP<sub>DAT2</sub>, i.e. the Goal argument projected by *manən*, plays the role of *pro<sub>y</sub>* in single-dative sentences – a syntactically projected ADDRESSEE coordinate. I tentatively propose that the DP<sub>DAT2</sub> is an overtly introduced ADDRESSEE itself, in the spirit of Baker (2008: 125): “All matrix clauses and certain embedded clauses have two special null arguments generated within the CP projection, one designated S (for SPEAKER) and the other A (for ADDRESSEE).” Taking the Mari data into account, Baker’s proposal can be elaborated to include exceptional cases when a discourse-oriented argument is overtly realized as an independent DP projected by the complementizer. A similar idea has recently been put forward by Spadine (2018) for English and Spadine (2019) for Tigrinya. In English, the syntactic presence of (covert) speaker and addressee in the left periphery of a clause is indicated by control into speech act modifying adjuncts, as in *Mary<sub>i</sub> told John<sub>j</sub> that, PRO<sub>i/\*j</sub> as a film critic, this movie deserves an Oscar*. In Tigrinya a special functional head *?il* arguably introduces an overt perspectival nominal at the left edge of a matrix or embedded clause. Unlike Baker (2008) and following Speas (2004), Sundaresan (2018), i.a., Spadine separates the speaker/addressee containing saP in English and the *?il*-headed Perspectival Projection in Tigrinya from CP. Thus, another direction for future research is to examine in more detail the (covert) left periphery in Mari to bring together various existing approaches to structurally present discourse-oriented items.

### **Data-availability statement**

The full original data generated by this study (i.e. the elicitation material) are available from the author upon request.

### **Acknowledgments**

I am immensely grateful to my Mari consultants for their judgments, especially to Elena Vedernikova and Tatiana Jefremova. I would like to thank Katalin É. Kiss, Marcel den Dikken, and my colleagues at the Hungarian Research Centre for Linguistics ELKH for their help with the research. Parts of the research were presented at *SICOGG22*, *NELS51*, *LSA95* and at a *SemanticsBabble* meeting (UC San Diego) and I am very thankful to the audiences for their feedback; a special thank you to Susi Wurmbrand and Idan Landau for their interest in the project. I would also like to thank Michael Diercks and an anonymous reviewer for their helpful comments on the manuscript. All mistakes are mine. The research was supported by the ÚNKP-21-4 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund and by the Hungarian National Research, Development and Innovation Office under the grant NKFI 129921.

### **References**

- Baker, Mark. 2008. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.
- Burukina, Irina. 2020. Mandative verbs and deontic modals in Russian: Between obligatory control and overt embedded subjects. *Glossa: A Journal of General Linguistics* 5.1. 54.
- Burukina, Irina. 2021. On P licensing dative subjects and AGR in infinitival clauses. Unpublished paper. Work conducted at: Hungarian Research Centre for Linguistics and Eötvös Loránd University.
- Cavirani-Pots, Cora. 2020. *Roots in progress: Semi-lexicality in the Dutch and Afrikaans verbal domain*. Doctoral thesis. Leuven, Belgium: KU Leuven.

- Chappell, Hilary. 2008. Variation in the grammaticalization of complementizers from *verba dicendi* in Sinitic languages. *Linguistic Typology* 12.1. 45–98.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000 Minimalist Inquiries: The framework. In: Roger Martin, David Michaels, Juan Uriagereka, and Samuel Jay Keyser, editors. *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: MIT Press. 89–155.
- Clark, Robin. 1990. *Thematic theory in syntax and interpretation*. London: Routledge.
- Daniel, Michael. 2007. Reported illocution: Data from several Daghestanian languages. Unpublished paper. Work conducted at: Moscow State University.
- Demirok, Omer, Deniz Özyıldız, and Balkiz Öztürk. 2020. Complementizers with attitude. In: Maggie Baird and Jonathan Pesetsky, editors. *Proceedings of the 49th annual meeting of the North East Linguistics Society*. Amherst, MA: GLSA.
- Diercks, Michael. 2013. Indirect agree in Lubukusu complementizer agreement. *Natural Language and Linguistic Theory* 31. 357–407.
- Driemel, Imke, and Maria Kouneli. 2020. ‘Say’-based complementation: Insights from Kipsigis. Unpublished paper. Work conducted at: Leipzig University.
- Fong, Suzana. 2019. Proper movement through Spec-CP: An argument from hyperraising in Mongolian. *Glossa: A Journal of General Linguistics* 4.1. 30.
- Galkin, I. S. 1964. *Istoricheskaya grammatika mariiskogo yazyka. Morfologija* [Historical grammar of Mari. Morphology]. Ioshkar-Ola: Mariiskoe Knizhnoe Izd.
- Heine, Bernd, and Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Hendrick, Randall. 1988. *Anaphora in Celtic and Universal Grammar*. Dordrecht: Kluwer.

Hock, Hans H. 1982. The Sanskrit quotative: A historical and comparative study. *Studies in the Linguistic Sciences* 12. 37–96.

Horn, Stephen W. 2008. *Syntax, semantics, and pragmatics of accusative-quotative constructions in Japanese*. Doctoral thesis. Columbus, OH: Ohio State University.

Isanbaev, N. I. 1961. Osobennosti upotreblenija manən i lijən v marijskom jazyke [Peculiarities of the use of manən and lijən in Mari.] In *Trudy Mariiskogo nauchno-issledovatel'skogo instituta yazyka, literatury i istorii* [Works of the Mari research institute of language, literature, and history] 15.

Klamer, Marian. 2000. How report verbs become quote markers and complementizers. *Lingua* 110. 69–98.

Koopman, Hilda, and Dominique Sportiche. 1989. Pronouns, logical variables, and logophoricity in Abe. *Linguistic Inquiry* 20.4. 555–588.

Landau, Idan. 2015. *A two-tiered theory of control*. Cambridge, MA: MIT Press.

Major, Travis, Justine Sikuku, and Michael Diercks. 2021. Agreeing complementizers are neither complementizers nor selected: Reanalyzing Bukusu complementizer agreement. Unpublished paper. Work conducted at: Pomona College and University of California, Los Angeles.

Major, Travis, and Harold Torrence. 2020. Say-chains, not say-complementation. To appear in the *Proceedings of the 38th West Coast Conference on Formal Linguistics*.

Özyıldız, Deniz, Travis Major, and Emar Maier. 2018. Communicative reception reports as hearsay: Evidence from indexical shift in Turkish. In: Richard Stockwell, Maura O'Leary, Zhongshi Xu & Z. L. Zhou, editors. *Proceedings of the 36th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings Project. 296–305.

Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge, MA: MIT Press.

- Savatkova, A. A. 2002. *Gornoe narechie mariyskogo jazyka* [Hill dialect of Mari language]. Savariae: Berzsényi Dániel Főiskola.
- Serdobolskaya, N. V., and S. Yu. Toldova. 2011. Grammaticalization of the verb of speech in Finno-Ugric languages. In: *Proceedings of the Congressus XI internationalis fenno-ugristarum*. Budapest: Reguly Társaság.
- Spadine, Carolyn. 2018. Control in illocutionary adjuncts as a diagnostic for discourse arguments. Unpublished paper. Work conducted at: Massachusetts Institute of Technology.
- Spadine, Carolyn. 2019. The Syntax of attitude holders: Evidence from Tigrinya. Unpublished paper. Work conducted at: Massachusetts Institute of Technology.
- Speas, Margaret. 2004. Evidentiality, logophoricity and the syntactic representation of pragmatic features. *Lingua* 114.3. 255–276.
- Sundaresan, Sandhya. 2018. Perspective is syntactic: Evidence from anaphora. *Glossa: A Journal of General Linguistics* 3.1. 128.
- Tanaka, Hidekazu. 2002. Raising to object out of CP. *Linguistic Inquiry* 33. 637–652.
- Toldova S. Yu., & N. V. Serdobolskaya. 2014. Glagol reči manaš v marijskom jazyke: osobennosti grammatikalizacii [The verb manaš in Mari: grammaticalization]. *Voprosy Jazykoznanija* 6. 66–91.
- Tóth, Ildikó. 2000. *Inflected infinitives in Hungarian*. Tilburg: Tilburg University dissertation.
- Truswell, Robert. 2007. Extraction from adjuncts and the structure of events. *Lingua* 117.8. 1355–1377.
- Williams, Edwin. 1980. Predication. *Linguistic Inquiry* 11:203–238.
- Williams, Alexander. 2005. *Arguments in Syntax and Semantics*. Cambridge: Cambridge University Press.

Wurmbrand, Susi. 2019. Cross-clausal A-dependencies. In: Eszter Ronai, Laura Stigliano and Yen-an Sun, editors. *Proceedings of the Fifty-Fourth Annual Meeting of the Chicago Linguistic Society*. Chicago: University of Chicago, Chicago Linguistic Society. 585–604.

Wurmbrand, Susi, and Magdalena Lohninger. 2020. Cross-clausal A-dependencies. Unpublished paper. Work conducted at: University of Vienna.

Yoon, James H. 2007. Raising of major arguments in Korean and Japanese. *Natural Language and Linguistic Theory* 25. 615–653.

Zanuttini, Raffaella. 2008. Encoding the addressee in the syntax: Evidence from English imperative subjects. *Natural Language and Linguistic Theory* 26. 185–218.