Grammaticalization of the verb of speech in Finno-Ugric languages
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-COMP²-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-MCONV  boast-PST.3SG

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

(3) jəvan  ola-ške  kaj-em  man-ən  şon-a.

Ivan  city-LAT  go-PRS.1SG  say-MCONV  think-PRS.3SG

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

The same phenomenon is observed in Besermian variant of Udmurt language with the converb of the verb *šuənə* ‘to say’:

**UDMURT (BESERMEN)**

(4) mar  pe  ta?  –  ź'ič'ə  šu-e.

what  CIT  DEM  fox  say-PRS.3SG

*“What is this?” – The fox says.*

(5) Vas’a  vera-z  brat-ez-la,  soje  žug-o-z  šu-sa.

Vasja  tell-PST.3  brother-P.3SG-DAT  dem.ACC  beat-FUT-3  say-MCONV

*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-MCONV

*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

---

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

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, Sordobolskaya 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. anaj4ez vera4z ataj4ez-Iə mone kwaret-i-z
   mother-P.3SG tell-PST.3 father-P.3SG-DAT I.ACC scold-PST.3
   kužaj4e šuə-sa.
   boss-P.1SG say-CONV
b. anaj4ez vera4z ataj4ez-Iə soje kwaret-i-z
   mother-P.3SG tell-PST.3 father-P.3SG-DAT dem.ACC scold-PST.3
   kužaj4ez šuə-sa.
   boss-P.3SG say-CONV

The mother said to the father that her boss had scolded her (a. ...that my boss scolded me,
b. ...that her boss scolded her).
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 küšť-en.
   father girl-DAT house wash-JUSS say-CONV order-PST.3SG
b. ača üdər-lan pört musk-aš küšť-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 šua-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 dastıš-k-o-m n’i turən
   July-ILL August-ILL dem hay prepare-MFUTM1PL already hay
   [život-lə tolalte med okm-o-z šu-sa].
   cattle-DAT in.winter OPT be.enough-MFUTM3 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.
In Besermyan, the choice of the reference strategy is influenced by the syntactic position of the coreferential NP. The subject of the dependent clause is more often encoded with the deictic strategy: it is the most preferred possibility with verbs of speech (13), it is equiprobable with verbs of speech causation. However, it is much more rarer observed with mental verbs, verbs of emotion (14) and causation, and it is totally excluded in adverbial clauses of purpose and reason. As for direct and indirect object in the dependent clause, they are only marginally encoded along the deictic strategy, cf. (15a) and (15b).

(13) so šu-i-z ʒ'etaž'e lokt-o šuə-sa.
DEM say-PST-3 in.the.evening come-FUT.1 say-CONV

He said he would come in the evening.

(14) pič'i pi kəška, so aldaš'k-o-z čaššaj4en šu4sa.
little boy be.afraid.PRS.3SG dem be.lost-FUT-3 forest-INS say-CONV

The boy is afraid that he will get lost in the forest.

(15) a. Vas'a vera-z brat-ez-lə, soje žug-i-z-ə šu-sa.
Vasja tell-PST brother-P.3SG-DAT DEM.ACC beat-PST-3-PL say-CONV

Vasja told his brother that someone beat him.

b. Vas'a vera-z brat-ez-lə, so-lo vit' pukt-i-z-ə šu-sa.
Vasja tell-PST brother-P.3SG-DAT demMDAT five se t say-CONV

Vasja told his brother that he was given a “5”.

It is an interesting peculiarity of Besermyan, that possessive suffixes on the subject of the dependent clause show the same behaviour as the subject itself, showing a strong preference for the deictic strategy with verbs of speech:

(16) turən4 eəvl šu-sa vera-š'k-e val n'i.
hay-P.1SG NEG say-CONV tell-DETR-PRS.3SG was already

She said she already had no hay.

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

<table>
<thead>
<tr>
<th>S 1 / IO 1</th>
<th>S 2, POSS 2</th>
<th>DO 2, IO 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>verbs of speech: ’say’, ’tell’</td>
<td>deictic (anaphoric)</td>
<td>anaphoric (deictic)</td>
</tr>
<tr>
<td>verbs of speech causation: ’demand’, ’request’</td>
<td>deictic/anaphoric</td>
<td>anaphoric (deictic)</td>
</tr>
<tr>
<td>mental verbs: ’think’, ’know’, ’believe’</td>
<td>anaphoric (deictic)</td>
<td>anaphoric (deictic)</td>
</tr>
<tr>
<td>verbs of emotion: ’rejoice’, ’be angry’</td>
<td>anaphoric (deictic)</td>
<td>anaphoric (deictic)</td>
</tr>
<tr>
<td>verbs of causation: ’make’, ’send’; verbs of intention: ’decide’, ’want’</td>
<td>anaphoric (deictic)</td>
<td>anaphoric (deictic)</td>
</tr>
<tr>
<td>adverbial purposive clauses</td>
<td>anaphoric only</td>
<td>anaphoric only</td>
</tr>
<tr>
<td>adverbial clauses of reason</td>
<td>anaphoric only</td>
<td>anaphoric only</td>
</tr>
</tbody>
</table>

The same parameters are relevant for the choice of the reference strategy in Mari; however, they are distributed in a different way. The deictic strategy is used more often than in Besermyan. The subject of the dependent clause can be encoded along the deictic or the anaphoric strategy with all the matrix verbs, the verbs of speech most often taking the deictic strategy (17). Direct and indirect objects and the possessive suffixes show preference towards the anaphoric strategy (18); (19).

3 The symbols “S 1”, “IO 1” denote subject and indirect object in the matrix clause; “S 2”, “DO 2”, “IO 2”, “POSS 2” denote subject, direct object, indirect object, and possessive suffixes on the subject of the dependent clause.
The brother told to the sister that he has killed the bear.

The brother told to the sister that the bear had wounded him.

The lad said that his 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:

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

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.

You said you wouldn’t come again. (Khanina 2007: 132)

Alsu thinks that I have left (lit. thinks about me “He’s left”). (Ibid.)
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


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.


