

**NEITHER A LEXICAL VERB  
NOR A COMPLEMENTIZER:  
EVIDENTIAL SAY-BASED COMPLEMENTS  
IN CHUVASH<sup>†</sup>**

Mikhail Knyazev

(Institute for Linguistic Studies, RAS)  
misha.knjazev@gmail.com

The paper discusses a potential challenge to the (full) decompositional analysis of ‘say’-complementizers according to which they are synchronically analyzed as respective non-finite 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* cannot be compositionally interpreted as a nominalization of a lexical verb ‘say’ since this would lead 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) analysis of 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

---

<sup>†</sup> 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. My heartfelt thanks also go to my Chuvash consultants, in particular Roza Nikolaevna Maksimova.

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.<sup>1</sup> For example, as shown in (1a)–(1b), clausal complements of most common speech and belief verbs such as *kala* ‘say’, *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).<sup>2</sup>

- (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* [*van’uk kaj-za te-ze*] *şotl-at*.<sup>3</sup>  
 Masha Vanyuk 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 & 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 until very recently (see Major 2021 for discussion).

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

<sup>1</sup> 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 were obtained in 2017–2023, originally during student linguistic expeditions to Poshkart organized by Masha Kholodilova and HSE University, St. Petersburg & Moscow. In the examples below, I use transcriptions and simplified glosses developed within this project.

<sup>2</sup> 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.

<sup>3</sup> Poshkart Chuvash also has finite uses of the -ZA converb (cf. (1b)), as well as of the -*n*( $\theta$ ) participle (cf. (2) below), to express past tense.

may retain phonological substance and be syntactically analyzable 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 several Turkic languages such as Turkish, Uyghur and Sakha (Özyıldız et al. 2019; Major 2021) as synchronically converbs (see also Knyazev 2016 for Kalmyk (Mongolic)).<sup>4</sup> On decompositional approaches, examples like (1a)–(1b) are analyzed as literally meaning ‘X said/ thought (something), *saying* p’ (see Section 3). Such approaches capitalize on recent proposals in formal semantics arguing that complement clauses (e.g. *that*-clauses) are not arguments of the predicate, as traditionally assumed, but *modifiers* (Kratzer 2006; Moulton 2015; Elliott 2020). The latter proposals are, in turn, based on the observation that *that*-clauses can be used in the predicative function (cf. *The claim is that he is liar*), like modifiers and unlike arguments (Stowell 1982; see also Krapova and Cinque 2016).

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.*<sup>5</sup>  
 hear-PST-1SG  
 ‘I heard from my mother that there has a been a fire in the neighbouring village.’<sup>5</sup>

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

<sup>5</sup> The marker *-n(ə)*, used in *tenine*, as well as in event nominalizations more generally (cf. (3a)), is traditionally referred to (and glossed) as the “past participle” marker.

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 (from Y) that Y/someone said that p’, which gives an intuitively wrong paraphrase (see Section 2). Therefore, in Knyazev (2022) I rejected the decompositional account and instead analyzed *tenine* as a complementizer that has syntactic constraints 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 a reportative evidential light verb SAY, which crucially allows to view the semantic contribution of *tenine* (‘X said that p’) as *presuppositional*.<sup>6</sup>

The paper is structured as follows. In Section 2, I review the distributional and semantic properties of *tenine*, as well as 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.

## 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

---

However, it is more accurately characterized as non-future participle (Logvinova 2021). Because the difference between event nominalizations based on  $-n(\vartheta)$  and on the future participle ( $-As$ ) is not relevant to me, I will simply gloss  $-n(\vartheta)$  as “participle” (Ptcp). Note that only the nominalization of *te* ‘say’ based on the  $-n(\vartheta)$  participle is used in complementizer function.

<sup>6</sup> 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. (See footnote 17.)

an experiencer subject; ii) an explicit (and sometimes implicit) information source argument, typically realized as oblique; and iii) they allow accusative-marked nominalized complements (cf. (3a) and (5a)). 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 *şan/ënen* ‘believe (someone)’ (so-called “response-stance” sense of *believe*, see Bogal-Albritten and Moulton 2018), as in (3c). Speakers 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’.

- (3) a. [*man taga san xajar-san<sup>1</sup>-a {eiz=er-ze*  
 I.GEN sheep you.SG.GEN cucumber-PL-ACC eat=let-CVB  
*te-n-in-e / eiz=er-n-in-e}*] ***elt-r-ëm.***  
 say-PTCP-POSS.3-ACC eat=let-PTCP-POSS.3-ACC hear-PST-1SG  
 ‘I heard that my sheep has eaten your cucumbers.’
- b. *petja [man kozak sumاران-za kaj-za te-n-in-e]*  
 Petya I.GEN cat get.sick-CVB go-CVB say-PTCP-POSS.3-ACC  
***pël-te-ë.***  
 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<sup>2</sup>-na şan-d-ëm.<sup>7</sup>*  
 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,

<sup>7</sup> In Chuvash, direct and indirect objects are marked by the same marker *-A*, which is sometimes glossed as the object marker (OBJ). For simplicity, I gloss this marker as accusative (ACC) throughout this paper.

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 (they require nominalized clauses instead).

- (4) a. \**maʃə* [ʃomər 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. \**maʃə* [vanʲuk kaj-za te-n-in-e] ʃotl-at.  
 Masha Vanyuk go-CVB say-PTCP-POSS.3-ACC say-NPST[3SG]  
 Intended: ‘Masha thinks that Vanyuk has left.’
- c. \**[es pørt tu-za lart-sa te-n-in-e] peʔə 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 the 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* proposed above. For example, we can see that *tenine* occurred, by far, most frequently with ‘hear’, whereas other predicates listed in the table 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).<sup>8</sup>

## 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 con-

<sup>8</sup> The presence of the verb *pëlter* ‘mean’ in this list, as well as its relatively high frequency, may appear surprising, but note that it is a causative of *pël* ‘know’ and thus may involve an implicit experiencer, as in ‘cause X to know that p (≈ suggest 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.

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

<i>ilt</i>	'hear'	110
<i>pëlter</i> ( <i>pël</i> -CAUS)	'mean, indicate (to someone)'	19
<i>ënen</i>	'believe (someone)'	16
<i>astu</i>	'remember'	15
<i>pël</i>	'know, learn'	14
<i>ənlan</i>	'understand'	12
<i>ənkär</i>	'guess'	12
<i>asa il</i>	'take in mind' (= 'remember')	11
<i>asra tit</i>	'keep in mind' (= 'remember')	12
<i>vula</i>	'read'	12
<i>jışan</i>	'admit'	12
<i>şirëplet</i>	'confirm'	12
<i>şan</i>	'believe'	12
<i>ənlantar</i>	'explain'	12
<i>kălar</i>	'discard'	11
<i>man</i>	'forget'	11
<i>păxmasărax</i>	'disregarding'	11
<i>përex</i>	'equals (to)'	11
<i>purnaşa kërť</i>	'put into life'	11
<i>tyrre kălar</i>	'put correctly' (= 'justify?')	11
<i>tërësle</i>	'make sure'	11
<i>şuta il</i>	'take into account'	11
<i>şuxăşla</i>	'assume'	11
<i>itla jivăra il</i>	'take too gravely' (= 'take to heart')	11

tribution. This can be shown by looking at verbs *pël* 'know' and *astu* 'remember' (and also *ənlan* 'understand'). As can be seen from the translations in (3b) and (3d), 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 until 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 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 požar pol-n-in-e]  
 grandma neighbor village-LOC fire be-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.’
- b. #*papi* [jonazar jal-da požar 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 (she was told so).’

In contrast to verbs like *pël* ‘know’, with verbs like *elt* ‘hear’ and *šan* ‘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, I assume that this is also in principle possible.

Incidentally, observe that the CA implication specifically targets the information source (and, accordingly, 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). Another related observation is that a cognitive factive verb *pël* ‘know’ retains a factivity implication when taking a *tenine*-clause, which can be shown by the fact that it is infelicitous when the speaker takes the proposition to be false, as in (7a), cf. *elt* ‘hear’ in (7b).<sup>9,10</sup>

<sup>9</sup> 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



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

*peʃə man kozak sumaran-za kaj-za te-n-in-e*

Petya I.GEN cat get.sick-CVB go-CVB say-PTCP-POSS.3-ACC

*pəl-te-ë / elt-r-ë.*

know-PST-3SG hear-PST-3SG

'Petya learned/heard 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. *#peʃə man kozak sumaran-za kaj-za te-n-in-e*

Petya I.GEN cat get.sick-CVB go-CVB say-PTCP-POSS.3-ACC

*pəl-te-ë.*

know-PST-3SG

'Petya learned that my cat has got sick (he was told so).'

b. *peʃə man kozak sumaran-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 may be taken to be the assumption of the decompositional approach. In Knyazev (2022), I presented arguments against this view, which I reproduce below in a more elaborate form. (Note that these arguments may not be automatically extended from one embedding 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),

---

speaker, but in such cases 'know' (and other factive verbs) may lose its factivity, as was shown by Simons (2007).

<sup>10</sup> 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 Poshkart Chuvash is generally nonfactive.

- (i) Context: = (7a)

*[peʃə man kozak sumaran-za kaj-za te-n-in-e / kaj-n-in-e]*

Petya I.GEN cat get.sick-CVB go-CVB say-PTCP-POSS.3-ACC go-PTCP-POSS.3-ACC

*asta-t.*

remember-NPST[3SG]

'Petya remembers that my cat has got sick (he was told so).'

according to which what is embedded is the proposition ‘*X/someone said that p*’, rather than *p* itself.<sup>11</sup> A possible exception is the verb *astu* ‘remember’ in (3d) but even in this case the simpler paraphrase appears more natural. While such paraphrases are in principle possible, they crucially 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).<sup>12</sup>

- (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))

A second argument against the decompositional approach 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).<sup>13</sup>

<sup>11</sup> Another conceivable paraphrase for sentences like (3a) would be ‘I heard someone/people *say* that *p*’, involving direct auditory perception. However, such paraphrase would a priori work only for the verb *elt* ‘hear’ but not other verbs that occur with *tenine*. Moreover, it would be incompatible with *elt* ‘hear’ with an ablative source (cf. (2)). In any event, this paraphrase is still rather implausible for (3a) since speakers never translate the corresponding sentences (or describe their meanings) using direct perception complements, realized in Russian with the complementizer *kak*. But I must leave the discussion of the difference between direct auditory perception vs. communicative reception reports (‘hear that’) in Chuvash for another occasion.

<sup>12</sup> 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. However, at least two examples do not contain an explicit predicate (cf. Knyazev 2022:416, (13c)).

<sup>13</sup> 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 *ʂan* ‘believe (someone)’ speakers sometimes use or at least ac-

While it would be wrong to assume that *tenine* is strictly disallowed in second-hand reports (see (10) below), such uses are at best marginal, weakening the premise of the decompositional analysis.

- (9) *maʃə* [*ʃomər bol-at* {*te-ze kala-n-in-e* /  
 Masha rain be-NPST[3SG] say-CVB say-PTCP-POSS.3-ACC  
 \**te-n-in-e*}] *kala-r-ë*.  
 say-PTCP-POSS.3-ACC say-PST-3SG  
 ‘Masha said that someone said it will rain.’

A third, and related, argument is that constructions involving *tenine* with an overt subject (which would uncontroversially signal its status as a nominalization of ‘say’), as in (10a)–(10b), are generally dispreferred (compared to *teze + kalanine*) and at least with verbs *elt* ‘hear’ and also *pël* ‘know’ (though perhaps not with *ʃan* ‘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.

- (10) a. *aməʃ* [*ul’-ə pilək il-e-p te-n-in-e*]  
 mother.POSS.3 son-POSS.3 five take-NPST-1SG say-PTCP-POSS.3-ACC

cept *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 ɛz=er-ze te-ze*  
 L.GEN sheep you.SG.GEN cucumber-PL-ACC eat=let-CVB say-CVB  
*kala-n-in-e*] *elt-r-ëm*.  
 say-PTCP-POSS.3-ACC hear-PST-1SG  
 ‘I heard that (allegedly) my sheep has eaten your cucumbers.’ (bracketed material not included in the original stimulus)
- b. [[*ul’-ə pilək il-et te-ze*] *utçitel’ kala-n-in-e*]  
 son-POSS.3 five take-NPST[3SG] say-CVB teacher say-PTCP-POSS.3-ACC  
*aməʃ ʃan-at*.  
 mother.POSS.3 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)

*şan-te-ë.*

believe-PST-3SG

i. 'Mother believed that her son<sub>i</sub> will get an A in the exam (as he<sub>i</sub> told her).'

ii. 'Mother believed that her son<sub>i</sub> said (to someone) that he<sub>i</sub> will get an A.'

b. *amaş* [ul'-ə pilëk 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 son<sub>i</sub> will get an A (as he<sub>i</sub> told her).'

ii. 'Mother heard (from someone) that her son<sub>i</sub> said that he<sub>i</sub> will get an A.'

c. *uteitel'* [vas'ə pilëk 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}] amaş elt-r-ë / pël-te-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 it be treated as an unanalyzed complementizer, as I proposed in Knyazev (2022)? The main problem with the latter analysis 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 *kilëş* '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

*tërës mar.*

true NEG

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

b. *peʔə [maʂə lajək {jorl-at           te-n-i-be           /*

Petya Masha good sing-NPST[3SG] say-PTCP-POSS.3-INS

*jorla-n-i-be}] kilëš-r-ë.*

sing-PTCP-POSS.3-INS 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 *tenë* or *tegen*, which are morphologically participles of *te* ‘say’, as shown in (12).<sup>14</sup>

(12) a. [*jonaz̄ar jal-da       požar 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))

<sup>14</sup> Note that *tenë* contains the same participle marker as found in *tenine* (*teni* and *tenibe*). However, the two have different morphosyntactic properties, since participles do not take case or possessive morphology and occur adnominally. It is plausible to analyze this as a case of morphological underspecification. An alternative view is that *-n(ə)* 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).

- b. [es     *kate-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*, the 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’.<sup>15</sup> At the 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 kate-a tok-sa te-ze*] *tërës 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 kate-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))

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*

<sup>15</sup> 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     *kate-a*     *kaj-za*] *t-et.*  
 news you.SG groom-ACC go-CVB say-NPST[3SG]  
 Intended: ‘The rumor says that you got married.’

*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 above paradox arises most clearly with nominalized forms such as *tenine* because for *converbial* forms of 'say' a decompositional account may still be a theoretical option (cf. Major 2021, see Section 3.6) in view of 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* [van'uk kaj-za te-ze] ʂotl-at.  
 Masha Vanyuk 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 conditions, in contrast to the case of *tenine*-clauses (cf. (8)).

### 3. Analysis

#### 3.1. Proposal in a nutshell

The main idea of the proposed account is that 'say' in 'say'-complementizers should be viewed as a *functional* or *light* (as opposed to lexical) verb. Specifically, I adopt Simeonova's (2020) identification of 'say'-complementizers with *reportative evidential* markers. What Simeonova proposes is that reportative evidentiality should be analyzed as a special kind of clausal embedding where the complement clause is embedded by a light verb SAY, realized as a functional head in the left periphery of the clause. Intuitively, the analysis is made plausible by the observation that reportative evidential statements can be

typically paraphrased using the verb ‘say’, as in ‘they say’, etc. The empirical motivation for Simeonova’s analysis comes from the fact that reportative evidential markers often develop from non-canonical ‘say’ (cf. Matic’ and Pakendorf 2013: 377). Such evidential uses of ‘say’ are also attested in Chuvash, as in (15).

- (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))

Crucially, the meaning component of SAY corresponding to the saying event is *presuppositional* and is not part of its *assertive* / *truth-conditional* content, which is treated as rather abstract and impoverished. This will be important for my analysis of *tenine* (and similar forms) because it allows to circumvent the problem of implausible paraphrases involving second-hand reports (see Section 2.2). At the same time, because the *presuppositional* content of SAY still implicates the existence of a communicative act, the analysis allows to capture the CA implication associated with *tenine*-clauses.

A further important aspect of Simeonova’s analysis is that it covers both matrix/unembedded as well as *embedded* evidentials. In fact, I will have little to say about the former, as in (15), but will instead focus on Simeonova’s analysis of the latter, which I will adopt for ‘say’-based complements in Chuvash, especially for *tenine*-clauses.

I will depart from Simeonova in two main respects. First, I will not follow Simeonova’s assumption that reportative evidential SAY corresponds to the *highest* projection of the embedded clause (specifically, a small *c* above CP) as this is hard to reconcile with the fact that in Chuvash *te* ‘say’/SAY in complementizer-like forms is followed by nonfinite morphology (see Section 2.3), standardly analyzed as generated above the verb in head-final languages, as in (16).

- (16) ...<sub>Clause</sub>] SAY] NMZ/PTCP/CVB]

Second, and more importantly, I will adopt Simeonova’s analysis *only for nominalized and participial* forms of ‘say’ (such as *tenine*, *teni*, *tenibe*, *tenë* and *tegen*) but not for the converbial form *teze*, which appears not to presuppose the existence of a saying event.

Before presenting Simeonova’s analysis, I will introduce the semantic framework in which it can be formally implemented.



3.2. *Semantic framework*

The proposed account follows recent proposals that treat complement clauses, traditionally analyzed as denoting propositions (type  $\langle s,t \rangle$ ), as instead denoting predicates (type  $\langle e,t \rangle$ ), more specifically predicates of *contentful individuals*, or *individuals with content* (represented as  $x_c$ ), i.e. what nouns like *rumor*, *claim*, *news*, etc. refer to (Kratzer 2006; Moulton 2015; Elliott 2020; Bondarenko 2022). The motivation behind analyzing complement clauses this way is that it allows to capture the fact that at least in English, *that*-clauses can be directly predicated of content nouns, as in *The claim is that there was a fire*. The predicate meaning of complement clauses is compositionally derived by way of combining the embedded proposition with an operator that takes a proposition and returns the set of contentful individuals whose content is specified by that proposition, where the content is recovered by the built-in function *CONT*. For concreteness, I assume that the said operator is realized by the complementizer (e.g. *that*) (but see Elliott 2020), as in (17a). On this view, *that there was a fire* denotes the set of contentful individuals whose content is that there was a fire, as in (17b).

- (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.  $\llbracket \text{that there was a fire} \rrbracket = \lambda x_c. [\text{CONT}(x) = \lambda w. \text{there was a fire in } w]$

How do complement clauses combine with attitude verbs on the predicate analysis? To address this question, I adopt a distinction in the literature between *presuppositional* and *nonpresuppositional* verbs (Kastner 2015; Bochnak & Hanink 2022). Nonpresuppositional verbs like *say* or *think* typically introduce new information into the discourse. By contrast, presuppositional verbs, including factive verbs like *know*, *remember* or *regret* and response-stance verbs (e.g. Hegarty 1992) like *agree*, *deny* or *doubt*, refer to old or familiar information which had been previously introduced into the discourse and which is presupposed to be part of the Common Ground. This distinction corresponds to how these two classes of verbs compose with their complements.

With nonpresuppositional verbs, the complement clause (type  $\langle e,t \rangle$ ) essentially functions as a modifier or restrictor of the contentful individual argument. For concreteness, I assume that the complement clause composes with the verb via the operation *Restrict* (Chung & Ladusaw 2004, see also Srinivas & Legendre 2022) with subsequent existential closure of the argument, as in (18). For example, the vP in (18c) would denote a set of thinking

events with John as their holder and some contentful individual as their theme such that this individual has the proposition that there was a fire as its content.

- (18) a.  $[[\text{think}]] = \lambda x_c. \lambda y. \lambda e. [\text{thinking}(e) \wedge \text{theme}(e) = x \wedge \text{holder}(e) = y]$   
 b.  $[[\text{think that there was a fire}]] = \lambda y. \lambda e. \exists x_c. [\text{thinking}(e) \wedge \text{theme}(e) = x \wedge \text{CONT}(x) = \lambda w. [\text{there was a fire in } w] \wedge \text{holder}(e) = y]$   
 c.  $[[\text{John think that there was a fire}]] = \lambda e. \exists x_c. [\text{thinking}(e) \wedge \text{theme}(e) = x \wedge \text{CONT}(x) = \lambda w. [\text{there was a fire in } w] \wedge \text{holder}(e) = \text{John}]$

By contrast, with presuppositional verbs the complement clause saturates the argument position of the verb in the usual manner, but prior to that it is nominalized and turned into an  $\langle e \rangle$ -type expression (otherwise it would not be able to compose with the verb). The nominalization is performed by a (possibly silent) anaphoric definite D, which takes some predicate P of contentful individuals and returns a unique contentful individual in the discourse that satisfies P. For concreteness, I assume that anaphoric D takes an additional index argument fixing the identity of this unique individual (Jenks and Konate 2022), as in (19a)–(19b). For example, the vP in (19d) would denote a set of knowing events whose holder is John and whose theme is the unique contentful individual which has the proposition that there was a fire as its content and which is identical to some specific contentful individual in the discourse (e.g. bearing index 8).

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

Given this semantic framework, we can now present Simeonova's (2020) analysis of reported evidentials.

### 3.3. Simeonova's analysis of reportative evidential 'say'

Simeonova's analysis of reportative evidential SAY is given in (20a). It is very similar to Moulton's (2015) analysis of the English complementizer *that* in

(17a), with two differences.<sup>16</sup> First, SAY takes an implicit holder argument in its specifier, whose reference is supplied by the context (see (20b) below). Second, and more importantly, SAY introduces the presupposition that there was a saying event with some contextually supplied individual (in the specifier of SAY) as its agent and the embedded proposition as its theme. For example, a SAY-based complement in (20b) would denote a set of contentful individuals whose content is that there was a fire (cf. (17b)) and would also presuppose that someone in the discourse (e.g. bearing index 7) said that there was a fire.

- (20) a.  $\llbracket \text{SAY} \rrbracket = \lambda p_{\langle s,t \rangle} \lambda y. \lambda x_c. [\text{CONT}(x) = p]$  (adapted from Simeonova 2020: 233, (224))  
 Defined only if  $\exists e. [\text{saying}(e) \wedge \text{theme}(e) = x \wedge \text{agent}(e) = y]$
- b.  $\llbracket \text{pro}_7 \text{ SAY } [\text{there was a fire}] \rrbracket^g = \lambda x_c. [\text{CONT}(x) = \lambda w. \text{there was a fire in } w]$   
 Defined only if  $\exists e. [\text{saying}(e) \wedge \text{theme}(e) = x \wedge \text{agent}(e) = g(7)]$

Simeonova's analysis of embedded evidentials with nonpresuppositional predicates 'say' and 'think', based on Bulgarian data, is given in (21) (Simeonova does not discuss embedding under other predicates). As can be seen, the analysis is very similar to the analysis of the corresponding examples with English *that*-clauses in (18c) except for the presence of the presupposition.

- (21) a.  $\llbracket \text{Zlati}_7 \text{ say } [\text{pro}_7 \text{ SAY } [\text{there was a fire}]] \rrbracket^g$   
 $= \lambda e. \exists x_c. [\text{saying}(e) \wedge \text{theme}(e) = x \wedge \text{CONT}(x) = \lambda w. [\text{there was a fire in } w] \wedge \text{holder}(e) = \text{Zlati}]$   
 Defined only if  $\exists e. [\text{saying}(e) \wedge \text{theme}(e) = x \wedge \text{agent}(e) = \text{Zlati}]$
- b.  $\llbracket \text{Zlati}_7 \text{ think } [\text{pro}_7 \text{ SAY } [\text{there was a fire}]] \rrbracket^g$   
 $= \lambda e. \exists x_c. [\text{thinking}(e) \wedge \text{theme}(e) = x \wedge \text{CONT}(x) = \lambda w. [\text{there was a fire in } w] \wedge \text{holder}(e) = \text{Zlati}]$   
 Defined only if  $\exists e. [\text{saying}(e) \wedge \text{theme}(e) = x \wedge \text{agent}(e) = \text{Zlati}]$   
 (adapted from Simeonova 2020: 239, (234))

<sup>16</sup> Simeonova's (2020) actual analysis for SAY involves existential quantification of the  $x_c$  argument (which yields a  $t$ -type denotation) to capture matrix evidentials. To achieve the meaning in (20a), which yields an  $\langle e, t \rangle$ -type denotation and thereby can capture *embedded* evidentials, Simeonova uses the special operation of existential disclosure. Because I am mostly concerned with embedded evidentials, I directly adopt the meaning in (20a) for SAY.

One may also note that the presupposition in (21) repeats or reinforces the meaning of the matrix verb, which fits with the semantic characterization of embedded evidentials in the literature (cf. Korotkova 2016), but Simeonova also argues that this presupposition places certain constraints on examples like (21) (see Section 3.5).

I will postpone the discussion of whether the analysis in (21) is appropriate for the corresponding examples with *teze*-clauses in Chuvash until Section 3.5. In the next section, I will show how this analysis can account for the properties of *tenine*-clauses.

### 3.4. Analysis of *tenine*-clauses

I assume that the set of verbs that combine with *tenine*, namely *elt* ‘hear’, *pël* ‘know, learn’, *şan* ‘believe (someone)’, *anlan* ‘understand’, *astu* ‘remember’ and others (see Section 2.1) belong to the presuppositional class. Thus, under the assumptions in Section 3.2, they should compose with the complement clause along the lines of (19). This allows to provide the following compositional analysis of *tenine*-clauses.

As suggested in Section 3.1, I identify *te-* with the reportative evidential SAY in (20a). I also assume that *-n(ə)-* in *tenine*, which is glossed as Ptcp, functions as a nominalizer (Nmz) and is semantically inert, i.e. an identity function (for simplicity, I ignore its denotation below). The contribution of *-i(n)-* in *tenine* (which is glossed as a third person possessive marker (Poss.3)), is less straightforward. But, as was showed by Kozhemyakina (2017), this marker in nominalizations does not track the person and number of the subject but is lexicalized (obligatory) in the nominative, accusative and the instrumental case. Thus, I will assume that *-i(n)-* realizes anaphoric D in (19a), which fits with its main function of expressing familiarity in ordinary noun phrases (Logvinova 2019a), as in (22).

- (22) *kəʒal ebë kërëk 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 (23) (repeated from (7b)) can be analyzed as in (24a). When combined with an embedding verb such as *elt* ‘hear’, this gives roughly the meaning in (24b) (disregarding tense): the sentence is true if there is a hearing event whose experiencer is Petya and whose theme is the unique contentful individual in the discourse

(bearing the index 8) which has as its content the proposition that my cat got sick. The sentence also presupposes that some individual in the discourse (here, Masha) said that my cat got sick. On this analysis, the meaning of sentences like (23) should come out as roughly synonymous with ‘Petya heard the rumor/claim/news that my cat got sick’, which seems to fit with native speakers’ intuitions about their meaning. A similar analysis can be given to sentences with verbs *pël* ‘know, learn’, *şan* ‘believe (someone)’ and others.

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

*peʃə 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.’

- (24) a.  $[[[DP\ 8\ [NP\ [pro_7\ [my\ cat\ has\ got\ sick]\ te_{SAY}]\ -n_{Nmz}]\ -in_e_D]]]^g$   
 $= \iota_{x_c}.[CONT(x) = \lambda w.[my\ cat\ got\ sick\ in\ w] \wedge x = g(8)]$   
 Defined only if  $\exists e.[saying(e) \wedge theme(e) = x \wedge agent(e) = Masha]$
- b.  $[[ (23) ]]^g = \exists e.[hearing(e) \wedge theme(e)]$   
 $= \iota_{x_c}.[CONT(x) = \lambda w.[my\ cat\ has\ sick\ in\ w] \wedge x = g(8)] \wedge experiencer(e) = Petya$   
 Defined only if  $\exists e.[saying(e) \wedge theme(e) = x \wedge agent(e) = Masha]$

The main advantage of the analysis in (24) is that because the existence of a saying event is not part of the asserted content it does not run into implausible paraphrases like (8a) (see Section 2.2). At the same time, the analysis correctly captures the CA implication with *tenine*-clauses because in order to satisfy the presupposition in question the sentence must occur in a context where the relevant propositional content was communicated by someone. As I suggested in Section 2.2, the presuppositional contribution of *tenine* becomes particularly relevant for the analysis of not inherently communicative verbs like *pël* ‘know, learn’ (and also *astu* ‘remember’ and *ənlan* ‘understand’) (cf. (3b) and (3d)), where the CA implication cannot be viewed as part of the lexical meaning of the verb (as opposed to verbs like *elt* ‘hear’ and *şan* ‘believe (someone)’) but must instead specifically arise from the use of *tenine*-clauses.<sup>17</sup>

<sup>17</sup> T. Bondarenko (p.c.) raises an interesting possibility that the presuppositional component of the meaning of SAY is redundant given that *tenine* refers to a familiar contentful individual, which may itself implicate the existence of a saying event. In other words, SAY should have a uniform analysis along the lines of (28b) below. In her (2022) dissertation, Bondarenko also argues (based on similar data in Buryat)

The assumption that the meaning component corresponding to the existence of a saying event indeed has a presuppositional status is suggested by the fact that the CA implication can survive under negation, as shown by (25).

- (25) Context: Someone said that there was a fire in the neighbouring village.  
*papi [jonazar jal-da pozar pol-za te-n-in-e]*  
 grandma neighbor village-LOC fire be-CVB say-PTCP-POSS.3-ACC  
*elt-m-en.*  
 remember-NEG-PTCP.RES  
 ‘Grandma did not hear that there was a fire in the neighbouring village  
 (but someone said so).’

Note that this presupposition can sometimes be suspended, as shown by the felicity of (25) in a context like ‘there is a survey asking people if they heard of any fires in neighboring villages lately’, suggesting that SAY is a *soft trigger* (cf. Abusch 2002, Abrusán 2011) with respect to the presupposition in question.

### 3.5. Extending the analysis to *teze*-clauses?

Whereas the analysis of SAY in (20) works relatively well with *tenine*-clauses, it is less clear whether it can be extended to *teze*-clauses, i.e. to complements of nonpresuppositional predicates (cf. (21)). Assuming that the converbial marker -ZA is semantically inert, for sentences like (1b) the analysis would predict the compositional meaning roughly as in (26): there is a saying event whose holder is Masha and whose theme is some contentful individual which has as its content the proposition that it will rain. The sentence also (vacuously) presupposes that some individual (here, Masha) said that it will rain.

---

that apparent projection of the saying event out of negation, as in (25) below, could be derived by having the nominalized clause take the wide scope over negation, without a built-in presupposition. One way to test this analysis would be to look at contexts where a *tenine*-clause refers to familiar/given content that was thought by someone but was not communicated, e.g. ‘Masha thought that my cat has got sick. Petya also believes/knows that my cat has got sick (even though no one told him so).’ As far as I can see, Bondarenko’s analysis predicts that *tenine* should be possible in the second sentence. My hunch is that this is infelicitous, leading to an over-generation problem, but this of course must be checked with Chuvash speakers. (See also Moulton et al. 2020: 14, where the same problem is raised for a comparable analysis of Korean/Japanese nominalized clauses.)



The analysis in (28b) correctly predicts that the meaning of *teze*-clauses is very similar to English *that*-clauses (with nonpresuppositional verbs), which seems intuitively correct.

#### 4. Conclusion and implications for theories of evidentiality

The central puzzle of this paper is the paradoxical status of ‘say’-complementizers like *tenine*: they seem to be syntactically analyzable as forms of the verb ‘say’ and also implicate the existence of a saying event (CA/communicative act implication), but at the same time they cannot be analyzed as full-fledged nominalizations of a lexical verb ‘say’, as this leads to incorrect paraphrases.

The solution proposed in this paper is to decompose *tenine* and other nominalization/participle-based ‘say’-complementizers as forms of the light verb SAY encoding reportative evidentiality (Simeonova 2020). The crucial assumption is that reportative evidential SAY does not refer to a saying event, thereby avoiding the problems with implausible paraphrases, but still *presupposes* the existence of a saying event, thereby capturing the CA implication.

An intuitive justification for treating *tenine*-clauses as evidential is that they specify the attitude holder’s information source of the embedded proposition as hearsay. This seems to fit the standard characterization of reportative evidentials. However, this potentially clashes with Korotkova’s (2016) generalization, namely that embedded evidentials must specify the information source of the *actual speaker* when they take the matrix scope (i.e. given from the perspective of the actual speaker, as opposed to the attitude holder). For example, on Korotkova’s (2016:138–139) view, embedded evidentials can only have meanings like ‘Jay said that—and *I’ve heard it*—Anna got a puppy’ but not ‘Jay said that—and *he has heard it*—Anna got a puppy’.<sup>18</sup> But the latter is precisely what we see in the case of *tenine*-clauses, which, despite taking the matrix scope (see (29c) below), express the information source of the *attitude holder* rather than the actual speaker (except of course for cases where the attitude holder coincides with the actual speaker). This is shown in (29) (cf. (6)).

<sup>18</sup> Conversely, embedded evidentials that take the embedded scope (i.e. given from the perspective of the attitude holder) must specify the information source of the attitude holder, thus allowing meanings like ‘Jay said that, as he has heard, Anna got a puppy’ but disallowing meanings like ‘Jay said that, as he thinks I’ve heard, Anna got a puppy’ (Korotkova 2016:138–139).



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

*petʃə* [*man kozʌk sumarlʌn-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

- a. ✓ 'Petya heard that my cat has got sick (he was told so).' (subject-oriented, matrix scope)
- b. # 'Petya heard that my cat has got sick (I was told so).' (speaker-oriented, matrix scope)
- c. # 'Petya heard that, as I/he was told, my cat has got sick.' (embedded scope)

Because Korotkova (2016) mainly discusses embedding of evidentials under verbs like 'say' or 'think' (but not under presuppositional verbs like 'hear' or 'know'), it remains unclear whether *tenine*-clauses constitute a genuine counterexample to her generalization (or perhaps a reason to reject the evidential analysis of *tenine*-clauses in the first place).<sup>19</sup> This question, as well as a more general theoretical justification for the evidential analysis of *tenine*-clauses, remains for future work.

## References

- Abrusán, Márta. 2011. Predicting the presuppositions of soft triggers. *Linguistics and Philosophy*, v. 34, 491–535.
- Abusch, Dorit. 2002. Lexical alternatives as a source of pragmatic presuppositions. In: B. Jackson (ed.) *Proceedings of SALT 12*, 1–19. Ithaca, NY: Cornell University, CLC Publications.
- Bochnak, M. Ryan & Emily A. Hanink. 2022. Clausal embedding in Washo: Complementation vs. modification. *Natural Language and Linguistic Theory*, v. 40, 979–1022.

---

<sup>19</sup> Note that Korotkova's (2016) generalization cannot be stated in Simeonova's (2020) framework, where the person whose information source is referred to by the evidential meaning (so-called ORIGO) is not explicitly represented but is determined pragmatically. 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). Accordingly, the corresponding problem with *tenine*-clauses does not arise in Simeonova's framework, which might provide a potential argument in favor of Simeonova's account of evidentiality compared to Korotkova's account.

- Bogal-Allbritten, Elizabeth & Keir Moulton. 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*, 215–232. Edinburgh: University of Edinburgh Press.
- Bondarenko, Tatiana I. 2022. *Anatomy of an attitude*. PhD Thesis. Cambridge MA: MIT.
- Chung, Sandra & William A. Ladusaw. 2004. *Restriction and saturation*. Cambridge MA: MIT press.
- Driemel, Imke & Maria Kouneli. 2022. *C-Agree is local subject-verb agreement in Kip-sigis*. Manuscript. [lingbuzz/005661](https://lingbuzz/005661)
- Elliott, Patrick D. 2020. Elements of clausal embedding. PhD Thesis. London: University College London.
- Hegarty, Michael Vincent. 1992. *Adjunct extraction and chain configurations*. PhD Thesis. Cambridge MA: MIT.
- Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd & Heiko Narrog. 2010. Grammaticalization and linguistic analysis. In: B. Heine, H. Narrog. (eds.). *The Oxford handbook of linguistic analysis*, 401–423. Oxford: Oxford University Press.
- Grimshaw, Jane. 2015. The Light Verbs Say and SAY. In: I. Toivonen, P. Csúri, and E. van der Zee (eds.). *Structures in the Mind*, 79–99. Cambridge, MA/London: MIT Press.
- Jenks, Peter & Rassidatou Konate. 2022. Indexed definiteness. *Glossa: a journal of general linguistics*, v. 7, 1–44.
- Kastner, Itamar. 2015. Factivity mirrors interpretation: The selectional requirements of presuppositional verbs. *Lingua*, v. 164, 156–188.
- Knyazev, Mikhail. 2016. Complementizers in Kalmyk. In: K. Boye and P. Kehayov (eds.). *Complementizer semantics in European languages*, 655–679. Berlin: de Gruyter.
- Knyazev, Mikhail. 2022. SAY-complementizers and indexical shift in Poshkart Chuvash: With emphasis on communicative reception reports. *Studies in Language*, v. 46, 402–452.
- Korotkova, Natalia. 2016. *Heterogeneity and uniformity in the evidential domain*. PhD Thesis. Los Angeles: UCLA.
- Kožemjakina, Anastasia. 2017. Nekotorye dannye o konstrukcijax s sentencijal'nym ak-tantom 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.
- Krapova, Iliana & Guglielmo Cinque. 2016. On noun clausal “complements” and their non-unitary nature. *Annali di Ca' Foscari. Serie Occidentale*, v. 50, 77–107.
- Kratzer, Angelika. 2006. Decomposing attitude verbs. Manuscript. <https://semanticsarchive.net/Archive/DcwY2JkM/attitude-verbs2006.pdf>

- Logvinova, Natalia. 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, Natalia. 2019b. Ob odnositel'nyx konstrukcijax i prišastijax v (malokaračkinskom) čuvašskom [On relative constructions and on participles in the Maloe Karachkino dialect of Chuvash]. Paper presented at a fieldwork seminar, Maloe Karachkino, August 2019.
- Logvinova, Natalia. 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, Travis. 2021. *On the nature of “say” complementation*. PhD Thesis. Los Angeles: UCLA.
- Major, Travis, Justine Sikuku & Michael Diercks. 2022. “Say”-complementation: reanalyzing agreeing complementizers as verbs. Manuscript.
- Matić, Dejan & Brigitte Pakendorf. 2013. Non-canonical SAY in Siberia: Areal and genealogical patterns. *Studies in Language*, v. 37, 356–412.
- Moulton, Keir. 2015. CPs: Copies and compositionality. *Linguistic Inquiry*, v. 46, 305–342.
- Moulton, Keir, Elizabeth Bogal-Allbritten & Junko Shimoyama. 2020. Things we embed. Talk given at the Brussels Conference on Generative Linguistics (BCGL) 13. <http://individual.utoronto.ca/keirmoulton/MoultonBCGL.pdf>.
- Özyıldız, Deniz, Travis Major & Emar Maier. 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, Vesela T. 2020. *The Syntax and semantics of light attitudes*. PhD Thesis. Ottawa: University of Ottawa.
- Simons, Mandy. 2007. Observations on embedding verbs, evidentiality, and presupposition. *Lingua* v. 117, 1034–1056.
- Srinivas, Sadhwi & Geraldine Legendre. 2022. Does D select the CP in light verb constructions? A reply to Hankamer & Mikkelsen (2021). *Linguistic Inquiry*, 1–27. [https://doi.org/10.1162/ling\\_a\\_00477](https://doi.org/10.1162/ling_a_00477).
- Stowell, Timothy A. 1982. *Origins of phrase structure*. PhD Thesis. Cambridge MA: MIT.