

# Grammaticalization of the verb of speech in Finno-Ugric languages<sup>1</sup>

Serdobolskaya Natalia, Toldova Svetlana

Russian State University for the Humanities, Moscow

## 1. Mari and Udmurt complementizers derived from the verbs of speech

It is well known that speech verbs in languages of the world can grammaticalize giving rise to complementizers (Lord 1976; Klamer 2000; Güldemann, von Roncador 2002). In this latter function, they are used in contexts that do not presuppose any speech situation, i.e. with mental verbs, emotional verbs, etc. For example, in Eastern Mari, the verb *manaš* (see (1) for the independent usage of this verb) in the form of converb is used as a complementizer that introduces indirect speech (2). Moreover, the same form is attested introducing subordinate clauses of mental and emotional verbs, as in (3). In that case, *manən* is desemanticized, since it does not denote a speech situation, hence, presenting an example of grammaticalized usage.

MARI (EASTERN)

- (1) *kugu-rak-še*      **man-eš**      "təj      ajda      ončal"  
big-COMP<sup>2</sup>-P.3SG    say-PRS.3SG      you      come.on      look.IMP

*The elder brother says: "You go and have a look" <...>.*

- (2) *üdər*      *ava-že*      *tud-əm*      *molo*      *joča*      *deč*  
girl      mother-P.3SG      dem-ACC      other      child      from  
*čot-rak*      *jörat-a*      **man-ən**      *moktan-en.*  
very-COMP      love-PRS.3SG      say-CONV      boast-PST.3SG

*The girl boasted that her mother loves her more than other children.*

- (3) *jəvan*    *ola-ške*    *kaj-em*      **man-ən**      *šon-a.*  
Ivan      city-LAT    go-PRS.1SG    say-CONV      think-PRS.3SG

*Ivan thinks that he will go to the city.*

The same phenomenon is observed in Besermyan variant of Udmurt language with the converb of the verb *šuanə* 'to say':

UDMURT (BESERMEN)

- (4) *mar*    *pe*    *ta?*    –    *ž'ič'ə*    **šu-e.**  
what    CIT    DEM    fox    say-PRS.3SG

*"What is this?" – The fox says.*

- (5) *Vas'a*    *vera-z*      *brat-ez-lə,*      *soje*      *žug-o-z*      **šu-sa.**  
Vasja    tell-PST.3    brother-P.3SG-DAT    dem.ACC    beat-FUT-3    say-CONV

*Vasja told to his brother that he would beat him.*

- (6) *pi*      *č'akla-š'k-e*      *so*      *baš't-o-z*      *vit'*      **šuə-sa.**  
boy      think-DETR-PRS.3SG    DEM    get-FUT-3    five      say-CONV

*The boy thinks that he will get the mark "5".*

The described grammaticalization pattern is attested in many genetically non-related languages, as in Indo-European (Slavic, Indo-Aryan, Iranian), Uralic (Finno-Ugric: Mari, Udmurt), Altaic

<sup>1</sup> The work is supported by Russian Fund of Fundamental Research (grant № 10-06-00550a), whose aid is gratefully acknowledged.

<sup>2</sup> Glosses: SG – singular number, PL – plural number, ACC – accusative case, DAT – dative case, INS – instrumental case, LAT – lative case, ILL – illative case, COMP – comparative suffix, DEM – demonstrative pronoun; PRS – present, PST – past tense ('first past' in reference grammars), FUT – future tense, IMP – imperative mood, JUSS – jussive mood, OPT – optative mood, INF – infinitive, CONV – converb, DETR – detransitivizing suffix, CAUS – causative suffix, CIT – citative particle.

(Mongolic, Turkic: Kononov 1953), Malayo-Polynesian (Klamer 2000), Nakh-Daghestanian (Daniel 2007).

Both Finno-Ugric and Turkic languages have complement constructions, where the simple converb of the verb ‘to say’ is used as a complementizer: the form *manən* in Mari and *šuəsa* in Udmurt (Majtinskaja 1982: 96; Isanbaev 1961; Timofeeva 1961). Other genetically related Finno-Ugric languages (Komi, Mordvin e.a.) lack such complementizers. Constructions with the converb formed along the same morphological pattern are attested in the neighbouring Tatar and Chuvash languages (*dip* in Tatar, *tese* in Chuvash), as in many other Turkic languages. Therefore, it has been suggested that this grammaticalization pattern in Mari and Udmurt is due to contact influence from Tatar and Chuvash.

The question arises, whether the syntactic and semantic properties of the complementizer constructions in Finno-Ugric languages correspond to the properties of Tatar and Chuvash constructions. The aim of this paper is to answer this question. This would permit to refine the ‘contact’ hypothesis with the exact information on whether the whole construction is borrowed or rather the morphologic model of forming the complementizer.

The data discussed in this paper have been gathered during the fieldwork in Eastern Mari (village of Staryj Torjal), and Besermyan variant of Udmurt (village of Shamardan).

## 2. The grammaticalization scale of speech verbs

Grammaticalization of speech verbs into complementizers is attested in many languages of the world (see e.g. Harris and Campbell 1995, Hopper and Traugott 1993, Lord 1976, Lehmann 2002, Saxena 1995 e.a.). There are two possible paths of grammaticalization, a speech verb can be grammaticalized into a citation marker, or it is grammaticalized into a subordinate conjunction that introduces both complement clauses and adverbial clauses of cause, purpose, measure etc. The grammaticalization pattern found in Mari and Besermyan follows the second path.

The opposition between direct and indirect speech constructions in languages of the world is based on the following (after Toldova 1999; Toldova, Serdobolskaya 2006; Aikhenvald 2009). The direct/indirect speech constructions have two speech acts, the matrix clause denoting the ‘real’ speech act and the complement clause denoting the imaginary speech act. The two situations, the one introduced by the matrix clause and the one introduced by the dependent clause have different “coordinates”, namely, participants, temporal and local characteristics. These coordinates can be encoded along the following strategies of reference: deictic strategy (direct speech strategy), where the NPs (temporal/local characteristics) in the complement clause are coindexed with the imaginary speech act participants (as in *He said: “I was in China yesterday”*), see (7a); or anaphoric strategy (indirect speech strategy), where the NPs (temporal/local characteristics) in the complement clause are coindexed with the real speech act participants, or, in case of no coreference to the real speech act participants, encoded by anaphoric devices used in this language (as in *He said he had been to China the day before*), see (7b) (see Toldova 1999 for the interlacing of these strategies in colloquial speech).

(7)	a.	anaj-ez	vera-z	ataj-ez-lə	<b>mone</b>	kwaret-i-z
		mother-P.3SG	tell-PST.3	father-P.3SG-DAT	<b>I.ACC</b>	scold-PST-3
		kužəj-e	šuə-sa.			
		boss-P.1SG	say-CONV			
	b.	anaj-ez	vera-z	ataj-ez-lə	<b>soje</b>	kwaret-i-z
		mother-P.3SG	tell-PST.3	father-P.3SG-DAT	dem.ACC	scold-PST-3
		kužəj-ez	šuə-sa.			
		boss-P.3SG	say-CONV			

*The mother<sub>i</sub> said to the father that her<sub>i</sub> boss had scolded her<sub>i</sub> (a. ...that my<sub>i</sub> boss scolded me<sub>i</sub>. b. ...that her<sub>i</sub> boss scolded her<sub>i</sub>).*

As for temporal and locative adverbials, these parameters will not be discussed here, since no shift similar to English is observed in Finno-Ugric.

The choice between direct and anaphoric reference strategy manifests itself in the choice of the mood of the dependent verb by the verbs of causation, speech causation, or intention: imperative is used by deictic reference strategy (8a), and infinitive by anaphoric reference strategy (8b).

MARI

- (8) a. ača üdər-lan pört muš-šo man-ən küšt-en.  
 father girl-DAT house wash-JUSS say-CONV order-PST.3SG  
 b. ača üdər-lan pört musk-aš küšt-en.  
 father girl-DAT house wash-INF order-PST.3SG

*The father ordered the girl to clean the house.*

### 3. Mari and Besermyan complementizer constructions compared to Tatar

#### 3.1. Semantic shifting of the speech verb

The discussed verb in Mari, Besermyan, and Tatar has the meanings 'to say' (illustrated in section 1) and 'to name':

BESERMYAN

- (9) 14 janvarja “vuž vil’ ar” šui-š’ko-m.  
 14 January old new year say-PRS-1PL

*We call the 14th of January “Old New year”.*

The groups of matrix verbs that can host the constructions with converb of speech as a complementizer include the following: speech verbs (7), mental verbs (6) (where an imaginary 'inner' speech can be supposed to occur), emotion verbs (10), adverbial constructions (with the semantics of purpose (11) and reason (12)). See the examples from Besermyan (the same semantic shift is observed in Mari and Tatar):

- (10) mon jara-t-iš’ko [so d’eš’ mad’-e šuə-sa].  
 I love-CAUS-PRS he well sing-PRS.3SG say-CONV

*It pleases me that he sings well.*

- (11) jul-e – avgust-e tin’ turən dastiš’k-o-m n’i turən  
 July-ILL August-ILL dem hay prepare-FUT-1PL already hay  
 [život-lə tolalte med okm-o-z šu-sa].  
 cattle-DAT in.winter OPT be.enough-FUT-3 say-CONV

*In July-August we prepare the hay in order that it should be enough for the*

- (12) parnik-ez... kal’ uš’t-i-m val n’i [pəš’ šu-sa].  
 hotbed-ACC now open-PST-1PL be.PST already hot say-CONV

*We’ve opened the hotbed because it was [too] hot.*

The grammaticalization path of this form can be characterized with the following scheme:

verbs of speech → mental verbs with → emotion verbs with → verbs that introduce events  
 ↓ propositional semantics propositional semantics  
 speech causation → verbs of causation and intention → adverbial clauses of purpose and reason

#### 3.2. Syntax of the constructions with grammaticalized verbs of speech: pronouns reference strategy

In Besermyan and Mari, as well as Tatar, both strategies of participants encoding, deictic and anaphoric, are used with the complementizer derived from the verb of speech. However, some groups of matrix verbs show preference towards deictic or anaphoric strategy.



MARI

- (17) iza-že                      šüžar-žə-lan                      [maska-m  
elder.brother-P.3SG    younger.sister-P.3SG-DAT    bear-ACC  
pušt-ən-**am**            man-ən]            kalas-əš.  
kill-PST-**1SG**            say-CONV            tell-PST.3SG

*The brother<sub>i</sub> told to the sister that he<sub>i</sub> has killed the bear.*

- (18) izaže<sub>i</sub>                      šüžar-žə-lan                      [maska    **tud-əm<sub>i</sub>**  
elder.brother    younger.sister-P.3SG-DAT    bear            dem-ACC  
susərt-en                      man-ən]                      kalas-əš.  
wound-PST.3SG                      say-CONV                      tell-PST.3SG

*The brother<sub>i</sub> told to the sister that the bear had wounded him<sub>i</sub>.*

- (19) [iza-že                      pört-əm                      nal-eš                      man-ən]                      tudo            ojl-en.  
elder.brother-**P.3SG**    house-ACC    take-PRS.3SG    say-CONV    dem            tell-PST.3SG  
*The lad<sub>i</sub> said that his<sub>i</sub> brother would buy a house.*

The distribution of the factors relevant for the choice of the reference strategy in Eastern Mari is shown in the following table:

S 1 / IO 1 =	S 2	DO 2, IO 2, POSS 2
verbs of speech: 'say', 'tell'	<b>deictic</b> (anaphoric)	anaphoric / deictic
verbs of speech causation: 'demand', 'request'	<b>deictic</b> /anaphoric	anaphoric / deictic
mental verbs: 'think', 'know', 'believe'	anaphoric / deictic	anaphoric
verbs of emotion: 'rejoice', 'be angry'	anaphoric / deictic	anaphoric
verbs of causation: 'make', 'send'; verbs of intention: 'decide', 'want'	anaphoric / deictic	anaphoric
adverbial purposive clauses	anaphoric / deictic	anaphoric
adverbial clauses of reason	anaphoric / deictic	anaphoric

These results are totally different from Tatar. According to (Khanina 2007), in Tatar (Mishar dialect) the choice of the reference strategy depends on the syntactic construction used. The peculiarity of the complement clauses with the grammaticalized verb of speech in Tatar (as well as in other Turkic languages) is the possibility of encoding the subject of the complement clause with accusative case. Accusative subject constructions most often take the anaphoric strategy of participants' encoding, while nominative subject constructions take the deictic strategy.

- (20) sin            [**min**            bütän            kil-m-i-m                      di-p]                      at-t-eŋ.  
you            **I(NOM)**    another            come-NEG-ST.IPFV-1SG    say-CONV            say-PST-3SG  
*You said you wouldn't come again. (Khanina 2007: 132)*

- (21) alsu            [**min**ϣ            [**ul**            kit-tϣ]                      di-p]                      ujl-ϣj.  
Alsu            I.ACC            **dem.NOM**    leave-PST            say-CONV            think-ST  
*Alsu thinks that I have left (lit. thinks about me "He's left"). (Ibid.)*

S 1 / IO 1 =	S2 = nominative	S2 = accusative
--------------	-----------------	-----------------

Reference strategy in the complement clause	deictic	anaphoric
---	---------	-----------

(After Khanina 2000; 2003; 2007)

Hence, the distribution of the syntactic properties of the discussed constructions differs in Finno-Ugric languages when compared to Turkic languages. However, the semantic shifts observed in Mari, Besermyan, and Tatar are the same.

The conclusion then can be made that it is not only the morphological model of forming the complementizer that is due to areal influence, but also the semantic constraints on the constructions formed with this complementizer. The syntactic features of the discussed constructions, on the contrary, have probably arisen in the discussed languages independently.

## References

- Daniel M.A. 2007. Reported Illocution: data from several Daghestanian languages. Conference on the languages of the Caucasus, Leipzig, 7-9 December 2007.
- Galkin I.S. 1964. Istoricheskaja grammatika marijskogo jazyka. Morphologija. Vol. 1. Joshkar-Ola. Grammatika udmurtskogo jazyka. 1970. Sintaksis prostogo predlozhenija. Izhevsk: "Udmurtija". Grammatika udmurtskogo jazyka. 1974. Sintaksis slozhnogo predlozhenija. Izhevsk: "Udmurtija".
- Güldemann T. & von Roncador M. (eds.), 2002. Reported discourse. Amsterdam: John Benjamins.
- Hopper P.J. & Traugott E.C., 1993. Grammaticalization. Cambridge: Cambridge University Press.
- Isanbaev N.I. 1961. Osobennosti upotreblenija *manyn* i *lijyn* v marijskom jazyke // Trudy Marijskogo nauchno-issledovatel'skogo instituta jazyka, literatury i istorii, vol. 15, pp. 95-111.
- Khanina O. 2000. Razlichnyje stadii grammatikalizacii glagola rechi v "podchinitelnyj sojuz" pri KSA // Computational Linguistics and Intellectual Technologies. Proceedings of the International Conference Dialogue'2000, vol. 1. Moscow: Nauka.
- Khanina O. 2003. Opyt opisanija semanticheskoi zony zhelanija i sredstv ee vyrazhenija (na materiale chuvashskogo jazyka) [Essay of desiderative semantic zone and its expression (based on Chuvash)], /Vestnik MGU, Filologija 5/ [Working Papers of Moscow State University, Philology 5]. 2003. 101-116.
- Khanina O. 2007. Konstrukcii s grammatikalizovannym konverbom glagola rechi. [Constructions with grammaticalized converb of speech verb], in Lyutikova E.A., Kazenin K.I., Tatevosov S.G., Solovyev V.D. (eds.). /Misharskij dialekt tatarskogo jazyka. Očerki po sintaksisu i semantike. /[Mishar dialect of Tatar. Essays on syntax and semantics]. Kazan. 2007, 126-140.
- Klamer M. 2000. How report verbs become quote markers and complementizers // Lingua, vol. 110, pp. 69-98.
- Kononov A.N. 1953. O sojuznom slove *diye* v tureckom jazyke // Akademiku V.A. Gordlevskomu k jego 75-letiju. Moscow: AN SSSR.
- Lehman Ch, 2002. Thoughts on grammaticalization. ASSIDUE. Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt, Nr. 9.
- Lord, Carol. 1976. Evidence for syntactic reanalysis: from verb to complementizer in Kwa.
- Majtinskaja K.E. 1982. Sluzhebnyje slova v finno-ugorskix jazykax. Moscow: Nauka.
- Serdobolskaja N.V., Toldova S.Ju. 2006. Puti grammatikalizacii glagola rechi *manaš* v konstrukcijax s sentencialnymi aktantami v marijskom jazyke // The grammar and pragmatics of complex sentences in languages spoken in Europe and North and Central Asia. International linguistic symposium LENCA-3. Book of abstracts. Tomsk: "Veter".
- Timofeeva V.T. 1961. Osobennosti kosvennoj rechi marijskogo jazyka // N.I. Isanbaev (ed.) Voprosy marijskogo jazyka. Joshkar-Ola, pp. 95-101.
- Toldova S.Ju. 1999. Osobennosti mestoimennoj referencii pri peredache chuzhoj rechi (mezhdue deiksisom i anaforoj) // Computational Linguistics and Intellectual Technologies. Proceedings of the International Conference Dialogue'1999, vol. 1. Moscow: Nauka.