

NEITHER A LEXICAL VERB NOR A COMPLEMENTIZER: EVIDENTIAL SAY-BASED
COMPLEMENTS IN CHUVASH*

Mikhail Knyazev

(Institute for Linguistic Studies, RAS; HSE University, St. Petersburg)

misha.knjazev@gmail.com

Abstract. The paper discusses a potential challenge to the (full) decompositional analysis of ‘say’-complementizers according to which they are synchronically analyzed as respective nonfinite forms of a lexical verb ‘say’ (cf. Major 2021), by focusing on *tenine*, the accusative-marked (event) nominalization of *te* ‘say’, which is used in the complementizer-like function with verbs like ‘hear’, ‘know’, ‘believe (someone)’, etc. in the Poshkart dialect of Chuvash. Based on data previously described by Knyazev (2022), it is shown that *tenine* is not compositionally interpreted as a nominalization of a lexical verb ‘say’ since this leads to incorrect paraphrases (‘heard/learned *that X/someone* said that *p*’). However, an analysis where *tenine* is a special complementizer (cf. Knyazev 2022) must also be rejected since *tenine* syntactically patterns like a nominalization. To resolve this paradox, it is proposed that *tenine* should be analyzed as involving a reportative evidential light verb ‘say’, with the saying event component having the status of a presupposition (Simeonova 2020). The possibility of extending this account to the more familiar converbial complementizer *teze* is briefly discussed, as well as implications for Korotkova’s (2016) generalization regarding restrictions on the interpretation of ORIGO in evidentials.

* This paper combines typology of reported speech constructions and formal semantics, two things Sasha was passionate about. It also brings back the memories of my short stay in LLACAN in Paris in June 2019, where I was working on the very same constructions discussed in this paper. During this stay I spent wonderful time with Sasha and her friends at Square Barye and at her place in Villejuif (twice), which was the last time I saw her. I don’t remember us discussing *teze* and *tenine* as somehow there were always more exciting things to talk about. But I am sure she would have enjoyed discussing these data and my current formal take on them.

1. Introduction[†]

Like other Turkic languages, (Poshkart) Chuvash makes extensive use of ‘say’-based forms to introduce finite complement clauses.¹ For example, as shown in (1a)–(1b), clausal complements of most common speech and belief verbs such as *kala* ‘say’ and *şotla* ‘think’ and others are introduced by *teze*, morphologically the same-subject (-ZA) converb of the verb *te* ‘say’, which can also function as a lexical verb, as shown in (1c).²

- (1) a. maşə [şomər bol-at **te-ze**] kala-r-ë.
Masha rain be-Npst[3Sg] say-Cvb say-Pst-3Sg
‘Masha said that it will rain.’
- b. maşə [van¹uk kaj-za **te-ze**] şotl-at.³
Masha rain go-Cvb say-Cvb say-Npst[3Sg]
‘Masha thinks that Vanyuk has left.’
- c. maşə [yrən şomər bol-at] **te-t**.
Masha tomorrow rain be-Npst[3Sg] say-Npst[3Sg]
‘Masha says it will rain tomorrow.’

In typological literature, elements like *teze* have traditionally been analyzed in terms of grammaticalization (Heine and Kuteva 2002), with an implicit assumption that they have a functional (i.e. complementizer) status. This has also been the default assumption in the generative literature at least until very recently (see Major 2021 for discussion).

However, as was pointed out by Matic’ and Pakendorf (2013), certain instances of ‘say’ may have an intermediate status in terms of the standard parameters of grammaticalization (Heine and Narrog 2010). For example, they may retain phonological substance and be syntactically analyzable

[†] This paper was presented (online) at 8th Workshop on Turkic and Languages in Contact with Turkic (Tu+ 8) (Harvard University, March 4–5, 2023). I thank the audience of the workshop, as well as the reviewers of the paper, for their useful comments.

1 The paper is based on fieldwork data from a dialect of Chuvash spoken in the village of Poshkart (Maloe Karachkino), Chuvashia Republic, Russian Federation. The data was obtained in 2017–2023, originally during student linguistic expeditions to Poshkart organized by Masha Kholodilova and HSE University, St. Petersburg and Moscow. In the examples below, I use transcriptions and simplified glosses developed within this project.

2 I adopt the convention of using capitalization to indicate that the suffix has multiple allomorphs conditioned by vowel harmony and the voicelessness of the preceding sound.

3 Poshkart Chuvash also finite uses of the *-za* converb (cf. (1b)), as well as of the *-n(ə)* participle (cf. (2) below), to express past tense.

as verb forms (converbs, etc.). In addition, because generic verbs of saying may have impoverished semantic content to begin with (e.g. they may express “internal speech” and in general be compatible with the absence of sound production (cf. Major 2021)), it is not obvious that forms like *teze* in (1a)–(1b) have undergone semantic bleaching.

A similar type of approach, referred to below as *decompositional*, has recently become popular in the generative framework, with attempts to analyze converbial ‘say’-complementizers in other Turkic languages such as Turkish, Uyghur and Sakha (Özyıldız et al. 2019; Major 2021) as synchronically converbs (see also Knyazev 2016 for Kalmyk (Mongolic)).⁴ These approaches often capitalize on proposals in the semantics literature, according to which complement clauses denote predicates of individuals with content (of type $\langle e, t \rangle$) and are often viewed as modifiers of the verb (Kratzer 2006; Moulton 2009; Elliott 2017), similarly to what would be usually assumed for ordinary converbial clauses and other adjunct-like expressions. Thus, it becomes possible to analyze examples like (1a)–(1b) as literally meaning ‘X said/thought (something), *saying* p’ (see Section 3).

The goal of this paper is to present challenges to decompositional approaches which arise with *nominalized* forms of ‘say’ in complementizer-like function. Such forms appear to be cross-linguistically less common than converbial forms and have received less attention in the literature. The paper focuses on the form *tenine*, illustrated in (2), which is the (participle-based) event nominalization of *te* ‘say’ in the accusative case in Poshkart Chuvash (Knyazev 2022).

- (2) ep abi-ren [jonazar jal-da pozar pol-nə **te-n-in-e**]
 I mother-Abl neighbor village-Loc fire be-Ptcp say-Ptcp-Poss.3-Acc
 elt-r-ëm.⁵
 hear-Pst-1Sg
 ‘I heard from my mother that there has a been a fire in the neighbouring village.’

As I pointed out in Knyazev (2022), elements like *tenine* pose problems for decompositional approaches because they are syntactically not modifiers and thus on such approaches the ‘say’-clause would be interpreted in the scope of the verb along the lines of ‘X heard *Y/someone say* that p’, which gives an intuitively wrong paraphrase (see Section 2). Therefore, in Knyazev (2022) I rejected the decompositional account and analyzed *tenine* as a complementizer with certain

4 Similar attempts have been made to analyze so-called ‘agreeing complementizers’ in African languages like Lubukusu (Major et al. 2022) and Kipsigis (Driemel and Koeneli 2022) as finite forms of the lexical verb ‘say’.

5 Marker *-n(ə)* participle marker, used in the *tenine*, as well as in event nominalizations more generally (cf. (3a)), is sometimes referred to and glossed as the “past participle”. However, because it can express both simultaneity and anteriority (see Logvinova 2021), I will simply gloss it as “participle”.

syntactic constraints that can be explained *diachronically* by its nominalization source. However, this account is not very satisfactory as it essentially restates the effects of syntactic analyzability (in the form of constraints) instead of directly deriving them from a synchronic analysis. In this paper, I propose a different solution to this problem based on Simeonova’s (2020) analysis of ‘say’-complementizers as involving on a reportative evidential light verb SAY, which crucially allows to view the semantic contribution of *tenine* (‘X said that p’) as *presuppositional*.⁶

The paper is structured as follows. In Section 2, I review the distributional and semantic properties of *tenine*, as well as review arguments for its analyzability from Knyazev 2022. In Section 3, I present my reportative evidential light verb analysis of *tenine*. Section 4 concludes the paper by discussing some potential challenges for the proposed analysis in view of Korotkova’s (2016) semantic analysis of evidentiality and “evidential shift”.

2. The puzzle of *tenine*

2.1. Predicates that occur with *tenine*-clauses

As described in Knyazev (2022), *tenine* introduces complement clauses with a specific class of predicates which have the following properties: i) they have an experiencer subject (and, less commonly, object); ii) an explicit (and sometimes implicit) information source argument, typically realized as oblique; and iii) they allow accusative-marked nominalized complements (cf. (3a) and (6a)). The verb with which *tenine* occurs most commonly is *elt* ‘hear’ (cf. (2) and (3a) below). Speakers also naturally produce examples with *tenine* with verbs *pël* ‘know, learn’, as in (3b), and *san* ‘believe (someone)’/*ënen* ‘believe (someone)’ (so-called “response-stance” sense of *believe*, see Bogal-Albritten and Moulton 2018), as in (3c). They also usually accept constructed examples with *vula* ‘read (somewhere)’, *ənlan* ‘understand’ and *astu* ‘remember’, as in (3d) (other verbs were not systematically tested). In the last two cases, the meaning of the verb might be more accurately translated as ‘infer (from someone’s words)’ and ‘have a recollection of having heard’, respectively.

- (3) a. [man taga san xəjar-san¹-a {ɕiz=er-ze te-n-in-e /
 I.Gen sheep you.Sg.Gen cucumber-Pl-Acc eat=let-Cvb say-Ptcp-Poss.3-Acc

6 Bondarenko’s (2022) dissertation discusses very similar constructions in Buryat (Mongolic) and also proposes an analysis of ‘say’-complements in terms of a contentful functional head (CONT). The analysis proposed here was developed independently of Bondarenko’s account, which I discovered only after I had written up this paper. I leave the comparison between the present and Bondarenko’s accounts for future work.

εiz=er-n-in-e}] **elt-r-ēm.**
eat=let-Ptcp-Poss.3-Acc hear-Pst-1Sg
‘I heard that my sheep has eaten your cucumbers’

b. petja [man kozak sumarlän-za kaj-za te-n-in-e] **pël-të-ë.**
Petya I.Gen cat get.sick-Cvb go-Cvb say-Ptcp-Poss.3-Acc know-Pst-3Sg
‘Petya learned that my cat has ran away (as he was told so).’

c. [kozak tar-za te-n-in-e] ep vas^{jə}-na **şan-d-ëm.**⁷
cat run.away-Cvb say-Ptcp-Poss.3-Acc I Vasya-Acc believe-Pst-1Sg
‘I believed Vasya that the cat has ran away.’

d. papi [jonazar jal-da pozar pol-za te-n-in-e]
grandma neighbor village-Loc fire be-Cvb say-Ptcp-Poss.3-Acc
asta-t.
remember-Npst[3Sg]
‘Grandma remembers that there has been a fire in the neighbouring village (as she was told so).’

Attitude verbs that do not express an information source like *şotla* ‘think’ and ordinary speech verbs like *kala* ‘say’, are incompatible with *tenine*-clauses, as shown in (4a)–(4b), and instead take *teze*-clauses (cf. (1a)–(1b)). *Tenine* is also incompatible with verbs of visual perception such as *kor* ‘see’, as in (4c), since they do not express a hearsay information source, and instead require nominalized complements (see Knyazev 2022:413).

(4) a. *maşə [şomər bol-at te-n-in-e] **kala-r-ë.**
Masha rain be-Npst[3Sg] say-Ptcp-Poss.3-Acc say-Pst-3Sg
Intended: ‘Masha said that the cat has run away.’

b. *maşə [van¹uk kaj-za **te-n-in-e]** **şotl-at.**
Masha Vanyuk go-Cvb say-Ptcp-Poss.3-Acc say-Npst[3Sg]
Intended: ‘Masha thinks that Vanyuk has left.’

⁷ Chuvash does not have a separate dative marker, which coincides with the accusative marker (both are sometimes glossed as “Obj”). However, for simplicity I will gloss it as “Acc” in both cases.

c. *[es pørt tu-za lart-sa **te-n-in-e** petʲə kor-za.
 you house do-Cvb put-Cvb say-Ptcp-Poss.3-Acc Petya see-Cvb
 Intended: ‘Petya saw that you built a house.’ (adapted from Knyazev 2022:413, (9))

In Table 1, I provide frequencies of occurrence of *tenine* (followed by a finite complement clause, which was not graphically marked as quotation) with different predicates in a sample of 181 examples from a (partly) bilingual Russian–Standard Chuvash corpus (<https://ru.corpus.chv.su/cgi-bin/corpus.cgi>). These data provide preliminary support for the distributional restrictions of *tenine* above. For example, we can see that *tenine* occurred, by far, most frequently with ‘hear’, whereas other predicates roughly conform to properties (i)–(ii) (subcategorization properties of these predicates were not checked but they all seem to allow accusative-marked nominalized clauses, based on their translations).⁸

Table 1. Frequency of predicates occurring with *tenine* in the corpus of Standard Chuvash

ilt	‘hear’	110
pëlter (pël-CAUS)	‘mean, indicate (to someone)’	19
ënen	‘believe (someone)’	16
astu	‘remember’	5
pël	‘know, learn’	4
ənlan	‘understand’	2
ənkār	‘guess’	2
asa il	‘take in mind’ (= ‘remember’)	1
asra tît	‘keep in mind’ (= ‘remember’)	2
vula	‘read’	2
jîşan	‘admit’	2
şirëplet	‘confirm’	2
şan	‘believe’	2
ənlantar	‘explain’	2
kəlar	‘discard’	1
man	‘forget’	1
pəxmasərax	‘disregarding’	1
përex	‘equals (to)’	1

8 The verb *pëlter* ‘mean’ in this list, and its relatively high frequency, may appear surprising, but note that it is a causative of *pël* ‘know’ and thus may be expected to involve an (implicit) experiencer, i.e. ‘mean to X that p’ (for more on experiencers with “verbs of demonstration” see Anand and Hacquard 2009). However, it remains to be understood whether and in what sense it may be taken to involve a hearsay information source. I did not specifically investigate the properties of *pëlter* ‘mean’ in Poshkart Chuvash.

purnaşa kært	‘put into life’	1
tyrre kəlar	‘put correctly’ (=‘justify’?)	1
tërəslə	‘make sure’	1
şuta il	‘take into account’	1
şuxəşla	‘assume’	1
itla jivəra il	‘take too gravely’ (= ‘take to heart’)	1

2.2. Arguments against semantic compositionality

Before addressing the question whether *tenine* in examples like (2) and (3a) can be compositionally interpreted as the nominalization of ‘say’, we need to establish whether *tenine* can in principle have an independent semantic contribution. This can be shown by looking at verbs *pəl* ‘know’ and *astu* ‘remember’ (and also *ənlan* ‘understand’). As can be seen from the translations in (3a) and (3c), when these verbs combine with *tenine*-clauses they imply that the information in the complement was communicated to the attitude holder, which I will refer to as the *communicative act (CA) implication*, leaving aside the discussion of its status for Section 3. I will understand CA here rather broadly to include any kind of transmission of verbal information, both in auditory and written form, although most typically it would involve an ordinary speech act. That the CA implication is obligatory with *tenine* is shown by the fact that in a context like (5), where the attitude holder receives the relevant information directly, *tenine*-clauses become infelicitous, as in (5b), and only nominalized clauses are possible, as in (5a).

(5) Context: Grandma saw the fire in the neighbouring village.

a. papi [jonazar jal-da pozar **pol-n-in-e**] pəl-et /
 grandma neighbor village-Loc fire be-Ptcp-Poss.3-Acc know-Npst[3Sg]
 asta-t.
 remember-Npst[3Sg]
 ‘Grandma knows/remembers that there has been a fire in the neighbouring village.’

b. #papi [jonazar jal-da pozar pol-za **te-n-in-e**]
 grandma neighbor village-Loc fire be-Cvb say-Ptcp-Poss.3-Acc

pël-et / asta-t.
 know-Npst[3Sg] remember-Npst[3Sg]

‘Grandma knows/remembers that there has been a fire in the neighbouring village.’

In contrast to verbs like verbs *pël* ‘know’, with verbs like *elt* ‘hear’ and *şan* ‘believe (someone)’ (and also *vula* ‘read’) the CA implication is presumably already included in the lexical meaning of the verb. Therefore, it becomes more difficult to establish that *tenine* has any semantic contribution with these verbs. However, for the sake of argument I will assume that this is also in principle possible.

Incidentally, observe that the CA implication specifically targets the information source (and also the degree of commitment to the proposition) of the attitude holder rather than actual speaker. This can be shown by the fact that a *tenine*-clause is still felicitous in a context where the speaker has *direct* knowledge of the proposition in the complement clause, as shown in (6). Conversely, a *tenine*-clause (on a par with a nominalized clause) is infelicitous with *pël* ‘know’, but not with *elt* ‘hear’, in a context where the speaker takes the proposition to be false, as in (7a). If the *tenine*-clause targeted the information source of the speaker, we would expect the verb *pël* ‘know’ to allow nonfactive uses (because hearsay evidence can be unreliable), like *elt* ‘hear’ in (7b). The fact that such uses are disallowed with *pël* ‘know’ suggests that *tenine* can only affect the cognitive relation between the proposition and the attitude holder.^{9,10}

(6) Context: My cat has got sick (I saw it).

petʰə man kozak sumarlän-za kaj-za **te-n-in-e** pël-tə-ë /
 Petya I.Gen cat get.sick-Cvb go-Cvb say-Ptcp-Poss.3-Acc know-Pst-3Sg
 elt-r-ë.
 hear-Pst-3Sg

‘Petya learned/heard that my cat has got sick (he was told so).’

9 In light of data like (7a), Knyazev’s (2022: 413) characterization of the meaning of *pël* ‘know’ with *tenine* as generally non-factive was inaccurate. A possible reason for that is that Knyazev (2022) only looked at examples where the attitude holder is the speaker, but in such cases ‘know’ (and other factive verbs) may lose its factivity, as was shown by Simons (2007).

10 Curiously, according to at least one of my consultants, the verb *astu* ‘remember’ *can* be used with *tenine* in contexts like (7a), as shown in (i). Crucially, however, this also holds also for a nominalized clause, suggesting that *astu* ‘remember’ in Poskart Chuvash is generally nonfactive.

(i) Context: = (7a)

[petʰə man kozak sumarlän-za kaj-za te-n-in-e / kaj-n-in-e] asta-t.
 Petya I.Gen cat get.sick-Cvb go-Cvb say-Ptcp-Poss.3-Acc go-Ptcp-Poss.3-Acc remember-Npst[3Sg]
 ‘Petya remembers that my cat has got sick (he was told so).’

(7) Context: My cat is in good health (I saw it) but Petya heard a rumour that it got sick.

a. #petʲə man kozək sumarlän-za kaj-za **te-n-in-e** pël-të-ë.
Petya I.Gen cat get.sick-Cvb go-Cvb say-Ptcp-Poss.3-Acc know-Pst-3Sg
'Petya learned that my cat has got sick (he was told so).'

b. petʲə man kozək sumarlän-za kaj-za **te-n-in-e** elt-r-ë.
Petya I.Gen cat get.sick-Cvb go-Cvb say-Ptcp-Poss.3-Acc hear-Pst-3Sg
'Petya heard that my cat has got sick (he was told so).'

Now, the crucial question is whether the CA implication arises due to the presence of a separate clause headed by the lexical verb *te* 'say', which we may take to be the assumption of the decompositional approach. In Knyazev (2022), I presented several arguments against this view, which I reproduce below in a more elaborate form. (Note that these arguments may not be automatically extended from one verb to the next and the question may ultimately have to be decided on a verb by verb basis.)

A first argument is that on the decompositional approach sentences with *tenine* like (2) and (3b)–(3d) would have implausible paraphrases, as in (8), perhaps with the exception of the verb *astu* 'remember' (cf. (3d)). While such paraphrases are in principle possible, they involve *second-hand* reports. Thus, they do not correspond to the intended meanings of the sentences, which describe *first-hand* reports, as is clear from the stimuli given to the speakers. Similarly, such sentences are almost never translated with uncontroversial second-hand reports (with an explicit verb 'say' or a similar predicate) in the bilingual corpus of Standard Chuvash (Knyazev 2022).¹¹

- (8) a. #'I heard from my mother that *she/someone said* that there was a fire in the neighboring village. (cf. (2))
b. #'Petya learned that *someone said* that my cat has got sick. (cf. (3b))
c. #'I believed Vasya that *he/someone said* that the cat has run away. (cf. (3c))
d. ?Grandma remembers that *someone said* that there was a fire in the neighboring village. (cf. (3d))

11 With the verb *astu* 'remember' and synonymous predicates (see Table 1), there is an explicit 'say' or similar predicate in the translation for the majority of the examples. Crucially, at least two examples do not contain an explicit predicate (cf. Knyazev 2022:416, (13c)).

A second argument is that speakers normally translate stimuli involving uncontroversial second-hand reports with the nominalization of the lexical verb *kala* ‘say’ (plus *teze*) (cf. (1a)) and usually disprefer and sometimes reject *tenine*-clauses, as shown in (9).¹² While it would be wrong to assume that *tenine* is strictly disallowed in second-hand reports (see below), this at least suggests that it is not its usual function, weakening the premise of the decompositional analysis.

- (9) maşə [şomər bol-at {te-ze **kala-n-in-e** / ***te-n-in-e**]
 Masha rain be-Npst[3Sg] say-Cvb say-Ptcp-Poss.3-Acc say-Ptcp-Poss.3-Acc
 kala-r-ë.
 say-Pst-3Sg
 ‘Masha said that someone said it will rain.’

A third, and related, argument is that while *tenine* is in principle possible with an overt subject (which would uncontroversially signal its status as a nominalization of ‘say’), as shown in (10a)–(10b), such constructions are generally dispreferred (compared to *teze* + *kalanine*) and at least with verbs *elt* ‘hear’ and also *pël* ‘know’ (but, interestingly, not with *şan* ‘believe (someone)’) must be interpreted as second-hand reports (cf. (10b)). There is an especially strong dispreference for such constructions (but not for constructions with *kalanine*) to have an *overt* subject in the most deeply embedded clause, as shown in (10c). Although the source of the latter constraint (or tendency) is unclear to me, the data show an overall marked status of *tenine* with second-hand reports, reinforcing the point made by the previous argument above.

12 In view of the possibility of expressing second-hand reports with *kalanine*, the first argument must be taken with caution as it relies on the assumption that the uncontroversial embedding of a lexical verb ‘say’ should always be interpreted as a second-hand report. However, this is not always clear. For example, with verbs *elt* ‘hear’ and especially with *şan* ‘believe (someone)’ speakers sometimes use or at least accept *teze* + *kalanine* as translations of stimuli that do not contain an explicit verb ‘say’ and that might thus be plausibly analyzed as first-hand reports (or at least are difficult to distinguish from first-hand reports), as shown in (ia)–(ib).

- (i) a. [man taga san xəjar-sanj-a ɛiz=er-ze te-ze **kala-n-in-e**
 I.Gen sheep you.Sg.Gen cucumber-Pl-Acc eat=let-Cvb say-Cvb say-Ptcp-Poss.3-Acc
 elt-r-ëm.
 hear-Pst-1Sg
 ‘I heard that (allegedly) my sheep has eaten your cucumbers.’ (bracketed material not included in the original stimulus)
- b. [[ul^l-ə pilək il-et te-ze] utɛitel^l kala-n-in-e] aməş
 son-Poss.3 five take-Npst[3Sg] say-Cvb teacher say-Ptcp-Poss.3-Acc mother.Poss.3
 şan-at.
 believe-Npst[3sg]
 ‘Mother believes the teacher (when he says) that her son will get an A.’ (bracketed material not included in the original stimulus)

- (10) a. aməş [ulʲ-ə pilək il-e-p **te-n-in-e**
 mother.Poss.3 son-Poss.3 five take-Npst-1Sg say-Ptcp-Poss.3-Acc
 şan-tə-ë.
 believe-Pst-3Sg
 i. ‘Mother believed that her son will get an A in the exam (as he told her).’
 ii. ‘Mother believed that her son₁ said (to someone) that he₁ will get an A.’
- b. aməş [ulʲ-ə pilək il-e-p **te-n-in-e**
 mother.Poss.3 son-Poss.3sg five take-Npst-1Sg say-Ptcp-Poss.3-Acc
 elt-r-ë.
 hear-Pst-3Sg
 i. # ‘Mother heard that her son₁ will get an A (as he₁ told her).’
 ii. ‘Mother heard (from someone) that her son₁ said that he₁ will get an A.’
- c. utəitelʲ [vasʲə pilək il-et {? te-n-in-e / te-ze
 teacher Vasya five take-Npst[3Sg] say-Ptcp-Poss.3-Acc say-Cvb
 kala-**n-in-e**] aməş elt-r-ë / pël-tə-ë.
 say-Ptcp-Poss.3-Acc mother.Poss.3 hear-Pst-3Sg hear-Pst-3Sg
 ‘Mother learned/heard (from someone) that the teacher said that Vasya will get an A.’

To summarize, while each individual argument above is probably not a knockdown argument against the decompositional analysis, taken together they make this analysis an unlikely possibility.

2.3. Arguments for syntactic analyzability

Now, if the decompositional analysis is not a viable option for *tenine*, should we treat it as an unanalyzed complementizer, as I proposed in Knyazev (2022)? The main problem with this move is that *tenine*-clauses morphosyntactically pattern like accusative-marked nominalizations, suggesting that they *are* nominalizations.

First of all, there is evidence for the syntactically active accusative marker *-(n)A* in *tenine*, as *tenine*-clauses occur only in the object position, whereas in the subject position the corresponding nominative (unmarked) form *teni* is required, as in (11a). Similarly, with verbs that govern instrumental case, such as *kiləş* ‘agree’, the instrumental form *tenibe* is used, as in (11b).

(11) a. [man eki kate-a {tok-sa **te-n-i** / tok-n-i]
 I.Gen sister groom-Acc go.out-Cvb say-Ptcp-Poss.3 go.out-Ptcp-Poss.3
 tērēs mar.
 true Neg

‘That my sister got married (as they say) is not true.’ (adapted from Knyazev 2022: 417,

(14a))

b. petʰə [maʃə lajək {jorl-at **te-n-i-be** / jorla-n-i-be}]
 Petya Masha good sing-Npst[3Sg] say-Ptcp-Poss.3-Ins sing-Ptcp-Poss.3-Ins
 kilēs-r-ē.
 agree-Pst-3Sg

‘Petya agreed (with the claim) that Masha sings well.’ (adapted from Knyazev 2022: 419,

(16a))

Note that *teni* and *tenibe* in examples like (11) have very similar semantic properties as *tenine* as both involve the CA implication (note that in (11a) the attitude holder is an implicit “judge”) and, crucially, resist a fully compositional interpretation: for example, in (11a) what the judge denies is the proposition itself, not the fact that it was communicated, and similarly for (11b). This suggests that all the three forms contain an underlying nominal projection.

This, however, is not sufficient to decompose *tenine* into a nominalization of *te* ‘say’ (as complementizers can also be embedded within a nominal projection). To show that *tenine* (and other nominalized forms) also contains a syntactically active nominalization marker, we shall look at the adnominal position, e.g. with nouns like *xibar* ‘news’. In this position, the ‘say’-complementizer takes the form *tenĕ* or *tegen*, which are morphologically participles of *te* ‘say’, as shown in (12).¹³

13 Note that *tenĕ* contains the same participle marker as found in *tenine* (*teni* and *tenibe*). However, the two they have different morphosyntactic properties since the participle marker does not take case or possessive morphology and occurs adnominally. It is plausible to analyze this as a case of morphological underspecification. An alternative view is that this *-n(ĕ)* is always syntactically a participial head, whereas nominalizing function is performed by a separate D head, realized by the 3Sg possessive marker *-i* (see Section 3) or possibly null. Note also that the marker *-AGAn* can only be used a participle (Logvinova 2019b).

- (12) a. [jonazar jal-da pozar pol-za **te-në** / **pol-nə**] xibar-a
 neighbor village-Loc fire be-Cvb say-Ptcp be-Ptcp news-Acc
 elt-r-ëm.
 hear-Pst-1Sg

‘I heard the news that there was a fire in the neighboring village (as they say).’ (adapted from Knyazev 2022:420, (18a))

- b. [es katə-a kaj-za **t-egen**] xibar por.
 you.Sg groom-Acc go-Cvb say-Ptcp.Prs news Cop

‘There is news that you got married (as they say).’ (adapted from Knyazev 2022:420, (18b))

Note, again, that like *tenine*, forms *tenë* and *tegen* in (12) are not compositionally interpreted since the respective sentences do not have the meaning ‘I heard the news that *someone said/says* that p’ and ‘there is news that *someone says* that p’, showing that these forms cannot be analyzed as ordinary nominalizations of a lexical verb ‘say’.¹⁴ At the time time, the CA implication is present, as in the case of *tenine*. This suggests that all of the above forms are underlyingly the same verbal root *te* ‘say’ plus a syntactically active nominalization or participle marker.

An additional reason to analyze *tenine* as containing a syntactically active nominalization (or participle) marker comes from restrictions on the converbial form *teze* (cf. (1a)–(1b)). As I showed in Knyazev (2022), *teze*-clauses cannot occur in the subject position, as shown in (13a), or in the adnominal position, as in (13b). This would be unexpected if *teze* were an unanalyzed complementizer but would directly follow if *teze* were syntactically a converb of *te* ‘say’, heading a VP adjunct (cf. Major 2021).

- (13) a. *[man eki katə-a tok-sa **te-ze**] tērəs mar.
 I.Gen sister groom-Acc go.out-Cvb say-Cvb true Neg

Intended: ‘That my sister got married is not true.’ (adapted from Knyazev 2022:417, (14a))

- b. *[es katə-a kaj-za **te-ze**] xibar por.
 you.Sg groom-Acc go-Cvb say-Cvb news Cop

Intended: ‘There is news that you got married.’ (adapted from Knyazev 2022:420, (18b))

14 Note that *tenë/tege*-clauses cannot be a subject relative clause (as in ‘the news saying that p’) since the verb *te* ‘say’ cannot take *xibar* as a subject, as shown in (i).

(i) *xibar [es katə-a kaj-za **t-et**].
 news you.Sg groom-Acc go-Cvb say-Npst[3Sg]
 Intended: ‘The rumor says that you got married.’

The most natural conclusion from the above discussion is that *tenine* (and other complementizer-like forms) should have syntactically active pieces of morphology, built on top of the verbal root *te* ‘say’. This, however, leads to a paradox: how can *tenine* and similar forms be *syntactically analyzable without being semantically decomposable* (see Section 2.2). One may try to view such forms as unanalyzed complementizers with constraints inherited from their diachronic origin as respective nonfinite forms, as I did in Knyazev (2022). However, this does not provide a plausible account of what is *synchronically* stored in speakers’ minds. As I hope to have shown now, it is likely that speakers implicitly perceive forms such as *teze*, *tenine*, etc. as syntactically analyzable. Another possibility is to assume that syntactic and semantic decomposition may not go hand in hand, which seems to be a default assumption in grammaticalization theory (Heine, Narrog 2010). However, I believe this does not seriously address the synchronic status of such forms in speakers’ mental grammar.

Before moving on to the account that I propose in the next section, I wish to point out the paradox at hand arises most clearly with nominalized forms such as *tenine*. This is because it is less obvious whether the decompositional account of *converbial* forms of ‘say’ is untenable, given the relative plausibility of paraphrases in (14a)–(14b) (cf. (1a)–(1b)).

(14) a. maʂə [ʂomər bol-at **te-ze**] kala-r-ë.
 Masha rain be-Npst[3Sg] say-Cvb say-Pst-3Sg
 Literally: ‘Masha spoke, saying that the cat has run away.’

b. maʂə [vanʷuk kaj-za **te-ze**] ʂotl-at.
 Masha rain go-Cvb say-Cvb say-Npst[3Sg]
 Literally: ‘Masha thinks, saying (to herself) that Vanyuk has left.’

While such paraphrases may sound redundant, they do not obviously lead to incorrect truth condition, in contrast to the case of *tenine*-clauses (cf. (8)). Therefore, the (full) decompositional analysis of *teze*-clauses remains a viable possibility (see Section 3.6)

3. Analysis

3.1. Proposal in a nutshell

The basic idea of the proposed account is that *te* ‘say’ in complementizer-like forms should be viewed as a *functional* or *light* (as opposed to lexical) verb, but in a stronger sense than suggested by Grimshaw (2015), where the light verb SAY (realized as *say* in English) is identified with the shared semantic component of various speech verbs. Specifically, I adopt Simeonova’s (2020) proposal according to which ‘say’ in complementizer-like forms is semantically equivalent to a *reportative evidential* marker, which is made initially plausible by the fact that *te* ‘say’ independently allows reportative evidential uses, as in (15) (on evidential uses of ‘say’ more generally see Matic’ and Pakendorf 2013: 377).

(15) [san kozak sumarlan-za kaj-nə] **t-eteë**.

you.Sg.Gen cat get.sick-Cvb go-Ptcp say-Npst.Pl

‘Your cat has got sick, they say.’ (= ‘They say that your cat has got sick’) (adapted from Knyazev 2022: 410, (4c))

The most important aspect of Simeonova’s proposal is that the meaning of ‘say’ (corresponding to the CA implication) is treated as a *presupposition* and is therefore not interpreted in the scope of the attitude verb (assuming that the relevant presupposition projects). This is in line with the view that the semantic contribution of grammatical markers is often presuppositional (see e.g. Heim and Kratzer 1998 on phi-features; see also Boye and Harder 2012 from a different perspective).

However, I will depart from Simeonova’s (2020) assumption that reportative evidential ‘say’ corresponds to the highest projection of the embedded clause (which she treats as a small *c* above CP) as this is hard to reconcile with the fact that in Chuvash evidential *te* ‘say’ occurs with nonfinite markers (see Section 2.3), which would normally be analyzed as functional heads above the verb. Instead, I will assume that evidential *te* ‘say’ corresponds to the functional head REP(ortative) in the left periphery of the clause but below the functional head corresponding to the nonfinite marker, as shown in (16) (adjusted for head-finality).

(16) ..._{Clause}] REP] Nmz/Ptcp/Cvb]

3.2. Semantic framework

Simeonova (2020) proposes a semantic analysis of REP within the general framework where propositions (elements of type $\langle s,t \rangle$) do not directly combine with attitude verbs but must first be turned into predicates of individuals with content, represented as x_c (elements of type $\langle e,t \rangle$), i.e. the denotation of nouns like *rumor*, *idea*, etc. (Kratzer 2006; Moulton 2015; Elliott 2017; Bondarenko 2021). This operation is often assumed to be performed by the complementizer itself, which can thus be analyzed as a function from propositions to predicates of contentful individuals, as in (17a) (but see Elliott 2017:§2.8).

- (17) a. $\llbracket \text{that} \rrbracket = \lambda p_{\langle s,t \rangle} . \lambda x_c . [\text{CONT}(x) = p]$ (adapted from Moulton 2015:312, (19b))
 b. $\llbracket \text{CP} \rrbracket = \lambda x_c . [\text{CONT}(x) = p]$
 c. $\llbracket \text{that there was a fire} \rrbracket = \lambda x_c . [\text{CONT}(x) = \lambda w . \text{there was a fire in } w]$

There are different approaches as to how attitude verbs compose with *that*-clauses. For consistency, I assume Simeonova's (2020) own analysis, illustrated in (18) (with Neo-Davidsonian logical representations, for more transparency), where verbs like *think* and *say* are analyzed as directly taking predicates of contentful individuals as arguments. On this analysis, the verb introduces an existentially quantified variable x_c , which is restricted by the *that*-clause.

- (18) a. $\llbracket \text{think} \rrbracket = \lambda P_{\langle e,t \rangle} . \lambda e . \exists x_c . [\text{thinking}(e) \wedge \text{theme}(e) = x \wedge P(x)]$
 (adapted from Simeonova 2020: 237, (230))
 b. $\llbracket \text{think that there was a fire} \rrbracket = \lambda e . \exists x_c . [\text{thinking}(e) \wedge \text{theme}(e) = x$
 $\wedge \text{CONT}(x) = \lambda w . [\text{there was a fire in } w]]$
 c. $\llbracket \text{John think that there was a fire} \rrbracket$
 $= \lambda e . \exists x_c . [\text{thinking}(e)(x) \wedge \text{CONT}(x) = \lambda w . [\text{there was a fire in } w] \wedge \text{holder}(e) = \text{John}]$

3.3. Analysis of the reportative evidential 'say'

Built on these general assumptions, Simeonova (2020) proposes an analysis of REP illustrated in (19). The analysis is similar to Moulton's (2015) analysis of complementizer in (17a) in that REP also takes a proposition and eventually returns a predicate of contentful individuals. But it differs from it in three respects. First, REP takes an implicit holder argument in its specifier, whose reference is supplied by the context (the assignment function g) (cf. (19b)). Second, and most

importantly, REP introduces the presupposition that there was a saying event involving the holder argument and the contentful individual introduced by the verb.¹⁵ Third, REP does not return a predicate of contentful individuals directly. Instead, Simeonova assumes that the variable x_c is originally existentially closed in the denotation for REP (to achieve a t-type meaning for unembedded uses of evidentials, cf. (19c)) but is later *disclosed* in the syntax by way of a special operation (cf. (19d)). As a consequence of this analysis, ‘say’-complements end up having the same denotation as *that*-clauses except that in addition they presupposes a saying event.

(19) a. $[[\text{REP}]] = \lambda p_{\langle s, t \rangle} . \lambda x_e . \exists y_c . [\text{CONT}(y) = p]$ (adapted from Simeonova 2020:233, (224))

Defined only if $\exists e . [\text{saying}(e) \wedge \text{theme}(e) = y \wedge \text{holder}(e) = x]$

b. $[[\text{REP} [\text{there was a fire}]]] = \lambda x_e . \exists y_c . [\text{CONT}(y) = \lambda w . \text{there was a fire in } w]$

Defined only if $\exists e . [\text{saying}(e) \wedge \text{theme}(e) = y \wedge \text{holder}(e) = x]$

c. $[[\text{pro}_7 \text{ REP} [\text{there was a fire}]]]^{\text{e}}$ (matrix evidential)

= $\exists y_c . [\text{CONT}(y) = \lambda w . \text{there was a fire in } w]$

Defined only if $\exists e . [\text{saying}(e) \wedge \text{theme}(e) = y \wedge \text{holder}(e) = g(7)]$

d. $[[\exists\text{-DIS} [\text{pro}_7 \text{ REP} [\text{there was a fire}]]]]^{\text{e}}$ (embedded evidential)

= $\lambda y_c . [\text{CONT}(y) = \lambda w . \text{there was a fire in } w]$

Defined only if $\exists e . [\text{saying}(e) \wedge \text{theme}(e) = y \wedge \text{holder}(e) = g(7)]$

Embedded evidentials with verbs ‘think’ and ‘say’, based on Bulgarian data, are analyzed as in (20) (Simeonova does not analyze embedded evidentials under verbs like ‘hear’ and ‘know’).

(20) a. $[[\text{Zlati think} [\text{pro}_7 \text{ REP} [\text{there was a fire}]]]]^{\text{e}}$

= $\lambda e . \exists y_c . [\text{thinking}(e, y) \wedge \text{CONT}(y) = \lambda w . [\text{there was a fire in } w] \wedge \text{holder}(e) = \text{Zlati}]$

Defined only if $\exists e . [\text{saying}(e) \wedge \text{theme}(e) = y \wedge \text{holder}(e) = \text{Zlati}]$

¹⁵ Simeonova does not explicitly discuss the lexical meaning of ‘say’. It may be assumed that the meaning is sufficiently abstract so as to include situations where there is no physical production of sound, along the lines of Grimshaw 2015 and Major 2021.

- b. $[[\text{Zlati say } [\text{pro}_7 \text{ REP } [\text{there was a fire}]]]]^{\text{S}}$
 $= \lambda e. \exists y_c. [\text{saying}(e, y) \wedge \text{CONT}(y) = \lambda w. [\text{there was a fire in } w] \wedge \text{holder}(e) = \text{Zlati}]$
 Defined only if $\exists e. [\text{saying}(e) \wedge \text{theme}(e) = y \wedge \text{holder}(e) = \text{Zlati}]$

(adapted from Simeonova 2020:239, (234))

On this view, embedded evidential complements have a very similar meaning as ordinary *that*-clauses (cf. (18)), except that they also presuppose a corresponding saying event, which basically repeats or reinforces the meaning of the matrix verb. This seems to fit with the semantic characterization of (speaker-oriented) embedded evidentials in the literature (cf. Korotkova 2016).

Note also that on Simeonova’s account the person whose information source is tracked by the evidential marker (so-called ORIGO) is not explicitly represented but is determined pragmatically. For example, in examples (20) the ORIGO is naturally understood as the actual speaker (rather than the attitude holder) presumably because the evidence of the speaker for the communicated content is more relevant since the speaker is marking the report about this content (see Section 4).¹⁶

3.4. Analysis of presuppositional verbs

Before presenting my analysis of *tenine*, I will make certain assumptions about the meanings of verbs which combine with *tenine*-clauses. Recall that *tenine*-clauses occur with a specific set of verbs such as by *elt* ‘hear’, *pěl* ‘know, learn’, *šan* ‘believe (someone)’, *ənlan* ‘understand’, *astu* ‘remember’ and a few others. I assume that these verbs are *presuppositional* in the sense that their complements refer to existing discourse referents, i.e. that are presupposed to be true or familiar to the addressee (as in the case of response-stance verbs like *agree*) (Kastner 2015). I also follow Kastner 2015 and similar proposals such as Bochnak and Hanink 2022, according to which presuppositional verbs select definite DP arguments (of type *e*), as illustrated in (21).

- (21) $[[\text{know}]] = \lambda x_c. \lambda e. [\text{knowing}(e) \wedge \text{theme}(e) = x]$

Since *that*-clauses, which have the $\langle e, t \rangle$ -type denotation (see Section 3.2), cannot directly combine with presuppositional verbs, I assume, following Bochnak and Hanink (2022), that they are embedded in a DP-shell headed by a definite D (more specifically, an anaphoric/familiar definite),

¹⁶ Cf. the following quote: “In reportative evidentials, the speaker is still Origo by virtue of sincerely making an utterance about the content of what was said, i.e. the speaker cannot make such a conversational move without being the ‘origo’ for independent reasons that do not necessitate the introduction of such a concept formally.” (Simeonova 2020:262)

which converts the predicate of individuals with content p into the unique such individual that exists in the discourse. I assume the implementation in Jenks and Konate (2022), where the anaphoric definite article (D^x) corresponds to the iota-operator with an additional index argument in its specifier, which returns the unique individual satisfying the restriction and identical to the index, as shown in (22a). On this view, familiar complements denote unique contentful individuals whose content is specified by the proposition in the complement and that are identical to some salient individual with content in the discourse, as shown in (22b)–(22c).

- (22) a. $[[D^x]] = \lambda P_{\langle e,t \rangle} . \lambda y_e . \iota x . [P(x) \wedge x = y]$ (adapted from Jenks and Konate 2022:17, (25b))
- b. $[[\text{8 } D^x [\text{that there was a fire}]]]^\text{B} = \iota y_c . [\text{CONT}(y) = \lambda w . [\text{there was a fire in } w] \wedge y = g(8)]$
- c. $[[\text{John know } [\text{8 } D^x [\text{that there was a fire}]]]]^\text{B} = \lambda e . [\text{knowing}(e) \wedge \text{theme}(e) = \iota y_c . [\text{CONT}(y) = \lambda w . [\text{there was a fire in } w] \wedge y = g(8)]]$

3.5. Analysis of *tenine*-clauses

Now, we are ready to give the analysis of *tenine*-clauses. As discussed above, the main assumption is that *te-* realizes the reportative evidential head REP, as in (19a). I also assume that *-n(ə)-* in *tenine* (glossed as Ptcp) is a nominalizer (Nmz) and is semantically inert (e.g. an identity function, but I will simply ignore its denotation below). The function of *-i(n)-* in *tenine* and nominalizations more generally (which is glossed as a third person possessive marker (Poss.3)), is less straightforward. As was showed by Kozhemyakina (2017), this marker does not track the person and number of the subject but is basically lexicalized (obligatory) at least in the nominative (unmarked), accusative and the instrumental case. I will tentatively assume that the main function of *-i(n)-* is to encode familiarity, which is also one of its main functions in ordinary noun phrases, as illustrated in (23) (see Logvinova 2019a). Thus, I will assume that it realized D^x (cf. (22a)).

- (23) *kəzəl ebë kërək il-d-ëm. kërëg-ë pet ozë-sker*
 this.year I fur.coat buy-Pst-1Sg coat-Poss.3 very warm
 ‘This year I bought a coat. **The coat** is very warm.’ (adapted from Logvinova 2019:18, (22))

With these ingredients, *tenine*-clauses in examples like (24) (cf. (7b)) can be analyzed as in (25a), which, when combined with the verb *elt* ‘hear’, gives roughly the meaning in (25b): the sentence is

true if there is a hearing event whose experiencer is Petya and whose theme is the unique individual y_c in the discourse context such that it has as its content the proposition that my cat got sick. The sentence also presupposes that there is a saying event with y_c as its theme is and with Masha as its holder. This provides intuitively correct truth conditions for (24). A similar analysis can be given to sentences with verbs *pël* ‘know, learn’, *şan* ‘believe (someone)’ and others.

(24) Context: Masha told Petya that my cat has got sick.

petʃə man kozək sumarlan-za kaj-za **te-n-in-e** elt-r-ë.
Petya I.Gen cat get.sick-Cvb go-Cvb say-Ptcp-Poss.3-Acc hear-Pst-3Sg
‘Petya heard that my cat has got sick.’

(25) a. $\llbracket \llbracket \text{DP } 8 \llbracket \text{NP } [\text{pro}_7 \llbracket \text{my cat has got sick} \rrbracket \text{te}_{\text{REP}}] \text{-n}_{\text{N}}] \text{-ine}_{\text{DX}} \rrbracket \rrbracket \rrbracket^{\text{g}}$

= $\iota y_c. [\text{CONT}(y) = \lambda w. [\text{my cat has got sick in } w] \wedge y = g(8)]$

Defined only if $\exists e. [\text{saying}(e) \wedge \text{theme}(e) = y \wedge \text{holder}(e) = \text{Masha}]$

b. $\llbracket (24) \rrbracket^{\text{g}} = \exists e. [\text{hearing}(e) \wedge \text{theme}(e)]$

= $\iota y_c. [\text{CONT}(y) = \lambda w. [\text{my cat has got sick in } w] \wedge y = g(8)] \wedge \text{experiencer}(e) = \text{Petya}$

Defined only if $\exists e. [\text{saying}(e) \wedge \text{theme}(e) = y \wedge \text{holder}(e) = \text{Masha}]$

Most importantly, because the saying event on this analysis is not part of the asserted context we do not run into implausible paraphrases, as in (8a). At the same time, the analysis correct captures the CA implication with *tenine*-clauses since in order to satisfy the presupposition in question, the sentence must occur in a context where the propositional content was communicated by someone. Now, with verbs like *elt* ‘hear’ and *şan* ‘believe (someone)’ the relevant presupposition may be satisfied almost vacuously since these verbs would normally imply the existence of a saying event by virtue of their lexical meaning (at least if treat the propositional sense of *elt* ‘hear’ and the response-stance sense of *şan* ‘believe’ as lexically specified, as opposed to derived compositionally). Therefore, the CA implication of *tenine* may be hard to detect in these cases. By contrast in the case of *pël* ‘know, learn’ (and also *astu* ‘remember’ and *ənlan* ‘understand’) (cf. (3b) and (3d)) the verb normally does not imply the existence of a saying event and thus the semantic contribution of *tenine* is more noticeable, leading to speakers’ robust intuition that the information in the complement was obtained though hearsay.

The analysis above also correctly predicts that *tenine*-clauses should be incompatible with *gotla* ‘think’, *kala* ‘say’ and similar verbs (cf. (4a)–(4b)) since such verbs take ⟨e,t⟩-type expressions as arguments (cf. (18a)) (or perhaps as modifiers, on Elliott’s (2017) and Bochnak and Hanink’s (2022) analysis), whereas *tenine*-clauses have type e, resulting in type clash.¹⁷

To summarize, the proposed evidential light verb analysis allows to capture both the syntactic analyzability of *tenine* and the unavailability of a full-fledged semantic decomposition with a lexical verb ‘say’.

3.6. Extending the analysis to *teze*-clauses?

A more controversial question is whether the same analysis should be extended to *teze*-clauses. On the one hand, on the analysis in (19) and assuming that the converbial marker *-ZA* is semantically inert, *teze*-clauses would denote predicates of contentful individuals (type ⟨e,t⟩), correctly predicting that they will felicitously combine with verbs like *gotla* ‘think’ and *kala* ‘say’ (cf. (1a)–(1b)). In addition, because *teze*-clauses have the basic meaning of *that*-clauses (cf. (18c)), they will lead to an initially plausible compositional meaning of the respective sentences, as shown in (26), corresponding to (1b).

$$(26) \quad \llbracket \text{Masha} [\text{C}_{\text{vbP}} [\text{pro}_7 [\text{it will rain}] \text{te}_{\text{REP}}] \text{-ze}_{\text{C}_{\text{vb}}}] \text{thinks} \rrbracket^{\text{e}}$$

$$= \lambda e. [\text{thinks}(e) \wedge \text{theme}(e) = x \wedge \text{CONT}(x) = \lambda w. [\text{it will rain in } w] \wedge \text{holder}(e) = \text{Masha}]$$

$$\text{Defined only if } \exists e. [\text{saying}(e) \wedge \text{theme}(e) = y \wedge \text{holder}(e) = \text{Masha}]$$

However, it is doubtful that *teze*-clauses in sentences like (1a)–(1b) should actually be analyzed as Bulgarian embedded evidentials in Simeonova 2020, which is what the analysis in (26) corresponds to (cf. (20a)). First, *teze*-clauses is the default, if not the only, strategy of embedding under *gotla* ‘think’ and *kala* ‘say’, implying that all such embedding should be analyzed as embedded reported evidentiality, whereas in the clear cases, including Bulgarian, Turkish and Georgian (cf. Korotkova 2016) embedded evidentials constitute a marked, or at least not the only, strategy. Second, *teze*-clauses do not pass Simeonova’s (2020) diagnostics for evidentiality as e.g. they can occur under negation and with a 1st person subject. Thus, it appears that a full decompositional analysis with a

17 Note that the verb *kala* ‘say’ allows nominalized complements in the ‘talk about’ sense (cf. (9)), suggesting that it is in principle compatible with e-type arguments. However, *kala* ‘say’ still does not combine with *tenine*-clauses (cf. (4a)). A possible explanation for this is that such constructions would be interpreted as second-hand reports, because the subject talks about some *existing* content individual uttered by someone else (or perhaps themselves), whereas, as I showed in Section 2.2, there is a general dispreference for *tenine*-clauses and a preference for *teze* + *kalanine* in second-hand reports.

lexical verb ‘say’ (as part of the asserted meaning) along the lines of (27) (with appropriate changes to the denotation of the ‘think’) may be a more appropriate option for *teze*-clauses, whereby the converbial clause is interpreted as a modifier of the thinking event, with the two events related by a direct causal link (‘~’) (cf. Özyıldız et al. 2019).

$$(27) \exists e. \exists e'. \exists y_c. [e' \sim e \wedge \text{thinking}(e) \wedge \text{holder}(e) = \text{Masha} \wedge \text{theme}(e) = y \\ \wedge \text{CONT}(y) = \lambda w. [\text{it will rain in } w] \wedge \text{saying}(e') \wedge \text{holder}(e') = \text{Masha}]$$

(modeled on Özyıldız et al. 2019:302, (23))

In view of the problems with the evidential analysis of *teze*, it remains to be seen in future research whether and how the analysis of *tenine* and *teze* can be unified.

4. Conclusion and theoretical implications

Before concluding this paper, I wish to say a few words about potential implications of the proposed account of *tenine* (and other nominalized and participial forms of *te* ‘say’) for the semantic analysis of (embedded) evidentiality as developed by Korotkova (2016) (or perhaps vice versa). In her dissertation, Korotkova (2016) proposes a generalization according to which ORIGO (the person whose evidence is tracked by evidentials markers) must coincide with the person from whose perspective the evidence is characterized as direct, hearsay, etc. (in other words, whether the source of information is in the scope of the attitude verb or outside it, i.e. projected to the matrix context). This generalization rules out patterns where: (i) ORIGO is the actual speaker but the source of information is in the subject-belief worlds (nonprojected); and where (ii) ORIGO is the matrix subject (attitude holder) but the source of information is in the speaker-belief worlds (projected). This is illustrated in (29) on the basis of the example (28) from Turkish.

(28) Jay [Anna bir köpek al-mış] di-yor. (TURKISH)
 Jay Anna Ndef dog get-Pst.Indir say-Prs
 ‘Jay said that Anna got a puppy’. (adapted from Korotkova 2016:138)

(29) a. speaker-oriented, projected (speaker thinks they have heard p)

✓Context 1: I was told by Mary, Anna's roommate, that Anna got a dog. Jay visited them recently and has seen the dog himself.

≈ 'Jay said that—and I've heard it—Anna got a puppy.'

b. speaker-oriented, not projected (speaker thinks they have not heard p)

Context 2: Jay visited Anna recently and found out that she finally got a dog. It's is exciting and he is sure that Anna has told me, as she wanted one for a long time. In fact, I was out of town and did not yet hear the news.

≈ 'Jay said that, as he thinks I've heard, Anna got a puppy.' (nonattested)

c. subject-oriented, projected (attitude subject thinks they have not heard p)

Context 3: I visited Anna recently and found out that she finally got a dog. It's is exciting and I am sure that Anna has told Jay, as she wanted one for a long time. In fact, she did call him, but he later forgot about it.

≈ 'Jay said that—and he has heard it—Anna got a puppy.' (nonattested)

d. subject-oriented, projected (*attitude subject thinks they have heard p*)

✓Context 4: I recently visited Anna and found out that she finally got a dog. Jay hasn't visited yet, but she called him to share the news.

≈ 'Jay said that, as he has heard, Anna got a puppy.'

As one may note, *tenine*-clauses appear to violate Korotkova's (2016) generalization, as illustrated in (30) (cf. (6)). On the one hand, they describe the *subject's* evidence since, as we saw earlier, they are felicitous in contexts where the speaker has direct evidence for the proposition. On the other hand, the source of information is characterized as reportative from the *speaker's* perspective.

(30) Context: I know that my cat has got sick (I saw it).

petjə [man kozək sumarlän-za kaj-za te-n-in-e] elt-r-ë.
Petya I.Gen cat get.sick-Cvb go-Cvb say-Ptcp-Poss.3-Acc hear-Pst-3Sg
✓ ‘Petya heard that my cat has got sick, as he was told.’ (subject-oriented, projected)
‘Petya heard that, as he was told, my cat has got sick.’ (subject-oriented, nonprojected)
‘Petya heard that my cat has got sick, as I was told.’ (speaker-oriented, projected)
‘Petya heard that, as I was told, my cat has got sick.’ (speaker-oriented, nonprojected)

Since Korotkova (2016) mainly discusses evidentials under *nonpresuppositional* verbs like ‘say’ and ‘think’, it is unclear whether her generalization is expected to hold also for *presuppositional* verbs like ‘hear’ and ‘know’ and whether the reportative evidential analysis of *tenine* as proposed in this paper provides a genuine counterexample to her generalization. If so, this may lend support for Simeonova’s (2020) treatment of ORIGO as not explicitly represented in the grammar. Alternatively, it may of course also suggest that the evidential analysis of *tenine* is on the wrong track. I leave this as a question for future research.

References

- Bochnak, R., Hanink, E. 2022. Clausal embedding in Washo: Complementation vs. modification. *Natural Language and Linguistic Theory*, v. 40, 979–1022.
- Bogal-Allbritten, E., Moulton, K. 2018. Nominalized clauses and reference to propositional content. In: R. Truswell, C. Cummins, C. Heycock, B. Rabern, and H. Rohde (eds.) *Processing of Sinn und Bedeutung 21*. Edinburgh: University of Edinburgh Press, 215–232.
- Bondarenko, T. 2021. How do we explain that CPs have two readings with some verbs of speech? To appear in *Proceedings of the 39th West Coast Conference on Formal Linguistics (WCCFL 39)*. http://tbond.scripts.mit.edu/tb/wp-content/uploads/2021/09/Bondarenko_TalkP aper.pdf
- Bondarenko, T. 2022. Anatomy of an attitude. PhD diss, MIT.
- Boye, K., Harder, P., 2012. A usage-based theory of grammatical status and grammaticalization. *Language*, v. 88, 1–44.
- Driemel, I., Kouneli, M. 2022. C-Agree is local subject-verb agreement in Kipsigis. Ms. lingbuzz/005661
- Elliott, P. D. 2017. Elements of clausal embedding. PhD diss., University College London.

- Kratzer, A., Heim, I. 1998. *Semantics in generative grammar*. Oxford: Blackwell.
- Heine, B., Kuteva, T. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Heine, B., Narrog, H. 2010. Grammaticalization and linguistic analysis. In: B. Heine, H. Narrog. (eds.) *The Oxford handbook of linguistic analysis*. Oxford: Oxford University Press, 401–423.
- Grimshaw, J. 2015. The Light Verbs Say and SAY. In: I. Toivonen, P. Csúri, and E. van der Zee (eds.). *Structures in the Mind*. Cambridge, MA/London: MIT Press, 79–99.
- Jenks, P., Konate, R. 2022. Indexed definiteness. *Glossa: a journal of general linguistics*, v. 7, 1–44.
- Kastner, I. 2015. Factivity mirrors interpretation: The selectional requirements of presuppositional verbs. *Lingua*, v. 164, 156–188.
- Knyazev, M. 2016. Complementizers in Kalmyk. In: K. Boye and P. Kehayov (eds.) *Complementizer semantics in European languages*. Berlin: de Gruyter, 655–679.
- Knyazev, M. 2022. SAY-complementizers and indexical shift in Poshkart Chuvash: With emphasis on communicative reception reports. *Studies in Language*, v. 46, 402–452.
- Korotkova, N. 2016. Heterogeneity and uniformity in the evidential domain. PhD diss, UCLA.
- Kožemjakina, A. 2017. Nekotorye dannye o konstrukcijax s sentencial'nym aktantom v čuvašskom jazyke (malokaračkinskij dialekt) [Some data on sentential complement constructions in the Maloe Karachkino dialect of Chuvash]. Paper presented at a fieldwork seminar, Maloe Karachkino, August 2017.
- Kratzer, A. 2006. Decomposing attitude verbs. Ms. <https://semanticsarchive.net/Archive/DcwY2JkM/attitude-verbs2006.pdf>
- Logvinova, N. 2019a. Neposessivnye funkcii pokazatelja posessivnosti tret'ego lica v malokaračkinskom govore čuvaškogo jazyka [Non-possessive functions of the third person possessive in Maloe Karachkino dialect of Chuvash]. *Acta Linguistica Petropolitana*, v. 15, 86–129.
- Logvinova, N. 2019b. Ob odnositel'nyx konstrukcijax i prišastijax v (malokarač kinskom) č uvašskom [On relative constructions and on participles in the Maloe Karachkino dialect of Chuvash]. Paper presented at a fieldwork seminar, Maloe Karachkino, August 2019.
- Logvinova, N. 2021. Zagadka *-nə*: upotreblenie form “pričastija prošedšego vremeni” [The puzzle of *-nə*: the use of the so-called “past participle”]. Paper presented at a fieldwork seminar, Maloe Karachkino, July 2021.
- Major, T. 2021. On the nature of "say" complementation. PhD diss., UCLA.

- Major, T, Sikuku, J., and Diercks. M. 2022. “Say”-complementation: reanalyzing agreeing complementizers as verbs. Ms. <https://research.pomona.edu/mjkd/files/2022/07/Lubukusu-Say-complementation-7.19.2022-to-distribute.pdf>
- Matić, D., Pakendorf, B. 2013. Non-canonical SAY in Siberia: Areal and genealogical patterns. *Studies in Language*, v. 37, 356–412.
- Moulton, K. 2015. CPs: Copies and compositionality. *Linguistic Inquiry*, v. 46, 305–342.
- Özyıldız, D, Major, T., and Maier, E.. 2019. Communicative reception reports as hear-say: Evidence from indexical shift in Turkish. In: R. Stockwell, M. O’Leary, Zh. Xu and Z.L. Zhou (eds.), *Proceedings of the 36th West Coast Conference on Formal Linguistics*, 296–305. Somerville, MA: Cascadilla Proceedings Project.
- Simeonova, V., 2020. The Syntax and semantics of light attitudes. PhD. diss, University of Ottawa.
- Simons, M. 2007. Observations on embedding verbs, evidentiality, and presupposition. *Lingua* v. 117, 1034–1056.