Non-canonical SAY with indexical shift in Poshkart Chuvash

Between verb and complementizer

Abstract

Although the development of generic verbs of speech (SAY) into complementizers is often viewed as a classic example of grammaticalization, recent research has drawn attention to "non-canonical" uses of SAY, which are not fully compositional, yet do not show clear signs of grammaticalization. This leaves the status of particular forms as syntactically decomposable uncertain. I explore this issue with reference to the distribution of the two instances of non-canonical SAY in Poshkart Chuvash (its converb *teze*, used with verbs of speech and thought, and its participial nominalization *tenine*, used with verbs of hearing). Focusing on complements with indexical shift, I show that *teze* and *tenine* differ in the choice of the "controller" for the shifted first person. I explore and reject a decompositional account of this pattern, arguing that the controller must be directly encoded as part of the construction, suggesting that "verb vs. complementizer" might be a false dichotomy.

Keywords: Turkic, verbs of saying, complementizer, indexical shift, constructionist approaches

1. Introduction

Generic verbs of speech (SAY) are commonly identified as an important source of complementizers. For example, SAY-derived complementizers have been found in a number of unrelated language families (see Heine & Kuteva 2002:261ff, 2007:236ff; Chappell 2008; Güldemann 2008; Matić & Pakendorf 2013, a.o.). Although it has become commonplace to view this process in terms of grammaticalization, Matić & Pakendorf (2013) have recently shifted the focus from "grammaticalized SAY" to *non-canonical* SAY, based on the fact that between fully compositional, or *canonical*, uses, on the one hand, and fully grammaticalized uses (i.e., non-compositional, accompanied by changes in inflection, phonological erosion and implying a word class change), on the other, there is a full spectrum of *conventionalized* but not fully grammaticalized uses of SAY.[1] Matić & Pakendorf illustrate such a use with the Sakha marker *dien*

---

[1] Matić & Pakendorf (2013) understand conventionalized constructions in terms of Heine & Kuteva's (2005: 4) *use patterns*, "recurrent pieces of discourse associated with the same grammatical meaning", which serve as early stages of grammaticalization. Perhaps it would also be consistent with Matić & Pakendorf's overall approach to view

in (1), historically a converb of SAY, which functions as a subordinator for reason clauses. Although in (1) *dien* does not express a speech event and thus cannot be considered "canonical", it is nevertheless not fully grammaticalized as it can in principle be viewed as syntactically a converb of SAY (in the 'internal monologue' sense) which has become conventionalized to express reason clauses (see also Chisarik & van der Wurff 2003).

(1) Sakha

Elbeχ   oɣo-nu   kïaj-an   kör-ü.   huoχ-pun   die-n
many   child-ACC   be.able-PF.CVB   see-FUT.3SG   neg-PRED.1SG   say-PF.CVB

akkaːs-ta-n-an   olor-obun.
refusal-VR-REFL-PF.CVB   sit-PRS.1SG

'Saying that I will not be able to look after many children, I refuse.'

'Since I will not be able to look after many children, I refuse.'

(Matić & Pakendorf 2013: 362)

Although Matić & Pakendorf (2013) use the term "non-canonical" primarily as a heuristic for their areal-typological analysis of languages of Siberia, I wish to show that their perspective can shed light on the synchronic analysis of non-canonical SAY in an individual language. Specifically, I will show that the attempt to determine the status of an instance of non-canonical SAY as a verb vs. complementizer in a given language, can lead to apparent paradoxes, which could only be resolved by analyzing the given form as part of a *conventionalized construction*. In other words, I will present an empirical argument in favor of treating particular instances of non-canonical SAY as genuinely conventionalized, as opposed to either a (grammaticalized) complementizer or a (lexical) verb, thus bringing into focus the constructionist perspective implicit in Matić & Pakendorf's (2013) notion of non-canonical SAY.

My argument will be based on the distribution of SAY-complementizers *teze* (a default complementizer used with verbs of speech and thought) and *tenine* (a specialized complementizer restricted to verbs of communicative reception, particularly 'hear') in the Poshkart dialect of Chuvash, illustrated in (2b)–(2c), which are identical to, respectively, the simultaneous converb and the past participial nominalization of the verb *te* 'say', cf. (2a).[2] Most of my argument will be dealing with pronominal deixis in clauses introduced by *teze* and *tenine*, as well as by the

---

conventionalized uses of SAY in terms of *constructionalization* (Traugott and Trousdale 2013), i.e., emergence of new constructions in the sense of conventionalized form-meaning pairings.

[2] I will refer to *teze* and *tenine* as complementizers for the sake of convenience, even though the analysis of these forms as complementizers is precisely the issue under discussion.

(contenful) verb *te* 'say', and in particular the "shifted" interpretation of the 1st person agreement on the embedded verb (where it cross-references the reported rather than the actual speaker), cf. (2b)–(2c). This pattern is reminiscent of direct speech but cannot be reduced to it since the 2nd person pronoun (*sanba*) inside the complement clause in (2a)–(2c) takes the perspective of the *actual* speaker, as in reported speech.

(2) a. (*sonʲa,*) *boris* [*san-ba* *ëɕl-e-p*] *te-rʲ-e.*
Sonya Boris 2SG-INST work-NPST-1SG say-PST-3SG
'(Sonya,) Boris₁ said that he₁ will work with you_addressee.'

b. (*sonʲa,*) *boris* *man-a* [*san-ba* *ëɕl-e-p*] *te-ze* *kala-rʲ-e.*
Sonya Boris 1SG-OBJ 2SG-INST work-NPST-1SG say-SIM.CVB say-PST-3SG
'(Sonya,) Boris₁ told me that he₁ will work with you_addressee.'

c. (*sonʲa,*) *ep* *boris-ran* [*vəl* *san-ba* *ëɕl-e-p*]
Sonya 1SG.NOM Boris-ABL 3SG.NOM 2SG-INST work-NPST-1SG
{*te-n-i-ne* / **te-ze*} *ëlt-rʲ-em.*
say-PST.PTC-POSS3-OBJ say-SIM.CVB hear-PST-1SG
'(Sonya,) I heard from Boris₁ that he₁ will work with you_addressee.'

d. (*sonʲa,*) *ep* *maṣa-ran* [*Boris* *san-ba* *ëɕl-e-t*]
Sonya 1SG.NOM Masha-ABL Boris 2SG-INST work-NPST-3SG
{*te-n-i-ne* / (?)*te-ze*} *ëlt-rʲ-em.*
say-PST.PTC-POSS3-OBJ say-SIM.CVB hear-PST-1SG
'(Sonya,) I heard from Masha that Boris will work with you_addressee.'

A similar pattern like that in (2a)–(2c), where indexicals within the same clause take different perspectives, has been reported in the (functional) typological literature, where it is described in terms of 'semi-direct', combined or bi-perspectival speech (see, e.g., Aikhenvald 2008, Güldemann 2008, Evans 2006, 2013), and sometimes characterized as a pragmatic/stylistic "anomaly" or deviation from the direct vs. indirect speech dichotomy. Instead, I will view the pattern in (2a)–(2c) as an instance of *indexical shift* (of the 1st person agreement), a phenomenon whereby indexicals take the perspective of the reported speaker in the absence of quotation (see, e.g., Shklovsky & Sudo 2014 and also Schlenker 2011, Deal 2017, Sundaresan 2018 for theoretical overviews and typological data), which potentially provides a different route to account for mixed pronominal deixis (see section 3). I will not, however, commit myself to specific theories of

indexical shift such as monster-centric theories, etc. (see Sundaresan 2018), but will use this notion primarily for descriptive purposes. I believe the advantage of this perspective is that it highlights, first, the status of mixed patterns like (2a)–(2c) as a fully conventional, default system of person alignment in a particular language (Poshkart Chuvash) and, second, its potentially neutral status from the point of view of cross-linguistic variation in the domain of reported speech (cf. Nikitina 2012:236 and also Spronck & Nikitina 2019 for a similar perspective).

I will be mostly concerned with the variability of the "controller" for the shifted 1st person, as we see in (2b)–(2c). For example, with verbs of speech (and thought) the controller is the matrix subject, as in (2b), whereas with verbs of hearing, as in (2c), the controller is the ablative source. Moreover, this difference is correlated with the choice of the complementizer: *teze* in the former case and *tenine* in the latter. The main puzzle is that this "switch" from *teze* to *tenine* under 'hear' is obligatory only in the presence of a shifted indexical, whereas it is at most a preference in non-shifted complements, cf. (2d).

I will show that although the pattern in (2), seems to argue for the decomposition of the complementizers as (syntactically) the converb and the past participial nominalization of the verb *te* 'say', there are important problems with this view.[3] To resolve this paradox, I will propose to view the patterns in (2) instead as instances of *constructions*, i.e. conventionalized form-meaning pairings which constitute speakers' knowledge of language (Goldberg 1995, 2006).

Apart from making a theoretical point, the goal of the paper is to document the (to my knowledge) previously undescribed pattern that we see in (2). While indexical shift under SAY-complementizers has been acknowledged as an important phenomenon in its own right (Messick 2017) and while the presence of more than one non-canonical SAY in a given language has also been described in typological work (see, e.g., Matić & Pakendorf 2013), there has not been a detailed language-internal comparative study of indexical shift with several SAY-complementizers in a given language. I believe that such a study, by focusing on the potentially varying degrees of verb vs. complementizer properties exhibited by different SAY-complementizers, can provide insights into the specific nature of conventionalization of non-canonical SAY. I also hope it will help to advance typological work on indexical shift, which so far has devoted little attention to the problem of the controller choice despite the generally acknowledged affinity between indexical

---

[3] Özyıldız, Major and Maier (2018) have recently proposed a decompositional analysis of the SAY-complementizer *diye* in Turkish, also based on evidence from indexical shifting (see section 4.1 for some further discussion, see also Yıldırım-Gündoğdu 2017 for a detailed study of indexical shift under *diye*).

shift and infinitival control (see, e.g., Anand & Nevins 2004 and also Landau 2015:33ff for discussion).[4]

The paper is based on the field work data collected by the author in the village of Poshkart, Chuvash Republic, Russian Federation in 2017–2019.

The paper is structured as follows. In section 2, I provide some background on SAY-complementizers in Poshkart Chuvash. Section 3 introduces the notion of indexical shift and discusses the parameters of indexical shift in Poshkart Chuvash. Section 4 presents the main puzzle and sketches an account of this puzzle in terms of syntactic decomposition of SAY-complementizers. Section 5 discusses some problems with the decompositonal account and presents a constructionist alternative. Section 6 concludes.

2. SAY-complementizers in Poshkart Chuvash

2.1. Poskart Chuvash

Poshkart Chuvash is a variety of Chuvash (< Turkic) spoken in the Poshkart village located in the Yadrin district of the Chuvash Republic (the Volga river region of Russian Federation). The Poshkart variety is identified as a distinct dialect of Chuvash and has a number of phonological and morphosyntactic features distinguishing it from the standard variety (see, e.g., Ašmarin 1898:344–392).[5] Poshkart has about 400 inhabitants, according to the 2010 census.[6] All speakers of Poshkart Chuvash also speak Russian and have at least some command of Standard Chuvash.

To facilitate the discussion of the data to be presented below, I will mention a few basic facts about Poshkart Chuvash morphosyntax.[7] Poshkart Chuvash has many morphosyntactic features commonly associated with Turkic languages, but there are also some peculiarities. It is primarily an agglutinative language with vowel harmony. It has a rich case system with differential accusative marking although there is no morphological distinction between direct and indirect objects, which are both marked with dative-accusative ("objective") case. By contrast, the system of possessive marking is very impoverished with the 1st and 2nd person markers mostly lost or

---

[4] The abovementioned study by Özyıldız et al. (2018) is a notable exception.
[5] In what follows, I will disregard any possible differences between Poshkart Chuvash and the standard variety even though this is an interesting topic in itself.
[6] The data are cited in the article on Poshkart from Chuvash Wikipedia
(https://cv.wikipedia.org/wiki/%D0%9F%D1%83%D1%88%D0%BA%C4%83%D1%80%D1%82_(%D0%95%D1%82%C4%95%D1%80%D0%BD%D0%B5_%D1%80%D0%B0%D0%B9%D0%BE%D0%BD%C4%95).
[7] The discussion is based on unpublished field work reports by participants of the linguistic expeditions to Poshkart in 2017–2019.

frozen and the 3rd person marker (which has no number distinction) emerging into the marker of definiteness.

Finite verb forms agree in person and number with their subjects. Whereas present and future tenses ("non-past") are not morphologically distinguished, there are several markers of past tense, including the simultaneous converb and the past participle used finitely. Aspectual distinctions are commonly expressed by complex predicates. Embedded clauses are typically non-finite and headed by various converbs and participles; finite clauses are embedded with the help of non-finite forms of the verb *te* 'say' (see below).

The basic word order is SOV although some material, especially subjects and complement clauses, may follow the verb and pre-verbal constituents can be more or less freely re-ordered.

Below I will give an overview of Poshkart Chuvash complementation system, focusing on the strategies involving SAY-complementizers or the contentful verb SAY. I will restrict my attention to finite declarative complements of verbs of speech and cognition.[8] But first I will briefly discuss clauses headed by participial nominalizations as they can express similar meaning and often alternate with finite complements introduced by SAY-complementizers.[9]

2.2. Participial clauses

Participial clauses, illustrated in (3a)–(3c), are arguably the most unmarked (non-controlled) complementation strategy in Poshkart Chuvash.[10] These are clauses headed by participial nominalizations, most commonly derived from the past participle (expressing either anteriority or simultaneity), but also the future and the debitative participle. The participial head is marked with an optional frozen 3rd person possessive marker (used with all person/number combinations) and a case or a postposition marker selected by the higher predicate.[11] Participial clauses have a sentence-like internal syntax. In terms of their external syntax, participial clauses usually follow

---

[8] Other non-finite complementation strategies include infinitival complements, which are used with a wide variety of typical subject and object control verbs, and complements headed by the simultaneous *-zA* converb, which are used with a handful of aspectual and modal predicates. Still more minor strategies include complements headed by the unmarked future participle, which appear with a few modal and volitional predicates, and complements headed by the *-zAn* converb, which appear with some subject experiencer and evaluative predicates. See Kožemjakina 2017 for details, and also Khanina MS for Standard Chuvash.

[9] Thus, I will disregard the uses of SAY(-complementizers) with embedded imperatives (cf. (13b) below), which appear in standard object control environments.

[10] I take the term "participial clauses" form Baker 2011, where it is applied to clauses headed by participial nominalizations in Sakha.

[11] The use of the (3rd person) possessive marker in participial nominalization is rather complex and appears to depend at least on the case marking and on the participial morphology. According to Kožemjakina 2017, it (almost) always appears with nominative- (cf. (10b)–(10c)), objective- (3a) and instrumental-marked (12a) participles but is typically absent or optional with other cases (3b) and with postpositions (11d).

the matrix arguments and immediately precede the matrix verb, as in (3a), but they can also sometimes precede the matrix arguments, as in (3b), and follow the matrix verb, cf. (3c). Distributionally, participial clauses are compatible with almost all verbs of speech and cognition.[12]

(3)   a. *ep*　　　　*poskil-den*　　[*maʂa*　*san-ran*　*enë*　{*il-n-i-ne*
　　　　1SG.NOM　neighbor-ABL　Masha　you-ABL　cow　get-PST.PTC-POSS3-OBJ

　　　　il-ess-i-ne}]　　　　　*ëlt-r-em.*
　　　　get-FUT.PTC-POSS3-OBJ　hear-PST-1SG

　　　　'I heard from [my] neighbor that Masha has bought/will buy a cow from you.'

　　b. [*peťa*　*koʒak*　*top-na-ʒən*]　　*maʂa*　*savən-a-t.*
　　　　Petya　cat　　find-PST.PTC-CSL　Masha　be.glad-NPST-3SG

　　　　'Masha is glad that Petya has found a/the cat.'

　　c. *es*　　　　(*pël-e-n*)　　　　[*man*　　*koʒak*　*tar-n-i-ne*]
　　　　2SG.NOM　know-NPST-2SG　1SG.POSS　cat　　get-PST.PTC-POSS3-OBJ

　　　　(*pël-e-n*).
　　　　know-NPST-2SG

　　　　'You know that my cat has run away.'

2.3. The contentful verb SAY

Finite complement clauses can be introduced either directly by the generic speech verb *te* 'say', which is the only verb with this property, or by complementizers, which are identical to non-finite forms of *te* 'say', especially the simultaneous converb *teze* (see section 2.4).

　　The (contentful) verb *te* 'say', illustrated in (4a)–(4e), can appear with different tense and person marking and is compatible with negation (4d) and imperative (4e), which shows that it cannot be considered morphologically defective. In addition, *te* 'say' can also take a dative (goal) argument, as in (4a), suggesting that it has argument structure, as is expected from a speech verb. Semantically, the verb *te* 'say' (in its contentful uses) also behaves like a speech verb and is typically used interchangeably with the other generic speech verb *kala* 'say' (modulo the change in the realization of the complement clause).

---

[12] One exception is the generic speech verb *te-* 'say', which requires a finite clause, cf. (5b). On special properties of *te-* 'say' in general see section 2.3. Another exception is *ʂotla* 'think', which requires *teze*-clauses, as in (8b).

(4) a. *ep*   *san-a*   [*ɨran*   *ɕomər*   *bol-a-t*]   *t-e-p.*
        1SG.NOM   2SG-OBJ   tomorrow   rain   become-NPST-3SG   say-NPST-1SG
        'I am saying to you that it will rain tomorrow.'

    b. [*san koʐak*   *sumarlan-za*   *kaj-nə*]   *t-etɕe.*[13]
        your cat   get.sick-SIM.CVB   go-PST.PTC   say-NPST.3PL
        'Your cat has got sick, they say.'

    c. *maʂa*   [*ɨran*   *xola-ja*   *kaj-a-p*]   *te-rʲ-e.*
        Masha   tomorrow   town-OBJ   go-NPST-1SG   say-PST-3SG
        'Masha said that she will go to town tomorrow.'

    d. *vasʲa*   *man-a*   [*man-ba*   *ëɕl-e-p*]   *te-m-en.*
        Vasya   1SG-OBJ   1SG-INST   work-NPST-1SG   say-NEG-RES.PTC
        'Vasya didn't tell me that he was going to work with me.'

    e. [*man-a*   *jorad-a-p*]   *te.*
        1SG-OBJ   love-NPST-1SG   say.IMP
        'Say that you love me.'

However, *te* 'say' also has some special syntactic properties, setting it aside from other complement-taking verbs such as *kala* 'say'. For example, *te* 'say' disallows nominal and participial complements, as shown in (5a)–(5b). It must also immediately follow the complement clause, disallowing extraposition of the clause to the right of *te* 'say', as shown in (6a), and the placement of the matrix subject between *te* 'say' and the clause, as shown in (7a). Note that both of these options are available for *kala* 'say' (with a *teze*-clause), cf. (6b)–(7b).

(5) a. *ep*   *on-a*   {*te-m-en   /   *kala-m-an*}.
        1SG.NOM   3SG-OBJ   say-NEG-RES.PTC   say-NEG-RES.PTC
        'I didn't say it.'

---

[13] The form *kajnə* in (4b) is used as the main predicate of the embedded clause, which is a general property of the past participle in Poshkart Chuvash, cf. (i).
(i) *san*   *koʐak*   *sumarlan-za*   *kaj-nə.*
    2SG.POSS   cat   get.sick-SIM.CVB   go-PST.PTC
    'Your cat has got sick.'

b. *maʂa* [*ïran* *xola-ja* *kaj-n-i-ne*] {**te-r�ʲ-e* / *kala-r̄ʲ-e*}.

Masha tomorrow town-OBJ go-PST.PTC-POSS3-OBJ say-PST-3SG say-PST-3SG

'Masha said that she will go to town tomorrow.'

(6) a. *boris* (*t-e-t) [*san-ba* *ëɕl-e-p*] (*t-e-t*).

Boris say-NPST-3SG 2SG-INST work-NPST-1SG say-NPST-3SG

'Boris says that he will work with you.'

b. *boris* (*kala-r̄ʲ-e*) [*san-ba* *ëɕl-e-p*] *te-ze* (*kala-r̄ʲ-e*).

Boris say-PST-3SG 2SG-INST work-NPST-1SG say-SIM.CVB say-PST-3SG

'Boris said that he will work with you.'

(7) a. (*lionila*) [*boris* *san-ba* *ëɕl-e-t*] (*??lionila*) *te-r̄ʲ-e*.

Lionila Boris 2SG-INST work-NPST-3SG Lionila say-PST-3SG

'Lionila said that Boris will work with you.'

b. (*lionila*) [*boris* *san-ba* *ëɕl-e-t*] *te-ze* (*lionila*) *kala-r̄ʲ-e*.

Lionila Boris 2SG-INST work-NPST-3SG say-SIM.CVB Lionila say-PST-3SG

'Lionila said that Boris will work with you.'

2.4. Converb of SAY

Finite complement clauses of verbs other than *te* 'say' are typically introduced by the complementizer *teze*, which is identical to the simultaneous (*-zA*) converb (cf. (26a)) of the verb *te-* 'say' (henceforth, **teze-clauses**).[14] *Teze*-clauses are common with many verbs of speech, including the generic speech verb *kala* (see (2b)), *moxtana* 'boast', as in (8a), *pëlder* 'announce, warn' (cf. (21a)) and others. *Teze*-clauses are also possible with (non-factive) verbs of propositional attitude, such as *ʂotla* 'think' and *ʂana* 'hope, believe', as in (8b). They are also possible, at least for some speakers, with the verb *ëlt* 'hear', cf. (2d). I will discuss this pattern in section 4.1.

---

[14] *Teze* can also be used to introduce purposive converbial clauses, as in (i). Such uses are not discussed in this paper.
(i) *Boris* *sonʲa-ba* *ëɕl-es* *te-ze* *kil-ze.*
  Boris Sonya-INST work-FUT.PTC say-SIM.CVB come-SIM.CVB
  'Boris went to the school to work with Sonya.'

(8) a. [*pet̸a   pilek   il-d-em*]   *te-ze*   *moxtan-a-t.*
         Petya   five   get-PST-1SG   say-SIM.CVB   boast-NPST-3SG
         'Petja boasts that he has got an A ("five").'

   b. [*ul-ë   pilek il-e-t*]   *te-ze*   *aməş*   {*şan-a-t*   /
       son-POSS3   five   get-NPST-3SG   say-SIM.CVB mother.POSS3 hope-NPST-3SG

       şotl-a-t}.
       think-NPST-3SG
       '[Petja's] mother hopes/thinks that her son will get an A ("five").'

By contrast, factive and semi-factive verbs, including verbs of perception such as *kor* 'see', as in (9a), verbs of knowledge such as *pël* 'know' (and also *anlan* 'understand' and *asta* 'remember'), as in (9b), and emotive factive verbs such as *sabən* 'glad', as in (9c), generally resist *teze*-clauses and require participial clauses.

(9) a. *es   pørt   tu-za   {*lart-sa   te-ze   /
         2SG.NOM   house   do-SIM.CVB   put-SIM.CVB   say-SIM.CVB

         lart-n-i-ne}   *pet̸a*   *kor-za.*
         put-PST.PTC-POSS3-OBJ   Petya   see-SIM.CVB
         'Peter saw that you had built a house.'

   b. *ep   eki-ren   on   koʐag-ë   {*tar-za   teze   /
       1SG.NOM   sister-ABL   3SG.POSS   cat-POSS3   run-SIM.CVB   say-SIM.CVB

       tar-n-i-ne}   *pël-d-em.*
       run-PST.PTC-POSS3-OBJ   know-PST-1sg
       'I learned from [my] sister that her cat has ran away.'

   c. *pet̸a*   *pilek*   {??*il-d-em*   *te-ze*   /   *il-ne-ʐen*}
       Petya   five   get-PST-1SG   say-SIM.CVB   get-PST.PTC-CSL

       sabən-a-t.
       be.glad-NPST-3SG
       'Petja is glad that he has got an A ("five").'

The compatibility between *teze*-clauses and verbs of propositional attitude suggests that *teze* has become at least partly grammaticalized given that semantic bleaching and distributional extension (beyond verbs of speech) are usually viewed as signs of (early stages of)

grammaticalization (see, e.g., Heine & Kuteva 2005). Moreover, informants never refer to the verb 'say' in their translations of examples with *teze*-clauses, which is sometimes used as a criterion for distinguishing complementizers from contentful 'say' (see Matić & Pakendorf 2013:373). At the same time, restriction of *teze*-clauses to non-factive verbs of speech and propositional attitude suggests that speakers might still analyze *teze* as related to the verb *te* 'say' and thus having some lexical content, perhaps expressing what Güldemann (2008:7) calls "internal speech". On the whole, the distribution of *teze*-clauses seems to be consistent with various analytical options, including viewing them as conventionalized without being fully grammaticalized. In section 4, I will discuss and assess more specific evidence from indexical shifting for treating *teze* in *teze*-clauses as syntactically converbs.

2.5. Past participial forms of SAY

Apart from the converbial form *teze*, complement clauses can also be introduced by the past participial forms of *te* 'say', both in their participial function (as heads of clausal adnominal modifiers) and in their nominalization function (as heads of clausal arguments).[15] The former function is illustrated in (10a), where the clause is semantically the argument of the noun *xɨbar* 'news'. In such cases the verb *te* 'say' appears in the non-inflected participial form *tene*. Note that in such cases the form *teze* is disallowed (I return to this fact in section 4.2). The latter function includes both clausal subjects and clausal complements. Clausal subjects are illustrated in (10b). In such cases the clause is introduced by the form *teni*, which is the past participial nominalization of *te* 'say', marked with a frozen 3rd person and a nominative case (see footnote 11). Note that the construction with *teni* is optional as the same meaning can also be realized with a participial clause with the respective morphological marking. Notably, the form *teze* is disallowed here. Note also that clausal subjects introduced by *teni* are disallowed with emotive factives, as shown in (10c), suggesting that such forms are also not fully grammaticalized (see the previous section).

(10) a. *ep*   [*san*   *taga*   *man*   *xəjar-a*   *ɕi-z=er-ze*]

       1SG.NOM   2SG.POSS   sheep   1SG.POSS   cucumber-OBJ   eat-SIM.CVB=let-SIM.CVB

       {*te-ne*   /   **te-ze*}   *xɨbar-a*   *ëlt-r-em*.

       say-PST.PTC   say-SIM.CVB   news-OBJ   hear-PST-1SG

       'I heard the news that your sheep has eaten my cucumbers.'

---

b. *man   eki   katɕɕa {tox-n-i   /   tox-sa*

1SG.POSS sister   groom   go.out-PST.PTC-POSS3.NOM go.out-SIM.CVB

*te-n-i   /   *tox-sa   te-ze}   teres mar.*

say-PST.PTC-POSS3.NOM   go.out-SIM.CVB say-SIM.CVB true NEG.COP

'That my sister has got married (as they say) is not true.'

c. *san   koʐak   {tar-n-i   /   ??tar-za*

2SG.POSS cat   run.away-PST.PTC-POSS3.NOM   run.away-SIM.CVB

te-n-i}   pët   ozal.

say-PST.PTC-POSS3.NOM   very bad

'That your cat has run away is very bad.'

Participial nominalizations of the verb *te* 'say' can also introduce clausal complements. The most common form used in this function is *tenine*, which is marked with the 3rd person possessive marker and objective case (see section 2.2 above).[16] Clauses introduced by *tenine* (henceforth, **tenine-clauses)** typically occur with the verb *ëlt* 'hear, as in (11a), cf. (3a), with the alternating participial clause (see also (2c) above). But they are also possible with other transitive verbs of acquisition of knowledge/communicative reception such as *pël* 'know/learn', as in (11b), *anlan* 'understand', as in (11c), and *vula* 'read', as in (11d), even though with these verbs, speakers will normally use a participial clause as their preferred option. In general, according to my informants, the use of *tenine*-clauses in examples like (11) appears to be optional and driven by pragmatic considerations, namely, to emphasize that the information was received via hearsay, which probably reflects the origin of *tenine* as a form of the speech verb *te* 'say'. Still, speakers never translate these examples with the actual verb 'say', suggesting that *tenine* here is at least partly conventionalized. However, its status as (syntactically) a participle still remains uncertain (see section 4).

(11) a. *ep   poskil-den   [maʂa enë san-ran   il-e-t]*

1SG.NOM   neighbor-ABL Masha cow you-ABL   get-NPST-3SG

te-n-i-ne   ëlt-r-em.

say-PST.PTC-POSS3-OBJ   hear-PST-1SG

'I heard from [my] neighbor that Masha will buy a cow from you.'

---

b.^(?)*ep*   *eki-ren*   [*on*   *koʐag-ë*   *tar-za*]
1SG.NOM   sister-ABL   3SG.POSS   cat-POSS3   run-SIM.CVB
te-n-i-ne   *pël-d-em*.
say-PST.PTC-POSS3-OBJ   know-PST-1sg
'I learned from [my] sister that her cat has run away.'

c. *ep*   *velosiped*   {*ɕemerl-ze*   *te-n-i-ne*   /
1SG.NOM   bike   break-SIM.CVB   say-PST.PTC-POSS3-OBJ
ɕemerl-n-i-ne}   *anlan-d-əm*.
break-PST.PTC-POSS3-OBJ understand-PST-1SG
'I realized (from someone's words) that [Masha's] bike had broken down.'

d. *director*   *pern*   *şkol*   *perremeş*   *virən*   {*jɨşən-za*
headmaster   our   school   first   place   get-SIM.CVB
te-n-i-ne   /   *jɨşən-nə*   *ɕinʥe*   *xaʐat-re*   *vula-na*.
say-PST.PTC-POSS3-OBJ   get-PST.PTC   about   newspaper-LOC   read-PST.PTC
'The headmaster [of our school] read in a newspaper that our school has got the first place.'

Tenine-clauses can also be used with verbs of speech, in which case they typically occur with overt subjects (see, e.g., (34b) below). I discuss these uses in section 5.1. Such uses are more marked and typically imply a separate speech event (other than the one expressed by the matrix verb). Also, speakers often translate them with the extra verb 'say'. This suggests that such uses might be compositional and involve the participle of the contentful verb *te* 'say' rather than a complementizer, rendering them less (if at all) grammaticalized. Interestingly, however, as I will argue in section 5, such uses still have some degree of conventionality and should be treated as constructions, on a par with the uses in (11).

As for the past participial nominalizations marked with cases other than objective, as illustrated in (12a)–(12b), their status as conventionalized/grammaticalized is very unclear as they were rarely (if ever) elicited in field work as responses to stimuli involving the relevant verbs. In addition, they also typically allow a compositional interpretation. Although these forms are certainly interesting to explore in some detail, I ignore them in the present paper and focus exclusively on *tenine*.

(12) a. *pet^ja*   *pilek*   {*il-d-em*   *te-n-i-be* /   *il-n-i-be*}

       Petya   five   get-PST-1SG   say-PST.PTC-POSS3-INST   get-PST.PTC-POSS3-INST

       moxtan-a-t.

       boast-NPST-3SG

       'Petja₁ boasts (by saying) that he₁ has got an A ("five").'

    b. *es*   *man-a*   *poleẕ-a-p*   *te-n-i-ẕen*   *ep*

       2SG.NOM   1SG-OBJ   help-NPST-1SG   say-PST.PTC-POSS3-INST   1SG.NOM

       sabən-d-əm.

       be.glad-PST-1SG

       'I was glad that you said that you would help me.'

To summarize, I have shown that *teze* and *tenine* can be used in a complementizer-like function to introduce complement clauses. Such uses are not fully compositional, but they are not fully grammaticalized either as they seem to retain at least some lexical content of the original verb 'say' (accounting for their distributional restrictions). The overall pattern seems to be compatible with viewing *teze* and *tenine* as conventionalized SAY, in accordance with Matić & Pakendorf 2013. The question that remains, however, is whether there is specific evidence for treating these forms as *syntactically* the converb and the past participle (what I will refer to as a "decompositional analysis"). To address this question, I will examine indexical shifting properties of *teze-* and *tenine-* clauses. But first, I will briefly discuss indexical shift in general and give an overview of its properties in Poshkart Chuvash.

3. Indexical shift in Poshkart Chuvash

3.1. Diagnosing indexical shift

In the literature, *indexical shift* is usually understood as a situation where indexical elements (i.e., 1st/2nd person pronouns and agreement markers, temporal and locative adverbs like *here*, *now*, etc.) pick up their reference from the reported context (introduced by a speech or thought predicate), crucially, in the absence of quotation (see Deal 2017, Sundaresan 2018 for an overview of the typology and the existing theoretical approaches to indexical shift).

Since it is important to distinguish indexical shift from quotation ("direct speech"), I will briefly illustrate the latter in Poshkart Chuvash. An example of a quotative use is given in (13a), where the quotation is marked in bold. The fact the string in bold is a quotation is shown by the

use of a vocative (*Vas^ja*), which is normally impossible in true embedding. Another feature concerns the interpretation of the 1st person possessive pronoun *man*, which takes the perspective of the reported speaker. While cross-linguistically this does not distinguish quotation from indexical shift, it can serve as a (negative) diagnostic of indexical shift in Poshkart Chuvash, where overt pronouns never shift, as I will show below (see section 3.2.1). Also observe that the matrix subject can optionally follow the main verb in (13a), which might be viewed as an instance of quotative inversion, cf. the English translation. Note that the use of the imperative in (13a) does not reliably indicate quotation, as embedded imperatives in Poshkart Chuvash are possible even in ordinary embedding, as in (13b), cf. the use of the 3rd person pronoun to refer to the reported speaker, which would be incompatible with a quotation construal.[17] As we see in (13a), both *teze*-clauses and the (contentful) verb *te* 'say' can introduce quotations.

(13) a. (*amƽş*) **vas^ja, man pat-n^ja kil** {*te-r^j-e /*
mother.POSS3 Vasya, 1SG.POSS side-POSS3.OBJ come.IMP say-PST-3SG

te-ze kala-r^j-e} (*amƽş*).
say-SIM.CVB say-PST-3SG mother.POSS3

'"Vasya, come to me$_1$," said [Vasya's] mother$_1$.'

b. *amƽş vaɕa [on pat-n^ja kil] {te-r^j-e /*
mother.POSS3 Vasya.OBJ 3SG.POSS side-POSS3.OBJ come.IMP say-PST-3SG

te-ze kala-r^j-e}.
say-SIM.CVB say-PST-3SG

'[Vasya's] mother$_1$ told Vasya to come to her$_1$.'

I will now illustrate indexical shift in Poshkart Chuvash. Consider (14b), where the 1st person agreement on the embedded refers to the reported speaker (*Boris*), rather than to the actual speaker. While (14b) can be understood as containing a quotation directly reporting Boris's original utterance in (14a), it can be shown that (14b) must also have an alternative (i.e., indexical shift) analysis where the complement clause (bracketed) is not quoted but embedded in an ordinary fashion.

---

[17] On embedded imperatives see, e.g., Kaufmann & Poschmann 2013 and references therein.

(14) a. *son^ja-ba   ëɕl-e-p.*

      Sonya-INST   work-NPST-1SG

      'I (will) work with Sonya.'

  b. *boris*   [*son^ja-ba*   *ëɕl-e-p*]   *te-ze*   *kala-r^j-e.*

      Boris   Sonya-INST   work-NPST-1SG say-SIM.CVB say-PST-3SG

      'Boris₁ said that he₁ will work with Sonya.' (lit.: 'Boris said I will work with Sonya.')

  c. *boris*   [*kam-ba*   *ëɕl-e-p*]   *te-ze*   *kala-r^j-e?*

      Boris   who-INST work-NPST-1SG say-SIM.CVB say-PST-3SG

      'Who did Boris₁ say he₁ will work with?' (lit.: 'Who did Boris say I will work with?')

One common way this is shown in the indexical shifting literature (cf. Shklovsky & Sudo 2014, a.o.) is by placing inside the complement clause a wh-phrase (*kamba*) taking scope over the whole sentence, as in (14c).[18] Since the wh-phrase is not part of the original utterance (cf. (14a)), example (14c) does not involve the faithful (verbatim) representation of this utterance and hence cannot be analyzed as a direct quotation.[19] Thus, by the informal definition given above, the reported perspective of the 1st person agreement in (14c) must be a shifted indexical. As a result, an analysis in terms of indexical shift should also be available for (14b), although, as such, this example is potentially ambiguous between a quotation and an indexical shift analysis.

While the wh-question diagnostic in (14c) is a useful one, it may render the examples to be judged by informants overly complex. Thus, instead of this diagnostic I will use a different one, namely, the presence of overt pronouns referring to the actual speaker/hearer inside the complement clause.[20] To use this diagnostic, one has to set up a context where one of the participants in the complement clause coincides with the actual speaker or addressee, as in (15a)–(15b). Since in Poshkart Chuvash, the speaker and the addressee of the actual speech act are normally expressed by 1st and 2nd person pronouns, as in ordinary "reported discourse", the direct

---

[18] This diagnostic is usually referred to as "long-distance wh-movement", incorporating the common analysis of such wh-questions in generative grammar. The two other diagnostics Shklovsky & Sudo (2014) list are (i) long-distance licensing of negative polarity items; and (ii) non-verbatim reports (see, e.g., Anand 2006). The diagnostic that is primarily used in this paper (see below) can be viewed as an instance of the latter.

[19] Maier (2015) discusses some potential problems with this diagnostic, arguing that while long-distance wh-movement in examples like (14c) is incompatible with "full quotation", it is not incompatible with the partial/mixed quotation, as shown by the attested examples, as in (i). This suggests that examples like (14c) can be viewed as instances of partial/mixed quotation (targeting only the embedded verb). However, the across-the-board partial/mixed analysis of indexical shift has been extensively criticized in the literature (see, e.g., Anand 2006, Schlenker 2011, Maier 2016) based on the fact that it is hard to reconcile with the observed morphosyntactic constraints on indexical shift (see Deal 2017 and section 3.2 below).

(i) Who did Mary say that she would "never misunderestimate ever again"? (Maier 2015:713)

[20] This diagnostic, e.g., is used by Anand and Nevins (2004) for Slave and discussed in Maier 2016:10.

quotation is excluded in such examples. Yet, the 1st person agreement may still retain the perspective of the reported speaker, which suggests that we are dealing with indexical shift in (15a)–(15b).

(15) a. Context: Boris said to Sonya, "I will work with you". Sonya reports this to someone else.

    boris   [*man-ba*   *ĕɕl-e-p*]   *te-ze*   *kala-r^j-e.*

    Boris   1SG-INST   work-NPST-1SG   say-SIM.CVB   say-PST-3SG

    'Boris$_1$ said that he$_1$ will work with me$_{speaker}$.' (lit.: 'Boris said I will work with me$_{speaker}$.')

  b. Context: Boris said to me, "I will work with Sonya". I report this to Sonya.

    (*son^{j}a,*) *boris*   *man-a*   [*san-ba*   *ĕɕl-e-p*]   *te-ze*   *kala-r^j-e.*

    Sonya   Boris   1SG-OBJ   2SG-INST   work-NPST-1SG   say-SIM.CVB   say-PST-3SG

    '(Sonya,) Boris$_1$ told me that he$_1$ will work with you$_{addressee}$.' (lit.: 'Boris told me I will work with you$_{addressee}$.')

One potential limitation of this diagnostic is that even in true quotations the verbatim requirement can sometimes be violated by way of adjustments (or "unquotation"), as in examples like *Obama said that the news "took [him] by surprise"* (see Maier 2015, 2016). Thus, one can potentially analyze the overt 1st/2nd person pronouns in (15a)–(15b) as instances of unquotation. Besides some general problems with this analysis (see footnote 19), there is a language-internal evidence against it. In Poshkart Chuvash, the presence of overt 1st person subject (*ep*), is incompatible with indexical shift, as shown by the wh-question diagnostic in (16a). This implies that the reported perspective of *ep* in (16b) must be an instance of direct quotation. The crucial observation is that if unquotation for the overt 1st person pronoun (*manba*) in the complement were generally available, we would expect it to be able to refer to the actual speaker. This, however, is not possible, as one must use a 2nd person pronoun (*sanba*), with the reported perspective. Given this reasoning, I conclude that an analysis of (15a)–(15b) and similar examples in terms of "unquotation" (mixed/partial quotation) is untenable in Poshkart Chuvash.

(16) a. *boris*   [(**ep*)   *kam-ba*   *ĕɕl-e-p*]   *te-ze*   *kala-r^j-e?*

    Boris   1SG.NOM   who-INST   work-NPST-1SG   say-SIM.CVB   say-PST-3SG

    'Who did Boris$_1$ say he$_1$ will to work with?'

b. Context: Boris told me, "I will work with you." I am reporting this to Roza.

boris	*man-a*	[*ep*	{*man-ba / *san-ba*}	*ëɕl-e-p*]
Boris	1SG-OBJ	1SG.NOM	1SG-INST 2SG-INST	work-NPST-1SG

te-ze	*kala-rʲ-e.*
say-SIM.CVB	say-PST-3SG

'Boris₁ told me that he₁ will work with me_speaker.'

The discussion of (15a)–(15b) has presupposed that indexical shift can be partial, i.e., only the 1st person agreement in the clause may shift while other indexicals remain unshifted. This has been a problematic assumption for some approaches to indexical shift, in particular those based on "monsters" (context-shifting operators), which usually adopt the so-called *Shift Together* constraint, requiring *all* indexicals in the clause to shift (see, e.g., Anand & Nevins 2004, Shklovsky & Sudo 2014, Deal 2017, 2018). However, Sundaresan (2018) argues that if particular classes of indexicals in a given language can be lexically specified as shiftable of unshiftable, then *Shift Together* will only be predicted to hold for shiftable indexicals and thus apparent "*Shift Together* violation" such as (15a)–(15b) will no longer pose a problem (assuming the overt indexical in the clause is unshiftable). Indeed, this is how Sundaresan (2018) analyzes similar examples in Tamil (cf. her (88)): she argues that only null 1st person pronoun/1st person agreement in Tamil is shiftable, whereas overt pronouns are unshiftable. The same analysis can be extended to Mishar Tatar, another language with partial indexical shift (Podobryaev 2014). Although I cannot do justice to the theoretical nuances of these arguments here, the purpose of this discussion was to show that partial shift need not be a problem for the indexical shift analysis of (15a)–(15b).[21]

Before concluding this section, I wish to comment on the treatment of mixed pronominal deixis in examples like (15a)–(15b) in the (functional) typological literature. As I mentioned in the Introduction, similar examples have been discussed in the typological literature where they are viewed as a special variety of reported discourse (alongside direct and indirect one), variously referred to as "semi-direct", "combined", or "bi-perspectival" speech (see, e.g., Evans 2006, Aikhenvald 2008, Güldemann 2008). Although this literature does not make explicit connection with the pragmatic ("mixed/partial quotation") analysis of indexical shift (see, e.g., Maier 2015), there seems to be an affinity between the two as both tend to view mixed patterns as a matter of stylistic choices. This tendency is criticized by Nikitina (2012:236), who argues that it may neglect

---

[21] Deal (2018) proposes an alternative analysis of partial "indexical shift" in Mishar Tatar in terms of *indexiphors*, which are roughly 1st person logophors (see also some discussion in section 5.2). For arguments against this view see Sundaresan 2018.

obligatory and hence grammatical nature of mixed patterns at least for some languages. Whether this criticism can be met (perhaps by allowing stylistic/pragmatic preferences to become conventionalized), there is a further problem for this view. The problem is that by singling out mixed patterns as a separate category on the direct-indirect continuum, this approach fails to capture the fact that mixed patterns in examples like (15a)–(15b) and "non-mixed" patterns in examples like (14b)–(14c) intuitively form a natural class as both are the unmarked/conventionalized expressions of the identity between the reported speaker and the subject of the embedded clause. This similarity, however, is easily captured in an indexical shift approach, which allows only for *some* indexicals in the clause to shift.

To conclude, I hope to have shown that an analysis in terms of indexical shift can be a fruitful way of characterizing the reported perspective of the 1st person in examples like (14b)–(14c) and (15a)–(15b). Although I deliberately ignored important theoretical problems underlying the notion of indexical shift, I believe that it can still be a useful analytical tool for typological description (independently of one's view of the specific theory of indexical shift).

In the following section I will discuss some parameters of indexical shift in Poshkart Chuvash, based on the existing typological/theoretical overviews in Deal 2017 and Sundaresan 2018.

3.2. Parameters of indexical shift in Poshkart Chuvash

3.2.1. Which indexicals can shift

As we saw in the previous section, indexical shift is possible for 1st person agreement but only in the absence of an overt subject (*ep*), cf. (16a). This suggests that overt subject pronouns do not shift and must take the perspective of the actual speaker (outside quotation). This is further illustrated in (17a), where *ep* can only refer to the actual speaker, cf. (17b) with a non-overt subject, where indexical shift *is* possible.

(17)   Context: Masha said to me, "I will get an A for Luiza's class." I am reporting this to Luiza.

    a. **maşa   man-a   [ep   sern-be   pilek il-e-p]   te-ze*

       Masha   1SG-OBJ 1SG.NOM   you.PL-INST   five   get-NPST-1SG   say-SIM.CVB

       kala-rʲ-e.

       say-PST-3SG

       Intended: 'Masha[1] told me that she[1] will get an A ("five") for your[addressee] [class].'

b. *maşa   man-a   [sern-be   pilek il-e-p]   te-ze   kala-rʲ-e.*
   Masha   1SG-OBJ   you.PL-INST   five   get-NPST-1SG   say-SIM.CVB   say-PST-3SG
   'Masha₁ told me that she₁ will get an A ("five") for your_addressee [class].'

Non-shiftability of overt indexicals seems to be a general property of Poshkart Chuvash and extends to non-subject pronouns and the locative adverbial 'here', as shown in (18)–(19). In (18), the participant coindexed with the reported speaker must be expressed with the 3rd person pronoun (*onba*) and cannot be expressed by the 1st person pronoun (*manba*), which refers to the actual speaker. Similarly, in (19), the adverbial *konda* 'here' cannot take the reported perspective (and refer to Moscow, where the original speech act took place) but only to the perspective of the actual speaker (i.e., to Poshkart, where the actual speech act takes place). In order to refer to 'here' of the reported speech act, one must use *onda* 'there'.

(18) *boris   [kam   {*man-ba / on-ba}   ёɕl-e-t]   te-ze*
      Boris   who.NOM   1SG-INST   3SG-INST   work-NPST-3SG   say-SIM.CVB
      kala-rʲ-e?
      say-PST-3SG
      'Who did Boris₁ say will work with him₁?'

(19) Context: One month ago, I met Boris in Moscow. He said to me, "I am living here, in Moscow". Now I am in Poshkart talking to Roza.
      *boris   [{*konda / onda}   porn-a-p]   te-ze   kala-rʲ-e?*
      Boris   here   there   live-NPST-1SG   say-SIM.CVB   say-PST-3SG
      'Boris₁ told me that he₁ is living there (= in Moscow).'

A similar correlation between overtness and non-shiftability of indexicals has been recently described for Mishar Tatar, another Turkic language, by Podobryaev (2014), who views (un)shiftability as a lexical property of individual pronouns.[22] In contrast to Mishar Tatar, where both 1st and 2nd null pronouns (*pro*) shift, however, Poshkart Chuvash disallows shift of the 2nd person *pro*, which must always refer to the actual addressee (outside quotation), as shown in (20). The other difference between Poshkart Chuvash and Mishar Tatar is that the latter allows shift of

---

[22] See Messick 2017: Ch. 3 for some discussion of the possible sources of the correlation between shiftability and overtness.

1st/2nd person possessive agreement (null possessive pronouns), whereas in Poshkart Chuvash 1st/2nd person possessive agreement is mostly unproductive.[23]

(20) Context: I am going to play chess with Boris. Masha thinks that Boris is superior than me and says to Boris, "You are going to beat him." I report this to Katya.

maʂa   boris-a   [man-a {*ɕinder-e-n   /   ɕinder-e-t}]   te-ze
Masha   Boris-OBJ   1SG-OBJ   win-NPST-2SG   win-NPST-3SG   say-SIM.CVB
kala-rʲ-e.
say-PST-3SG
'Masha told Boris₁ that he₁ that he will beat me_speaker [in chess].'

As a result, in Poshkart Chuvash the only shiftable element is the 1st person agreement/1st person null subject (*pro*). This pattern is not typologically unusual. Thus, in Sundaresan's (2018) mini-typological database of 26 languages, there are at least four languages, unrelated to Poshkart Chuvash, that show this pattern, namely Telugu, Tamil (both Dravidian), Nuer (Nilotic) and Dargwa (Norheastern Caucasian). It is also in line with the implicational hierarchy of "shifty" indexicals proposed by Deal (2017), according to which 1st person is most likely to shift, followed by 2nd person followed by HERE.

3.2.2. Which verbs and complement types allow shifting

In the previous section, we saw that the only shiftable indexical element in Poshkart Chuvash is the 1st person subject agreement / null subject pronoun (*pro*). Since 1st person agreement in Poshkart Chuvash is restricted to finite verbs, we only find shift in finite complements, which is again in line with typological generalizations, according to which indexical shift is generally restricted to finite clauses (Deal 2017, see also Messick 2017: Ch. 4 for some suggestions for why it might be the case).

Further, since finite complements are introduced either by the verb *te* 'say' or by the forms of *te* 'say' functioning as complementizers, we expect the distribution of indexical shift to follow the distribution of the complementizers (*teze*, *tenine*, etc.). This is indeed what we find. We have already seen indexical shift with *teze* under the speech verb *kala* 'say' (see, e.g., (15a)–(15b)). Similar examples with another speech verb *pëlder* 'announce' and a cognitive verb *ʂotla* 'think' are given in (21a)–(21b). Example (21c) illustrates indexical shift with the complementizer *tenine* under the verb *ëlt* 'hear'. Note that in (21c) the 1st person agreement on the embedded verb is

---

[23] For example, according to Remizova's (MS) unpublished field report on Poshkart Chuvash, 1st person (singular) possessive is only used with kinship terms, primarily in vocatives.

"controlled" by the ablative source (*Boris*) rather than by the matrix subject; also note that the embedded subject is realized as the 3rd person pronoun *vəl*. I will discuss these peculiarities in some detail in section 4. Finally, (21d) shows indexical shift under the (contentful) verb *te* 'say' itself.

(21)   a. Context: Vasya said to me, "I am going to do military service with you." I am reporting this to Katya.

vas^ja	*man-a*	[*man-ba*	*perle*	*sluʐit^j*	*t-a-p*]	*te-ze*
Vasya	1SG-OBJ	1SG-INST	together	service	win-NPST-1SG	say-SIM.CVB

pëlder-tɕ-e.
announce-PST-3SG

'Vasya$_1$ announced that he$_1$ is going to do military service with me$_{speaker}$.'

b. Context: I am playing chess with Vasya. Vasya says to me, "I will beat you in chess." I am reporting this to Katya.

boris	[*man-a*	*ɕinder-e-p*]	*te-ze*	*ʂotla-t.*
Boris	1SG-OBJ	win-NPST-1SG	say-SIM.CVB	think-NPST.3SG

'Boris$_1$ thinks he$_1$ that he will beat me$_{speaker}$ [in chess].'

c. Context: Boris said to me, "I will work with Sonya." I am reporting this to Sonya.

(*son^ja,*)	*ep*	*boris-ran*	[*vəl*	*san-ba*	*ëɕl-e-p*]
Sonya	1SG.NOM	Boris-ABL	3SG.NOM	2SG-INST	work-NPST-1SG

te-n-i-ne	*ëlt-r^j-em.*
say-PST.PTC-POSS3-OBJ	hear-PST-1SG

'(Sonya,) I heard from Boris$_1$ that he$_1$ will work with you$_{addressee}$.'

d. Context: = (21c)

(*son^ja,*)	*boris*	[*san-ba*	*ëɕl-e-p*]	{*t-e-t*	/	*te-r^j-e*}.
Sonya	Boris	2SG-INST	work-NPST-1SG	say-NPST-3SG		say-PST-3SG

'(Sonya,) Boris$_1$ said/says that he$_1$ will work with you$_{addressee}$.'

3.2.2. Optionality of shift

In the two previous sections we identified which elements in Poshkart Chuvash *can* shift and when. Another parameter concerns whether they *must* shift if they can. Indexicals in Poshkart Chuvash are generally not required to shift, a situation parallel to Mishar Tatar (Podobryaev 2014). For example, 1st person agreement with a non-overt subject is systematically ambiguous between a

shifted interpretation, where it refers to the matrix subject (cf. (15b) above), and an unshifted interpretation, where it refers to the actual speaker, as shown in (22a). It should be noted, however, that speakers tend to realize the embedded subject overtly (as the 1st person pronoun *ep*), in order to refer to the actual speaker (presumably for reasons of disambiguation, see also the discussion surrounding (16a) above).

(22) a. (*son^{j}a,*) *boris*   *man-a*   [*san-ba*   *ëɕl-e-p*]   *te-ze*
           Sonya   Boris   1SG-OBJ   2SG-INST   work-NPST-1SG   say-SIM.CVB

           kala-r^j-e.
           say-PST-3SG

           '(Sonya,) Boris$_1$ told me that {I$_{speaker}$ / he$_1$} will work with you$_{addressee}$.'

      b. (*son^{j}a,*) *boris*   *man-a* [*ep*   *san-ba*   *ëɕl-e-p*]   *te-ze*
           Sonya   Boris   1SG-OBJ 1SG.NOM   2SG-INST   work-NPST-1SG   say-SIM.CVB

           kala-r^j-e.
           say-PST-3SG

           '(Sonya,) Boris$_1$ told me that {I$_{speaker}$ / *he$_1$} will work with you$_{addressee}$.'

This completes the overview of the basic parameters of indexical shift in Poshkart Chuvash.[24] To summarize, indexical shift in Poshkart Chuvash (i) targets 1st person subject agreement/null subject pronoun (*pro*); (b) is restricted to finite clauses introduced by *te* 'say' and complementizers derived from *te* 'say' (*teze*, *tenine*, etc.); (c) is optional. Now we can turn to the main puzzle related to the choice of the controller for the shifted 1st person with *teze* and *tenine*.

4. The puzzle: shift of first person under 'hear'

4.1. Introducing the puzzle

As we saw in (21c), indexical shift can occur under the verb *ëlt* 'hear', in which case the complementizer *tenine* is used and the shifted 1st person is controlled by the ablative source. Interestingly, *teze* is disallowed in the presence of indexical shift, as shown in (23a), as opposed to *tenine*, as we saw in (21c).[25] Interestingly, in the absence of indexical shift, *teze* is in principle

---

[24] Apart from the three parameters discussed in section 3.2, Deal (2017) also identifies another parameter of indexical shift, namely, whether a given indexical has a *de se* interpretation. This parameter was not tested in the present work.
[25] The unacceptability of (23a) does not depend on the presence of the embedded subject *vəl*, a point I return to in section 5.2

compatible with *ëlt* 'hear' (along with *tenine*, which is the unmarked option), as shown in (23b). It must be acknowledged that some speakers disprefer *teze* in favor of *tenine* in examples like (23b). However, there is still a clear contrast in acceptability between *TEZE* + HEAR with indexical shift, as in (23a), which is uniformly rejected by all speakers, and *TEZE* + HEAR without indexical shift, as in (23b), which is accepted by most speakers.

(23)   a. Context: Boris said to me, "I will work with Sonya." I am reporting this to Sonya.

*(*son^j a,*)   ep   boris-ran   [(*vəl*)   san-ba   ëɕl-e-p]

Sonya   1SG.NOM   Boris-ABL   3SG.NOM   2SG-INST   work-NPST-1SG

te-ze   *ëlt-r^j-em.*

say-SIM.CVB   hear-PST-1SG

Intended: '(Sonya,) I heard from Boris$_1$ that he$_1$ will work with you$_{addressee}$.'

   b. Context: Masha said to me, "Boris is going to work with Sonya." I am reporting this to Sonya.

(*son^j a,*)   ep   maʂa-ran   [*Boris*   san-ba   ëɕl-e-t]

Sonya   1SG.NOM   Masha-ABL   Boris   2SG-INST   work-NPST-3SG

{*te-n-i-ne*   /   ^{(?)}*te-ze*}   *ëlt-r^j-em.*

say-PST.PTC-POSS3-OBJ   say-SIM.CVB   hear-PST-1SG

'(Sonya,) I heard from Masha that Boris will work with you$_{addressee}$.'

To complete the paradigm, we also need to consider whether shifted 1st person under 'hear' can be controlled by the matrix subject, i.e., the hearer/recipient of the report. It turns out that this option is not possible neither with *tenine*, as shown in (24a), nor with *teze*, as shown in (24b).[26] To express this meaning, one has to use the 3rd person on the embedded verb, as in ordinary embedding without shift.

---

[26] The diacritic in (24b) reflects the fact that *teze* is slightly marked under *ëlt* 'hear' in general (see (23b)).

(24) Context: The teacher said to Vasya, "Vasya, you sing better than everyone else." I am reporting this to Roza.

a. *vasʲa   utɕitelʲ-dan   [vəl   tɕi   ləjk   {*jorl-a-p   /   jorl-a-t}]*
   Vasya   teacher-ABL   3SG.NOM   most   good   sing-NPST-1SG   sing-NPST-3SG
   te-n-i-ne   ёlt-se.
   say-PST.PTC-POSS3-OBJ   hear-SIM.CVB
   'Vasya$_1$ heard from the teacher that he$_1$ sings best [of all].'

b. $^{(?)}$*vasʲa   utɕitelʲ-dan   [tɕi   ləjk   {*jorl-a-p   /   jorl-a-t}]   te-ze*
   Vasya   teacher-ABL   most   good   sing-NPST-1SG   sing-NPST-3SG   say-SIM.CVB
   ёlt-se.
   hear-SIM.CVB
   'Vasya$_1$ heard from the teacher that he$_1$ sings best [of all].'

The whole pattern is summarized in Table 1. As we can see, shifted 1st person under *ёlt* 'hear' must be controlled by the ablative source and requires the complementizer *tenine* (row E); control by the matrix subject (hearer) is generally prohibited (row D). In the absence of indexical shift, both complementizers are in principle possible (row C). Verbs of speech and thought take the complementizer *teze* and the shifted 1st person under is controlled by the matrix subject (rows A–B).[27]

Table 1. Acceptability of the two SAY-complementizers depending on the verb class and the choice of the controller

	Verb class & controller for shifted 1st person	Acceptability of complementizer	
		teze	*tenine*
A	SAY/THINK + no shift	OK	*
B	SAY/ THINK + shift to the matrix subject	OK	*
C	HEAR + no shift	$^{(?)}$OK	OK
D	HEAR + shift to the matrix subject	*	*
E	**HEAR + shift to the ablative source**	*	**OK**

---

[27] The unacceptability of *tenine* with indexical shift under 'say' is illustrated in (i).
(i) *(*sonʲa,*) *boris   man-a   [(vəl)   san-ba   ёɕl-e-p   te-n-i-ne]*
   Sonya   Boris   1SG-OBJ   3SG.NOM   you-INS   work-NPST-1SG   say-PST.PTC-POSS3SG-OBJ
   kala-rʲ-e.
   say-PST-3SG
   Intended: '(Sonya,) Boris$_1$ told me that he$_1$ will work with you.'

It is difficult to assess how unusual the pattern in Table 1 is from a cross-linguistic perspective as there is little data on indexical shifting under verbs of hearing, most research being focused on verbs of speech and thought (see Özyıldız et al. 2018). Moreover, alternation of complementizers of the kind we see in Poshkart Chuvash has not been discussed in connection with indexical shift. Nevertheless, it is noteworthy that the pattern in Poshkart Chuvash that we see in Table 1 is at least different from the patterns reported for Turkish (Özyıldız et al. 2018) and Uyghur (Sudo 2010).[28] In Uyghur, 1st person under 'hear' must shift to the matrix subject (the hearer), whereas shift to the (oblique) source is disallowed, which is the opposite of the pattern we see in Poshkart Chuvash. Turkish has yet a different pattern, illustrated in (25), where 1st person under 'hear' can shift not only to the ablative source, as in Poshkart Chuvash, but also to the matrix subject (the hearer), unlike in Poshkart Chuvash (Özyıldız et al. 2018). I will come back to the Turkish pattern in the next section.

(25) *Ayşe Mercan'-dan* [*kahraman-ım   diye*] *duy-du*.
      Ayşe mercan-ABL   hero-COP.1SG   DIYE hear-PST.3SG
      'Ayşe1 heard from Mercan2 that she1/2 / Ispeaker was a hero.'
      (Özyıldız et al. 2018)

To summarize, the pattern of indexical shift under verbs of hearing that we see in Poshkart Chuvash has not yet been described and thus can be useful for understanding the typology of controller choice in indexical shift. In the following sections, I will discuss some possible explanations for the pattern in Table 1.

4.2. A sketch of a decompositional account

In this section, I will explore an analysis of the pattern in Table 1 in terms of decomposition of complemenizers *teze* and *tenine* as syntactically forms of the verb *te* 'say', which I will ultimately reject.

As a starting point, let's focus on the question why shift of the 1st person to the ablative source under 'hear' only allows the complementizer *tenine* and disallows *teze* (row E in Table 1). It seems natural to account for this pattern by analyzing *teze* as syntactically the simultaneous (*-zA*) converb of the verb *te* 'say' and *tenine* as the past participle of *te* 'say'. The *-zA* converb is a same-subject

---

[28] Özyıldız et al. (2018) also cite Polinsky 2015 on Tsez, where 1st person indexicals can shift to the matrix subject under verbs of hearing, and Sundaresan 2018 on Tamil, where shift to the subject of 'hear' is significantly harder than under speech verb. Unfortunately, the data on shift to the ablative source is not reported in these works.

simultaneous converb in Poshkart Chuvash, which requires a null subject (represented as *PRO*) controlled by the matrix subject, as shown in (26a), cf. the unacceptability of an overt subject in (26b). By contrast, the subject of the participial nominalization has a free reference. It can be overt, as shown in (27a), or it can be dropped (as represented by *pro*) and take a non-subject or a contextually-supported antecedent, as shown in (27b).

(26) a. *vasʲa₁* [{*vəl / *PRO₁*} *divan* *ɕindʑe* *virt-sa*] *televizor* *pog-a-t*.
 Vasya 3SG.NOM sofa on lie-SIM.CVB TV watch-NPST-3SG
 'Vasya is watching TV lying on the sofa.'

 b. *[*vasʲa* *divan* *ɕindʑe* *virt-sa*] *aməş* *televizor* *pog-a-t*.
 Vasya sofa on lie- SIM.CVB mother.POSS3 TV watch-NPST-3SG
 Intended: 'His mom is watching TV while Vasya is lying on the sofa.'

(27) a. *ep* [*eki* *katɕtɕa* *kaj-n-i-ne*] *ëlt-r-em*.
 1SG.NOM sister groom go-PST.PTC-POSS3-OBJ hear-PST-1SG
 'I heard that [my] sister got married.'

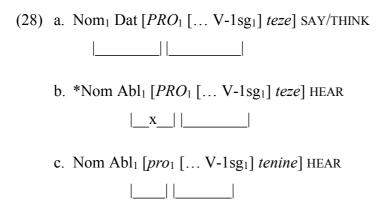 b. *ep* *ekiren₁* [*pro₁/₂* *katɕtɕa* *kaj-n-i-ne*] *ëlt-r-em*.
 1SG.NOM sister groom go-PST.PTC-POSS3-OBJ hear-PST-1SG
 'I heard from [my] sister₁ that she₁/₂ got married.'

The decompositional analysis allows to account for the unacceptability of *teze* in examples like (23a), cf. (21c) with *tenine*, in the following way. Suppose that shifted 1st person in Poshkart Chuvash is invariably controlled by the subject of *te* 'say' (including *teze* and *tenine*, under the decompositional analysis). Now, given that the subject of *teze* is a subject-oriented *PRO*, the 1st person agreement on the embedded verb will have to be controlled by the matrix subject. This correctly predicts the controller choice under speech/thought verbs, as schematically represented in (28b), cf. (21a)–(21b). Crucially, shift to the ablative source under *ëlt* 'hear' will be disallowed with *teze*, as shown in (28b), as it violates the subject-orientation requirement associated with *teze*. By contrast, the subject of *tenine*, which is a participial nominalization under the decompositional analysis, has a free reference and thus can take the ablative source as its antecedent without violating the subject-orientation requirement, as shown in (28c).[29]
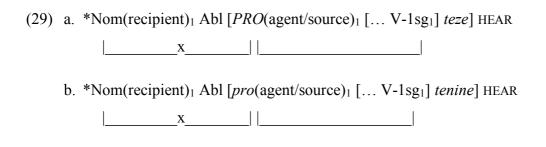
---

[29] The remaining case, corresponding to shift (to the subject) under speech/thought predicates with *tenine* (see row B in Table 2 and footnote 29), is not directly excluded by this account as there is nothing intrinsically wrong with the

(28) a. Nom₁ Dat [PRO₁ [… V-1sg₁] *teze*] SAY/THINK

        |_____| |_____|

    b. *Nom Abl₁ [PRO₁ [… V-1sg₁] *teze*] HEAR

        |__x__| |_____|

    c. Nom Abl₁ [pro₁ [… V-1sg₁] *tenine*] HEAR

        |_____| |_____|

The converb analysis of *teze* is independently supported by the fact that *teze*-clauses are disallowed with noun predicates such as *xɨbar* 'news', as we saw in (10a) above. This follows from the standard assumption that converbial clauses do not function as adnominal modifiers (the standard function of participial clauses). The converb analysis also accounts for why *teze*-clauses cannot appear as sentential subjects, as we saw in (10b). Given that *-zA* is a same-subject converb, it must be controlled by the matrix subject. However, the matrix subject position is occupied by the converbial clause itself, leading to the failure to control the subject of *teze*.

Having accounted for why *teze* is disallowed under *ëlt* 'hear' with the ablative controller (row E in Table 1), let's turn to the unacceptability of control by the matrix subject (row E). As it stands, the decompositional account predicts that it should be possible (and, in fact, obligatory) given that *teze*, being a same-subject converb (under the decompositional analysis), will force subject control. One way to exclude control by the matrix subject (the hearer) is to capitalize on the fact that *te* 'say' and hence *teze* (under the decompositional analysis) is a speech verb, whose subject argument expresses an agent or a source of speech. Under this view, it follows naturally that the subject of *teze* or *tenine* and hence the shifted 1st person (controlled by the former) cannot be coindexed with the subject of 'hear', which is a recipient of speech. This is schematically represented in (29a)–(29b).

(29) a. *Nom(recipient)₁ Abl [PRO(agent/source)₁ [… V-1sg₁] *teze*] HEAR

        |_____x_____| |_____|

    b. *Nom(recipient)₁ Abl [pro(agent/source)₁ [… V-1sg₁] *tenine*] HEAR
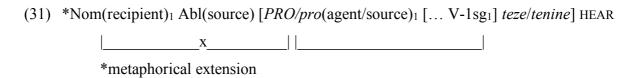
        |_____x_____| |_____|

---

pro subject of *tenine* being co-indexed with the matrix subject. To account for it, one may perhaps view these examples as instances of "subject obviation" (see, e.g., Farkas 1992), although the details remain to be worked out.

(i) *Nom [pro₁ [… V-1sg₁] *tenine*] SAY/THINK

    |__x_| |_____|

One must be careful not to incorrectly rule out indexical shift under verbs of thought/propositional attitudes, as we saw in (21b), cf. (28a). If the shifted 1st person must invariably express an agent of speech, shift to the matrix subject in such cases will be excluded. To deal with this problem, however, we can assume that the agent of speech can be metaphorically extended to an attitude holder.[30] Indeed, an attitude holder can semantically be construed as an agent of speech in the sense of "internal speech/dialogue" (see Güldemann 2008, Matić & Pakendorf 2013), which arguably provides a mechanism for extension of non-canonical SAY from speech to thought predicates. This is shown in (30).

(30) Nom(holder)$_1$ [*PRO*(agent/source)$_1$ [… V-1sg$_1$] *teze*] THINK

|_____| |_____|

metaphoric extension

The question now is what prevents the same metaphorical extension from applying in examples with indexical shift under 'hear', as in (29a)–(29b). Note that non-canonical SAY is compatible with verbs of hearing at least in some languages, including Uyghur, Turkish and, in fact, Poshkart Chuvash (in the absence of shift), suggesting that an agent of speech could also, at least in principle, be extendable to a recipient of speech, perhaps via an intermediate attitude holder construal. Indeed, there is a natural pragmatic inference from the recipient of a proposition to the attitude holder (cf. a similar proposal in Özyıldız et al. 2018, to which I will return shortly). In order to deal with this problem, however, one may tentatively assume that a metaphoric construal (recipient < attitude holder < agent/source) is blocked precisely when a *literal* construal as independently available, i.e. in the presence of an ablative source, as shown in (31).

(31) *Nom(recipient)$_1$ Abl(source) [*PRO/pro*(agent/source)$_1$ [… V-1sg$_1$] *teze/tenine*] HEAR

|_____x_____| |_____|

*metaphorical extension

To conclude, the proposed decompositional account, supplemented with the notion of metaphorical extension, goes a long way towards explaining the pattern in Table 1. Before turning to some problems, it is interesting to compare this account with Özyıldız et al. (2018)'s account of the corresponding pattern in Turkish.

---

[30] On metaphorical extension as a mechanism for the development of new constructions and subconstructions, see Goldberg 1995.

Recall that in Turkish both the matrix subject (the hearer) and the ablative source can control the shifted 1st person, as we saw in (25). To account for this, Özyıldız and colleagues decompose the complementizer *diye* as syntactically the form of the verb *demek* 'say', projecting a null subject. The important assumption that they make is that the subject of *diye* is lexically specified as the *logophoric center* of the matrix clause, i.e., the participant who is the author of the report or with respect to whose consciousness or point of view the report is being made (see Sells 1987). Now, to account for why the shifted 1st person can be controlled either by the matrix subject (the hearer) or by the ablative source, the authors argue that communication reception reports (such as 'hear', etc.) are fundamentally ambiguous in that they can be construed either as speech reports, with the ablative source as the logophoric center, or as attitude reports, with matrix subject/hearer as the logophoric center. Thus, whether shifted 1st person will be controlled by the ablative source or the matrix subject depends on the construal of the communication reception report.

The question that arises is how we can explain the difference between Turkish and Poshkart Chuvash under the decompositional account of non-canonical SAY. One possible answer is that we can view the role of the logophoric center as an advanced stage of desemantization/grammaticalization of the more concrete semantic role of the agent/source of speech — a stage which presumably follows the stage of metaphoric extension, as shown in (32). Thus, whereas *teze*/*tenine* in Poshkart Chuvash might still be analyzed under the decompositional analysis as projecting the agent of speech (even if metaphorically extended), this is no longer the case for the Turkish *diye*.
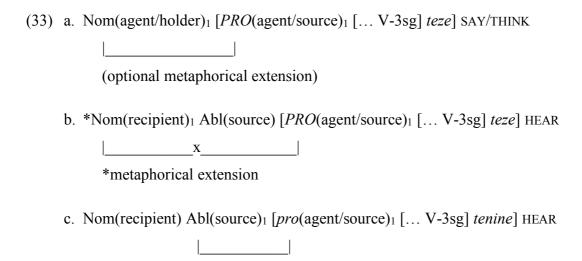
(32) logophoric center < agent of speech (metaphorically extended) < agent of speech

To summarize, the proposed analysis of *teze* as (syntactically) the same-subject converb of the verb *te* 'say' and *tenine* as a participial nominalization of the same verb seems to neatly account for the acceptability of control of the shifted 1st person by the ablative source only with *tenine* but not with *teze*. At the same time, the lack of control of the shifted 1st person by the matrix subject (the hearer/recipient of speech) with both *teze* and *tenine* follows from incomplete desemantization of the verb *te* 'say', whose subject still expresses the speaker (albeit metaphorically extended to the attitude holder). Below I turn to some problems of the proposed account.

4.3. Modifications of the decompositional account

In the previous discussion I have deliberately omitted the case of HEAR without shift (row C in Table 1), cf. (23b). If we return to this case, we can see that the decompositional account predicts

that *teze* must be ungrammatical, contrary to fact. Under this account, the subject of *teze* (*PRO*), which, by hypothesis, expresses the agent of speech, must be co-indexed with the matrix subject (the recipient of speech). By hypothesis, this should be achieved by the mechanism of metaphorical extension, cf. (30). This mechanism, however, will be blocked since there is a "literal" agent of speech in the structure (the ablative source), just like in the case with shift discussed in (31) above. As a result, the example is incorrectly predicted to be ungrammatical, as shown in (33b). Note that *tenine* is correctly predicted to be possible in this case, as the subject of *tenine* (*pro*) can be co-indexed with the ablative source, as in (33c). Similarly, in the case of SAY/THINK without shift in (33a) the understood subject of *teze* (*PRO*) can be co-indexed with the matrix subject.[31]

(33) a. Nom(agent/holder)$_1$ [*PRO*(agent/source)$_1$ [… V-3sg] *teze*] SAY/THINK

      |_____|

      (optional metaphorical extension)

    b. *Nom(recipient)$_1$ Abl(source) [*PRO*(agent/source)$_1$ [… V-3sg] *teze*] HEAR

      |_____x_____|

      *metaphorical extension

    c. Nom(recipient) Abl(source)$_1$ [*pro*(agent/source)$_1$ [… V-3sg] *tenine*] HEAR

      |_____|

How to deal with this problem? Recall that examples like (23b), represented in (33b), are somewhat marked and even rejected by some speakers. Perhaps this precisely reflects the difficulty of metaphorical extension of the agent of speech to the recipient of speech in the presence of an ablative source. In other words, we might be able to take the predictions of the decompositional account regarding examples like (33b) to be essentially correct. The problem, however, is that examples with shift under TEZE + HEAR like (23a), cf. (31), are significantly worse than examples without shift like (23b). The examples with shift under TEZE + HEAR are categorically rejected by all speakers. This difference remains unexplained under the decompositional account. Thus, either we need to seek a separate source of the extra-unacceptability of (31) or we need to drop the decompositional account of (33b) and find a different explanation of its markedness.

One possibility is to assume that *teze* has an additional *non-decompositional* analysis (as syntactically a complementizer). Under this account, the problem with metaphorical extension in

---

[31] The structure with *tenine* under SAY/THINK without shift should again be blocked by some additional mechanism like obviation, see footnote 28.

(33b) does not arise simply because *teze* need not project a subject. Note that, for this account to work, we must further assume that the non-decompositional analysis of *teze* in examples without shift, as in (33b), is not available or dispreferred for some speakers, accounting for their marked status. The crucial assumption of this account is that *the complementizer must be decomposed with indexical shift but need not be decomposed without shift* (in order to account for why *teze* is categorically blocked in examples like (31)). Accordingly, this account can be referred to as the *partial ambiguity account*. The partial ambiguity account can be viewed as a minimal modification of the account presented in the previous section, which I will now call the *uniform decompositional account*, as it leaves the explanation for the puzzle in Table 1 (see section 4.1) essentially intact.

Another possibility is that *teze* is unambiguously a complementizer, leaving aside *tenine* for the moment. At first glance, under this account we seem to lose the explanation of the unacceptability of indexical shift with *teze* under HEAR, cf. (31). However, suppose that indexical shift of 1st person is *generally* subject-oriented in Poshkart Chuvash (independently of the specific construction).[32] This correctly predicts the fact that shift to the ablative phrase with *teze* is blocked. As for the unacceptability of shift to the matrix subject (recipient of speech) in (33b), one may assume that indexical shift under HEAR is again generally blocked in Poshkart Chuvash, independently of the specific construction (indeed, this seems to be a marked option cross-linguistically, cf. Sundaresan 2018). Although indexical shift is ostensibly possible under HEAR with *tenine*, one may adopt a decompositional analysis of *tenine*, under which it is syntactically a form of the verb *te* 'say' and hence the restriction of shift under HEAR no longer applies. In its turn, the subject-orientation restriction will be circumvented by way of shifting the 1st person to the subject of *tenine*. I will call this account the *non-uniform decompositional account*, reflecting the fact that only *tenine*, but not *teze* is syntactically decomposed under this account.

To summarize, I have briefly sketched two modifications of the original uniform decompositional account. They are compared to the original account in Table 2. Both modifications seem to solve the problem of *teze* without shift under HEAR, cf. (33b), and thereby accurately capture the main puzzle in Table 1. Crucially, all the three alternatives rely on the decomposition of *tenine* (as syntactically the past participle of *te* 'say'), as we see in Table 2. In the next section, I will examine this assumption in some detail and show that it is problematic.

---

[32] This seems to be an implicit assumption in Deal 2017, 2018.

Table 2. Status of the two SAY-complementizers depending on the type of the decompositional account

Type of account	Status of complementizer	
	teze	*tenine*
uniform decompositional account	decomposed	decomposed
partial ambiguity account	decomposed only under shift	decomposed
non-uniform decompositional account	non-decomposed	decomposed

5. Against decomposition of SAY-complementizers

5.1. SAY with overt subjects

At first glance, there seems to be evidence *in favor of* treating *tenine* in examples like (21c), cf. (28c), as syntactically the past participle of the verb *te* 'say'. However, on closer inspection it turns out that this evidence argues against such an analysis.

The first piece of evidence, which I discuss in this section, concerns uses of *tenine* with an overt subject, as in (34a)–(34b). Before turning to the semantics of these uses, we need to make sure that *tenine* does have an overt subject in these uses. The reason to analyze *Boris* as the subject of *tenine* in (34a)–(34b), rather than that of the embedded verb is that we independently know that full NPs such as *Boris* cannot control 1st agreement. This is shown by the unacceptability of examples like (35), with *teze*. Whether *teze* is analyzed as a verb or a complementizer, it disallows an overt subject and, thus forcing the NP in question (*Boris*) to be analyzed as the subject of the embedded verb. Such examples, however, are strongly rejected by all speakers, suggesting that full NPs like *Boris* cannot control 1st person agreement.

(34)   a.   Context: Boris said to someone, "I will work with Sonya." Lionila said to me, "Boris said that he will work with Sonya". I am reporting to Sonya what Lionila told me.

(*sonʲa,*)   *ep*   *lionila-ran*   [*boris*   [*san-ba*   *ëɕl-e-p*]
Sonya   1SG.NOM   Lionila-ABL   Boris   2SG-INST   work-NPST-1SG

te-n-i-ne]   *ëlt-rʲ-em.*
say-PST.PTC-POSS3-OBJ   hear-PST-1SG

'(Sonya,) I heard from Lionila that Boris$_1$ said that he$_1$ will work with you$_{addressee}$.'

b. Context: =(34a)

(*son^j a,*) *lionila*   *man-a*   [*boris*   [*san-ba*   *ëɕl-e-p*]

Sonya   Lioniola   1SG-OBJ   Boris   you-INST   work-NPST-1SG

te-n-i-ne]   *kala-r^j-e.*

say-PST.PTC-POSS3-OBJ   say-PST-3SG

'(Sonya,) Lionila told me that Boris₁ said that he₁ will work with you_addressee.'

(35) **lionila*   *man-a*   [*boris*   *san-ba*   *ëɕl-e-p*]   *te-ze*   *kala-r^j-e.*

Lionila   1SG-OBJ   Boris   you-INST   work-NPST-1SG   say-SIM.CVB   say-PST-3SG

Intended: 'Lionila told me that Boris₁ (said that he₁) will work with you_addressee.'

The fact that *tenine* allows overt subjects and hence must project a subject position, thus necessitating a decompositional analysis, provides independent evidence for the analysis of *tenine* in examples without overt subjects like (21c) as also projecting a subject position realized by the null pronoun (*pro*), as shown in (36), cf. (28c).

(36) Context: Boris said to me, "I will work with Sonya." I am reporting this to Sonya.

(*son^j a,*) *ep*   *boris-ran₁*   [*pro₁* [*vəl₁*   *san-ba*   *ëɕl-e-p₁*]

Sonya   1SG.NOM   Boris-ABL   3SG.NOM   2SG-INST   work-NPST-1SG

te-n-i-ne]   *ëlt-r^j-em.*

say-PST.PTC-POSS3-OBJ   hear-PST-1SG

'(Sonya,) I heard from Boris₁ that he₁ will work with you_addressee.'

There are two reasons, however, to reject a uniform analysis of examples without overt subjects like (36) and examples with overt subjects like (34a)–(34b). First, there is a semantic difference between the two. As seen from the translation, examples in (34a)–(34b) involve *two* reports rather than one: there is an original speech act by Boris, a report of this speech act by Lionila to the actual speaker, and, finally, the secondary report of Lionila's report to Sonya by the actual speaker. By contrast, examples like (36) involve only *one* speech report: an original speech act by Boris reported by the actual speaker to Sonya. Clearly, examples like (36) need not be understood along the lines of (37), where the actual speaker is reporting Boris's report about Boris's original speech act. Rather the actual speaker is reporting Boris's original speech act ("I am going to work with Sonya") directly.

(37) #'Boris said to me, "I said that I will work with Sonya."'

What this semantic difference suggests is that *tenine* in examples with overt subjects in (34a)–(34b) is compositionally interpreted and functions as an ordinary contentful speech verb expressing a separate speech report. By contrast, *tenine* in (36) is not compositionally interpreted (at least not in the same sense as *tenine* in (34a)–(34b)): it does not express an independent speech report as distinct from the communicative reception report expressed by the verb *ëlt* 'hear'.[33] At best, *tenine* expresses the same report as *ëlt* 'hear', but viewed, as it were, from the perspective of the ablative source. Although a precise characterization of the semantic mechanism behind examples like (36) remains to be worked out, this brief discussion shows that these examples are semantically very different from (34a)–(34b). Given that lack of compositionality is usually associated with a more grammaticalized/conventionalized status, the semantic difference between (34a)–(34b) and (36) can be viewed as an argument against a decompositional analysis of *tenine* in (36) and in favor of its status as a complementizer.

Second, examples (34a)–(34b) and (36) seem to have a slightly different acceptability status. While examples like (36), with the non-compositional *tenine*, are unmarked and uniformly accepted by all my informants, examples with overt subject and compositional *tenine* appear to be dispreferred and sometimes rejected by informants (although they are definitely not ungrammatical as speakers occasionally produce them in elicitation tasks). For example, some informants are initially puzzled by examples like (34a)–(34b) and instead suggest replacing *tenine* with (*teze*) *kalanine*, i.e., with an unambiguously contentful speech verb. One informant suggested "correcting" the 1st person agreement on the embedded verb with the 3rd person agreement, as in (39), presumably avoiding the analysis of *Boris* as the subject of *tenine* and hence its construal as a contentful verb. By contrast, informants never suggested similar alternatives for examples with a non-compositional *tenine* in (36).

(38) a. *ep*   *lionila-ran*   [*boris*   [*san-ba*   *ëɕl-e-p*]
         1SG.NOM   Lionila-ABL   Boris   2SG-INST   work-NPST-1SG
         {?*te-n-i-ne*   /   *te-ze*   *kala-n-i-ne*]   *ëlt-rʲ-em*.
         say-PST.PTC-POSS3-OBJ   say-SIM.CVB   say-PST.PTC-POSS3-OBJ   hear-PST-1SG
         'I heard from Lionila that Boris₁ said that he₁ will work with you_addressee.'

---

b. *lionila*   [*boris*   [*san-ba*   *ëɕl-e-p*]   {$^?$*te-n-i-ne* /

Lioniola   Boris   you-INST   work-NPST-1SG   say-PST.PTC-POSS3-OBJ

te-ze   *kala-n-i-ne*]   *kala-rʲ-e.*

say-SIM.CVB   say-PST.PTC-POSS3-OBJ   say-PST-3SG

'(Sonya,) Lionila told me that Boris₁ said that he₁ will work with you_addressee.'

(39) *ep*   *lionila-ran*   [*boris*   *san-ba*   *ëɕl-e-t*]   *te-n-i-ne*

     1SG.NOM   Lionila-ABL   Boris   2SG-INST   work-NPST-3SG   say-PST.PTC-POSS3-OBJ

     ëlt-rʲ-em.

     hear-PST-1SG

     Intended: 'I heard from Lionila that Boris₁ (said that he₁) will work with you_addressee.'

Although this difference in speaker behavior need not directly reflect the grammatical status of the corresponding constructions, it suggests that examples with the compositional *tenine* like (34a)–(34b) cause certain processing difficulties for the speakers in a way that examples with the non-compositional *tenine* in (36) do not. This difference in processing is hard to explain if both types of *tenine* receive a uniform (decompositional) analysis. By contrast, if the two types of *tenine* are taken to be structurally distinct in that only the compositional *tenine* (with an overt subject) is syntactically a verb (past participle), whereas the non-compositional *tenine* is a complementizer, the observed difference could be accounted for by assuming that speakers generally avoid analyzing *tenine* as a verb and prefer the complementizer analysis. This assumption seems rather natural from the point of view of mechanisms of grammaticalization/conventionalization. Assuming that the non-compositional *tenine* is more grammaticalized/conventionalized than the compositional *tenine*, there must have been a strong pressure for speakers of Poshkart Chuvash to start using *teze* and *tenine* non-compositionally in the first place. Presumably, the same pressure is still present at the synchronic level, forcing speakers to use *teze* and *tenine* as complementizers. Whatever the ultimate explanation of these facts, it must rely on the availability of a non-decompositional analysis of *tenine* in examples like (36).

To summarize, the examination of the uses of *tenine* with overt subjects provides evidence against a uniform analysis of (36) and (34a)–(34b) and against the syntactic decomposition of *tenine* in examples like (36). This conclusion has important implications for the analysis of the puzzle in Table 1 (see section 4.1). Recall that all variants of the account of the puzzle (see Table 2) crucially rely on the decompositional analysis of *tenine*. By destroying evidence for this

analysis, we have also discarded the proposed account of the puzzle and thus opened up the possibility of a different (constructionist) account, to which I will come back in section 5.3.

Before turning to the second argument against syntactic decomposition of *tenine* in (36), I would like to dwell on the uses of *tenine* with overt subjects for some time. The argument above presupposed that such uses require a decompositional analysis of *tenine* (in order to accommodate the overt subject). Interestingly, there are some problems with this analysis.

Assuming that *tenine* with an overt subject in (34a)–(34b) is syntactically the past participle of the verb *te* 'say' expressing an (original) speech report, we expect that *tenine* will also be compatible with an embedded clause with an *overt* subject. Note that so far, we have only seen embedded clauses with *tenine* that had shifted 1st person agreement, cf. (34a)–(34b). Interestingly, such examples are strongly rejected by all informants, as shown in (40a). This suggests that they have a different grammaticality status compared to examples like (34a)–(34b), with a non-overt (most deeply) embedded subject, which are in principle accepted (even if dispreferred) by most informants. Note that, although examples like (40a) are expected to be difficult to parse/comprehend (due to double embedding), their unacceptability is unlikely due to processing factors alone as the corresponding examples with (*teze*) *kalanine* like (40a), cf. also (38b) above, are acceptable (in fact, (40b) was suggested by an informant as an alternative for (40a)).[34]

(40) a. *(*son^j^a*,)*   *ep*   [*lionila* [*boris*   *san-ba*   *ëɕl-e-t*]
        Sonya   1SG.NOM   Lionila   Boris   2SG-INST   work-NPST-1SG
        te-n-i-ne]   *ëlt-r^j^-em.*
        say-PST.PTC-POSS3-OBJ   hear-PST-1SG
        Intended: '(Sonya,) I heard that Lionila said that Boris will work with you_addressee.'

    b. (*son^j^a*,)   *ep*   [*lionila* [*boris*   *san-ba*   *ëɕl-e-t*]   *te-ze*
        Sonya   1SG.NOM   Lionila   Boris   2SG-INST   work-NPST-1SG   say-SIM.CVB
        kala-n-i-ne   *ëlt-r^j^-em.*
        say-PST.PTC-POSS3-OBJ   hear-PST-1SG
        '(Sonya,) I heard that Lionila said that Boris₁ will work with you_addressee.'

The ungrammaticality of (40a) is surprising if *tenine* with an overt subject is syntactically the past participle of the verb *te* 'say'. The verb *te* 'say' clearly allows embedding of clauses with overt

---

[34] It cannot be excluded that the unacceptability of in principle possible that examples like (23a) are rejected because of a combined difficulty of using *tenine* as a verb and an independent difficulty of multiple embedding, Experimental evidence.

subjects, as we saw in section 2.3, cf. (4b). There seems to be no mechanism to block a free combination of *tenine* with such clauses, as we see in examples like (40a). Thus, there is a problem for a decompositional analysis of *tenine* even in its uncontroversially non-complementizer uses.

A similar problem arises for examples like (41a) with overt (dative) addressees. Such examples are rejected as ungrammatical even though the verb *te* 'say' in principle allows dative arguments, as shown in (41c), cf. also (4d) above. Moreover, the unacceptability of (41a) cannot be due to a semantic infelicity as shown by the acceptability of the corresponding example with (*teze*) *kalanine* in (41b).

(41) a. (*sonʲa,*) *ep* *ëlt-rʲ-em* [*boris* (**vaɕa*) [*san-ba*
         Sonya 1SG.NOM hear-PST-1SG Boris Vasya.OBJ 2SG-INST

         ëɕl-e-p] *te-n-i-ne*].
         work-NPST-1SG say-PST.PTC-POSS3-OBJ

         '(Sonya,) I heard that Boris₁ said (to Vasya) that he₁ will work with you_addressee.'

      b. (*sonʲa,*) *ep* *ëlt-rʲ-em* [*boris* *vaɕa* [*san-ba*
         Sonya 1SG.NOM hear-PST-1SG Boris Vasya.OBJ 2SG-INST

         ëɕl-e-p] *te-ze* *kala-n-i-ne*].
         work-NPST-1SG say-SIM.CVB say-PST.PTC-POSS3-OBJ

         '(Sonya,) I heard that Boris₁ said to Vasya that he₁ was will work with you_addressee.'

      c. (*sonʲa,*) [*boris* *Maʂa-na* [*san-ba* *ëɕl-e-p*] *te-rʲ-e*].
         Sonya Boris Masha-OBJ 2SG-INST work-NPST-1SG say-PST-3SG

         '(Sonya,) Boris₁ told Masha that he₁ will work with you_addressee.'

The ungrammaticality of examples in (40a)–(41a) suggests that *tenine*, even in its non-complementizer uses, does not show a full syntactic productivity of the kind we expect from a genuine participial nominalization such as (*teze*) *kalanine*. To anticipate the analysis to be presented in section 5.3, such partial productivity of *tenine* with an overt subject seems to indicate that we are dealing with a conventionalized construction. In its turn, the constructionist analysis of non-complementizer *tenine* should, a fortiori, be extendable to non-compositional/complementizer

uses of *tenine* in (36), for which, as I showed in this section, a decompositional analysis is independently excluded.[35]

5.2. Logophoric third person pronoun with shifted agreement

Before I turn to the constructionist account of the puzzle in Table 1, I wish to discuss one more potential argument for the decompositional analysis *tenine* in examples like (36). Again, I will show that upon closer inspection the data argue *against* the decompositional analysis.

The argument comes from the fact that examples with indexical shift under *tenine* often contain a nominative 3rd person pronoun *vəl* '(s)he', co-indexed with the ablative source and the embedded verb, cf. (36). Interestingly, the corresponding examples with *teze*, as in (42a), are perceived as marked by the speakers: some informants do not find them fully acceptable (without, however, categorically rejecting them) and all prefer the version without *vəl*. By contrast, examples with *tenine* show the opposite pattern, as we see in (42b): all informants prefer the version with *vəl* and some find the version without *vəl* less than fully acceptable.

(42) a. (*sonʲa,*) *boris* *man-a* [([?]*vəl*) *san-ba* *ëɕl-e-p*] *te-ze*

Sonya Boris 1SG-OBJ 3SG.NOM you-INST work-NPST-1SG say-SIM.CVB

kala-rʲ-e.

say-PST-3SG

'(Sonya,) Boris₁ told me that he₁ will work with you_addressee.'

  b. (*sonʲa,*) *ep* *boris-ran* [[?](*vəl*) *san-ba* *ëɕl-e-p*]

Sonya 1SG.NOM Boris-ABL 3SG.NOM 2SG-INST work-NPST-1SG

te-n-i-ne *ëlt-rʲ-em.*

say-PST.PTC-POSS3-OBJ hear-PST-1SG

'(Sonya,) I heard from Boris₁ that he₁ will work with you_addressee.'

Before discussing how this pattern can supply an argument in favor of the decompositional analysis of *tenine*, I will focus on the analysis of *vəl* in examples like (42a) with *teze*. The analysis should capture two facts: (a) why *vəl* is in principle acceptable in these examples; (b) why it, nonetheless, has a marked status. First of all, it is clear that *vəl* in (42a) could not be the subject of

---

[35] The data presented in this section are, of course, compatible with treating the non-compositional *tenine* simply as a complementizer (and thus do not require a constructionist analysis). This analysis, however, is unable to account for the puzzle in Table 2 (see section 4), as opposed to the constructionist alternative (see section 5.3).

teze (assuming that *teze* allowed a decompositional analysis), since the *-zA* converb disallows overt subjects, cf. (26a). In addition, there is direct evidence that *teze* disallows *vəl* as its subject, as shown by the total unacceptability of examples like (43), where the embedded verb already has a subject, thus forcing the analysis of *vəl* as the subject of *teze*. As a consequence, *vəl* must at least in principle be analyzable as the subject of the embedded verb, controlling 1st person agreement.

(43) *lionila   man-a   [vəl   [boris   san-ba   ëɕl-e-t]   te-ze]
      Lioniola   1SG-OBJ   3SG.NOM   Boris   you-INST   work-NPST-1SG   say-SIM.CVB

      kala-rʲ-e.
      say-PST-3SG

      Intended: 'Lionila told me (saying) that Boris₁ said that he₁ will work with you_addressee.'

This conclusion is not as far-fetched as it might seem. Control of the 1st person agreement by a 3rd person pronoun has been described cross-linguistically. Most if not all of such patterns have been discussed in connection with the marking of logophoricity, specifically with logophoric systems which employ 1st person agreement on the embedded verb (with or without a designated logophoric pronoun), a situation referred to as "first person logophoricity" and attested in a number of African languages (see, e.g., Culy 1994, Curnow 2002 and Nikitina 2012 for an overview). Particularly interesting are the systems without a designated logophoric pronoun where the logophoric 1st person agreement is effectively controlled by an ordinary 3rd person pronoun, as illustrated in (44).[36] Commenting on patterns like (44), Nikitina (2012) observes that they should not be very unusual in view of the fact that logophoric pronouns typically develop from 3rd person/demonstrative pronouns (see, e.g., Dimmendaal 2001). Thus, 3rd person pronouns in such examples can be viewed as "emerging logophors".

(44) Karimojong (Eastern Nilotic; Uganda)

      àbò papà tolim ɛbè àlózì   iŋèz morotó
      AUX father say   that 1SG.go.NPST 3SG Moroto

      'The father said that he was going to Moroto.' (Novelli 1985: 531, quoted from Curnow 2002: 9 and Nikitina 2012:249)

---

[36] Recently, a similar case has been described in a Dravidian language Telugu (see Messick 2016).

Given the discussion above, it is plausible to analyze *vəl* with shifted 1st person ("logophoric *vəl*") in examples like (42a) as an (emerging) logophor controlling 1st person agreement.[37] This provides an answer to question (a) above.

Let's turn to question (b), namely, why "logophoric *vəl*" has a marked status in examples with *teze* (but not with *tenine* in examples like (42b)). At first glance, the answer might be provided by the decompositional analysis of *tenine*. Suppose that "logophoric *vəl*" is *generally* marked in Poshkart Chuvash. This would account for the intermediate status of *vəl* in examples with *teze*, as in (42a). By contrast, under the decompositional analysis of *tenine*, *vəl* in examples with *tenine*, as in (42b), has an option to be analyzed as the subject of *tenine* rather than a logophor. Thus, the pattern in (42a)–(42b) can potentially provide an argument in favor of the decompositional analysis of *tenine*.

I would like to argue, however, that the pattern in (42a)–(42b) can be accounted for in a natural way without assuming that *vəl* is the subject of *tenine* (i.e., by assuming that it is logophoric in both examples). Thus, it does not provide an argument for the decomposition of *tenine*.

It has been observed in the literature on anaphora resolution that when two pronouns alternate, it is usually the case that the more phonetically/structurally deficient pronoun (e.g., null pronoun) will tend to take a more prominent antecedent (e.g., the subject or discourse topic), whereas the less deficient pronoun will tend to take a less prominent antecedent (see, e.g., Mayol and Clark 2010 and references therein for Romance languages and Patel-Grosz & Grosz 2017 and references therein for a similar pattern with demonstrative and personal pronouns in German). Although this distributional difference is often stated as a preference, it can be quite robust and lead to contrasts in acceptability.

We may try to extend the same reasoning to the alternation between *vəl* vs. the null subject (*pro*) in examples like (42a)–(42b). The ablative source which controls the shifted 1st person in (42b) is syntactically less prominent than the matrix subject which controls the shifted 1st person in (42a). Thus, assuming that more deficient pronouns are generally associated with more prominent antecedents, we predict that if there were a contrast between *vəl* and the null subject in examples like (42a)–(42b), it would be such that the null subject would be more preferable in examples with a subject antecedent, as in (42a), and *vəl* would be more preferable in examples with a non-subject (ablative) antecedent, as in (42b). This is exactly what we observe.

---

[37] Interestingly, this pattern is not attested in other Turkic languages discussed in connection with indexical shifting such Turkish, Mishar Tatar and Uyghur.

Still, one has to be careful to distinguish between predicting the actual pattern and predicting the pattern *given that there is a contrast*. As such, the contrast between *vəl* vs. the null subject in (42a)–(42b) is not predicted by the contrast in the relative prominence of the potential antecedents *across* the two sentences. This is because the logic of a competition account is based on choosing the more deficient pronoun for the more prominent potential antecedent *in a given sentence*, cf. Mayol and Clark 2010. Below I will give a very brief sketch of how such a competition account might work, leaving the detailed investigation of pronominal reference in Poshkart Chuvash for future research.

If we consider the potential antecedents for *vəl* and the null subject in (42a)–(42b), we will see that the null subject controlling the 1st person agreement in (42a)–(42b) is potentially ambiguous between a shifted interpretation, where it refers to the matrix subject/the ablative source, and an unshifted interpretation, where it refers to the actual speaker. This is shown in (45a)–(45b). By contrast, *vəl* controlling 1st person agreement unambiguously has a shifted interpretation.

(45) a. (*sonʲa,*)   *ep*   *boris-ran*   [*san-ba*   *ëɕl-e-p*]   *te-n-i-ne*
   Sonya   1SG.NOM   Boris-ABL   2SG-INST   work-NPST-1SG   say-PST.PTC-POSS3-OBJ
   ëlt-rʲ-em.
   hear-PST-1SG
   '(Sonya,) I heard from Boris₁ that {I / he₁} will work with you_addressee.'

  b. (*sonʲa,*)   *boris*   *man-a*   [*san-ba*   *ëɕl-e-p*]   *te-ze*
   Sonya   Boris   1SG-OBJ   you-INST   work-NPST-1SG   say-SIM.CVB
   kala-rʲ-e.
   say-PST-3SG
   '(Sonya,) Boris₁ told me that {I / he₁} will work with you_addressee.'

Now, we can derive the pattern in (42a)–(42b) by assuming the prominence hierarchy in (46). Under (46), the matrix subject (the reported speaker) in (42a) is more prominent than the actual speaker. Hence the competition account will favor the more deficient pronoun, namely *pro*, for its realization. By contrast, in (42b) the ablative source is less prominent than the actual speaker, hence the competition account will favor the less deficient pronoun, namely *vəl*, for its realization.[38]

---

[38] One potential concern with the account above is that it predicts that the preference for *pro* or *vəl* in (42a)–(42b), respectively, is totally symmetric. Yet, the pattern of informant judgments suggests that there might be some asymmetry involved. Whereas most informants in principle allow either pronominal option in both (42a) and (42b)

(46)   subject (reported speaker) > actual speaker > ablative source

Although the proposed competition account requires a more detailed elaboration, the general point is that the distributional pattern of *vəl* with 1st person agreement can in principle be made to follow from independent principles of anaphora resolution, crucially without the need to analyze *vəl* in examples like (42b) as the subject of *tenine*. As a result, the pattern in (42a)–(42b) does not consitute evidence in favor of the decompositional account of *tenine*.

Interestingly, it turns out that the observed pattern might actually provide an argument *against* syntactic decomposition of *tenine*. According to the proposed account, an overt pronoun (*vəl*) is chosen over a null pronoun in examples with *tenine* such as (42a) because a non-subject antecedent (ablative source) is less prominent than the actual speaker (the other potential antecedent for the 1st person). Note, however, that under a decompositional analysis, *tenine* would have a null subject (co-indexed with the 1st person agreement on the embedded verb). Given the hierarchy in (46), the (null) subject of *tenine* would be a more prominent antecedent than the actual speaker and thus we would expect a null pronoun chosen over *vəl*, contrary to fact. Indeed, in examples with an overt subject of *tenine*, as in (47a)–(47b), *vəl* is dispreferred in a neutral context, with some informants suggesting that a contrastive focus on the subject and still others rejecting these examples outright. To conclude, rather than providing an argument in favor of a decompositional analysis of *tenine*, the pattern in (42a)–(42b) might actually be taken to argue against it.

(47) a. *ep*   [*boris*   [([?]*vel*)   *san-ba*   *ëɕl-e-p*]
    1SG.NOM   Boris   3SG.NOM   2SG-INST   work-NPST-1SG

    te-n-i-ne]   *lionila-ran*   *ëlt-rʲ-em*.
    say-PST.PTC-POSS3-OBJ   Lionila-ABL   hear-PST-1SG

    'I heard from Lionila that Boris₁ said that he₁ will work with you_addressee.'

   b. *lionila*   *man-a*   [*boris*   [([?]*vel*)   *san-ba*   *ëɕl-e-p*]
    Lioniola   1SG-OBJ   Boris   3SG.NOM   you-INST   work-NPST-1SG

    te-n-i-ne   *kala-rʲ-e*.
    say-PST.PTC-POSS3-OBJ   say-PST-3SG

    '(Sonya,) Lionila told me that Boris₁ said that he₁ will work with you_addressee.'

(with the preferences as described above), at least one informant consistently disallowed *vəl* in (42a) while in principle allowing both options in (42b). This pattern, if real, might reflect a general economy principle favoring the use of more deficient pronouns in default circumstances (such as Patel-Grosz & Grosz's (2017) *Minimize DP*). A full understanding of this pattern requires further study, including experimental evidence. I leave it for future research.

To summarize the discussion so far, we have arrived at an apparent paradox. On the one hand, all the potential solutions to the puzzle in Table 1 (see section 4.1) presuppose a decompositional analysis of *tenine*, i.e., that *tenine* is syntactically the past participle of the verb *te* 'say' (see sections 4.2 and 4.3). On the other hand, there is evidence against syntactic decomposition of *tenine* and in favor of treating it as a complementizer. Thus, it appears that either we lose an account of the puzzle in Table 1 or we are faced with the problematic aspects of the decompositional account.

In addition to these problems, I have shown that even uses of *tenine* with overt subjects, which are incompatible with a complementizer analysis, still resist a decompositional analysis in view of their partial productivity (cf. ban on embedding of non-shifted clauses and on overt dative addressee). In the next section I present a constructionist alternative as a potential way to resolve these apparent paradoxes.
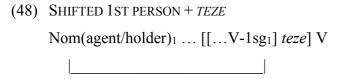
5.3. A sketch of a constructionist account

I wish to argue that the only way to meet both desiderata — to account for the restrictions on *teze* and *tenine* with indexical shift (see Table 1 in section 4.1), at the same time without syntactically decomposing them — is to encode these restrictions as part of a *construction* in the sense of a conventionalized form-meaning pairing (see, e.g., Goldberg 1995, 2006).

The crucial part of the puzzle in Table 1 concerns the subject-orientation requirement on shift of 1st person with *teze*, which captures the unacceptability of the ablative source as the potential controller under 'hear'. I proposed two possible accounts for this requirement: either by analyzing *teze* as the same-subject converb of the verb *te* 'say' (and assuming that 1st person shifts to the understood subject of *teze*) or by a general subject-orientation on indexical shift. Both accounts presuppose that *tenine* is the past participle of the verb *te* 'say', which projects its own subject, to allow the shifted 1st person with *tenine* to be controlled by the ablative source (without violating the subject orientation requirement). In other words, the problem is to enforce the subject-orientation requirement on shift with *teze* without automatically extending it to *tenine*.

The idea I wish to pursue is that the subject-orientation requirement must be stated *at the level of the construction*. In other words, it is directly encoded in the indexical shift construction with *teze*, shown in (48), rather than derived from decomposition of *teze* as the same-subject converb or from a general subject-orientation requirement on indexical shift in Poshkart Chuvash. The specification in (48) correctly captures the fact that shifted 1st person with *teze* will not be able to be controlled by the ablative source.

The construction in (48) must also have an additional constraint, restricting control of the shifted 1st person to the agent of speech or the attitude holder (metaphorically construed as the agent of speech). This ensures that the subject of a verb of hearing (the recipient of speech) cannot control shifted 1st person (due to the presence of an ablative source, or the agent of speech). Note that this constraint is not extended to the construction with *teze* under 'hear' without indexical shift (cf. the discussion around (33b) above) precisely because it is a specific property of the construction with indexical shift in (48).

(48)   SHIFTED 1ST PERSON + *TEZE*

Nom(agent/holder)$_1$ … [[…V-1sg$_1$] *teze*] V

|_____|

Constraint: Shifted 1st person is construed as the agent of speech act or — metaphorically — as the attitude holder.

At first glance, encoding the restrictions on shifted 1st person with *teze* as part of the respective construction as opposed to deriving them from decomposition of *teze*/*tenine* or general properties of indexical shift in Poshkart Chuvash (as was suggested in section 4.3) might look like a restatement of the puzzle in Table 1. However, note that the constructionist approach has means to derive the relevant properties of the construction in (48) from the properties of other independently existing constructions in the language. In other words, the properties of (48) are *motivated* by other constructions in Poshkart Chuvash (see, e.g., Goldberg 1995).

Starting from the restriction on the referent of the shifted 1st person to the agent of speech, we may take it to be general property of all constructions with shifted 1st person under the verb *te* 'say', including its grammaticalized/conventionalized uses such as *teze* in (48) and *tenine* in (49b) as well as its contentful uses in (49c), cf. (21d).[39] The relevant constraint can be encoded as the property of the more abstract construction in (49a) inherited by its more concrete instances, including the construction with *teze* in (48) and the construction with *tenine* in (49c). Note that we do not need to specify that shifted 1st person with *tenine* will be controlled by the ablative source as this follows from the fact that the ablative source is the agent of speech.

---

[39] To this list we may also include the construction with *tenibe*, cf. (12a). Indexical shift with other forms of *te*, such as *teni* and *tenə*, cf. (10b) and (10a), has not been investigated.

(49) SHIFTED 1ST PERSON + SAY (*TE*-)

    a. …NP(agent/holder)… […V-1sg$_1$] *te*   (generalized form)

        |_____|

    Constraint: Shifted 1st person is construed as the agent of speech act or — metaphorically — as the attitude holder.

    b. …NP(agent)… […V-1sg$_1$] *te*$_{contentful}$

        |_____|

    c. …NP(agent)… [[…V-1sg$_1$] *tenine*] V$_{hearing}$

        |_____|

In informal terms, the function of the constructions in (48)–(49) can be characterized as reporting a "perspectival" speech or thought act in the sense of originating from an individual X which is the subject argument in the reported proposition and from whose perspective the respective propositional content is presented.[40] For example, (14b) reports the propositional content 'Boris will work with Sonya' as given from Boris's perspective and originating from Boris. The constraints on the shifted 1st person in (48)–(49) can then be viewed as a way of specifying how the identity of the "perspectival center" is determined. The constructions in (48)–(49) can, in turn, be viewed as instances of still more general speech/thought report constructions, expressed by the verb *te* 'say' and its forms (*teze*, *tenine*, etc.) with a complement unspecified for the perspective (i.e., with or without shifted 1st person). However, a comprehensive treatment of the entire network of related constructions in Poshkart Chuvash is beyond the scope of this paper.

Now, turning to the subject-orientation requirement of shift with *teze* in (48), it can be taken to be motivated by the same-subject construction with the *-zA* converb, as shown in (50), cf. (26a)–(26b). Given that *teze* is derived from the *-zA* converb of the verb *te* 'say', it is clear that speakers of Poshkart Chuvash at least at some point of its development used to analyze it as such and thereby complied with the same-subject requirement (assuming that the shifted 1st person was linked to the understood subject of *teze*). It is natural to assume that throughout the process of grammaticalization/conventionalization of *teze*, this requirement has stuck as a residue of the diachronic origin of the construction.

---

[40] For a formal description of the semantics of constructions with indexical shift including the problem of *de se* attitudes see Schlenker 2011.

(50)   Nom$_1$… [PRO$_1$ …V-zA] V

      |_____|

Two important points are in order. First, under a (diachronic) constructionist account, we do not need to assume that *teze* is syntactically a same-subject converb to account for its subject-orientation requirement. This is because an uncontroversial category change is in principle compatible with the preservation of some grammatical properties of the original element. As an example, consider the subcategory of Russian prepositions derived from converbs (see, e.g., Kortmann & König 1992). This class includes prepositions which assign accusative case such as *ne sčitaja* 'not including' and *isključaja* 'excluding', which are related to transitive verbs *sčitat'* 'count' and *isključat'* 'exclude' (cf. *isključaja Vasju*$_{\text{ACC}}$ 'excluding Vasya'). These elements are uncontroversial prepositions by a number of diagnostics (e.g., they do not require a controller, lack perfective counterparts, etc.). As a result, the most natural explanation of the case-marking properties of these prepositions is that they have retained, or inherited, these properties from the respective verbs despite undergoing grammaticalization (semantic bleaching, category change, etc.).[41] Thus, the constructionist perspective allows us to provide a motivation for the same-subject requirement without adopting an otherwise problematic decompositional analysis or *teze*/*tenine*.

Second, a diachronic account of the subject-orientation requirement like the one above is not merely compatible with but necessitates a constructionist perspective because it is simply impossible to state the relevant constraint at the level of the complementizer (*teze*) as it involves the *combination* of the complementizer and the embedded verb with a 1st person marking, i.e., a construction. Importantly, while the proposed account does not argue against analyzing *teze* and *tenine* as complementizers, it crucially assumes that control of the shifted 1st person cannot be stated as a general property of indexical shift in Poshkart Chuvash but must be encoded as a property of the constructions with the respective complementizers. This ensures that the subject-orientation requirement will be restricted to *teze* and will not be automatically extended to *tenine*.

Before concluding, I wish to briefly come back to the uses of *tenine* with overt subjects, as discussed in section 5.1, cf. (34a)–(34b). Previously I have shown that although these uses seem to require a decompositional analysis — as they clearly disallow a complementizer analysis and have a compositional interpretation — there are still problems with analyzing *tenine* in such uses as a past participle given their limited syntactic productivity (e.g., ban on non-shifted complements and dative addressees). As is well known, partial productivity is a hallmark feature of a

---

[41] Interestingly, some deverbal prepositions such as *blagodarja* 'thanks to', derived from the transitive verb *blagodarit'* 'to thank', have lost their accusative case-marking in favor of the dative case.

construction (Goldberg 1995, 2006). What I would like to suggest is that the apparent paradox can be resolved by extending the constructionist approach proposed above to these cases as well. Specifically, I will assume that Poshkart Chuvash has a special construction for expressing doubly embedded "perspectival" speech reports (in the sense above), as shown in (51).

(51)   [Nom(agent)… V-1sg *tenine*] V~speech/hearing~
         |_____|

     Constraint: Shifted 1st person is construed as the agent of the original speech act, which is reported by a secondary speech act (reported by the actual speaker).

The distinctive feature of the construction in (51) is that it realizes the "perspectival center" as the subject of *tenine*, which is thus viewed as a past participle of the verb *te* 'say'. The important thing, however, is that while *tenine* is analyzed here as syntactically a participle, it is still part of a conventional form-meaning pairing with its partly idiosyncratic properties. This would account for why the uses of *tenine* with overt subjects are not automatically extended to non-shifted (non-perspectival) embedded clauses and are not used with overt addressees. In addition, its marked status (see section 5.1) can be viewed as reflecting a particular pragmatic/register feature, which can also be part of a constructional specification (Goldberg 2006).

6. Conclusion

I started this paper by discussing the conceptual shift in the study of "grammaticalized SAY" from the emphasis on the distinction between a lexical verb 'say' vs. a grammatical element (complementizer) to the emphasis on the degree of conventionalization of a particular use pattern involving non-canonical SAY (see Matić & Pakendorf 2013). As Matić & Pakendorf observe, this view is motivated by the fact that particular instances of non-canonical SAY can lack full compositionality without being grammaticalized (i.e., without full word class change, changes in inflection, phonological erosion, etc.). Implicit in this view, I believe, is the notion that characterizing a particular instance of non-canonical SAY in terms of the dichotomy *verb vs. complementizer* may not be the most fruitful approach as it fails to capture the fact that grammatical change may target larger units in the process Traugott and Trousdale (2013) called *constructionalization* (i.e., emergence of new constructions), cf. Heine & Kuteva's (2005) notion of *use pattern* cited by Matić & Pakendorf 2013.

The present paper may be seen as a concrete instantiation of this view on the basis of the properties of two instances of non-canonical SAY in the Poshkart variety of Chuvash, namely *teze*

and *tenine*, which are identical to the same-subject converb and the past participle nominalization of the verb *te* 'say'. The argument in the paper dealt with the patterns of indexical shift shown by the complements introduced by *teze* and *tenine*. In particular, it concerned the "switch" from *teze* to *tenine* with indexical shift under verbs of hearing, accompanied by the corresponding "switch" in the choice of the controller of the shifted 1st person from the matrix subject to the ablative source, the pattern which constituted the main puzzle of this paper. In the core part of the paper (sections 4.2 and 4.3), I presented several ways to account for this puzzle, which relied on the idea that either both *teze* and *tenine* or only *tenine* are syntactically decomposed as the respective forms of the verb *te* 'say'. Then I presented several problems with the decompositional analysis of *tenine* (entailed by all the variants of the decompositional account). In particular, *tenine* has a non-compositional interpretation and has a less marked acceptability status, as compared to the "compositional *tenine*" (with overt subjects). In addition, the use of the (logophoric) *vəl* with shifted 1st person is better accounted for under the complementizer analysis of *tenine*.

As a more promising alternative to the decompositional account, I have argued for a constructionist approach where the restrictions on the controller choice with *teze* and *tenine* are directly encoded as part of the respective constructions (conventionalized form-meaning pairings). The advantage of this view is that it allows to motivate the observed restrictions on *teze*/*tenine* — as inherited from other constructions in Poshkart Chuvash (in particular, the same-subject construction and other "perspectival" speech/report constructions) — without, however, assuming otherwise problematic syntactic decomposition.

It must be admitted that the theoretical point advanced in this paper may not strike some readers as particularly dramatic. In the end, one could assume a constructionist approach as a starting point and not even consider a decompositional analysis of non-canonical SAY in the first place. Still, although I have presented the data from Poshkart Chuvash in the form of an argument for a constructionist approach, the goal of this paper was not so much to argue for a theory as to draw the attention of researchers in the domain of indexical shift and non-canonical SAY to a number of issues that, to my mind, deserve more research, namely, the choice of the "controller" for a shifted indexical; the distribution of several instances of non-canonical SAY in a given language; and, last but not least, the interaction between the domain of indexical shift and the domain of non-canonical SAY in general, which so far have been mostly divided.

Glosses

abl	ablative	obj	objective
acc	accusative	pass	passive
aux	auxiliary	pf	perfective
cop	copula	pl	plural
cvb	converb	poss	possessive
csl	causal	pred	predicative
fut	future	prs	present
gen	genitive	pst	past
inst	instrumental	ptc	participle
ipf	imperfective	refl	reflexive
loc	locative	sg	singular
neg	negation	sim	simultaneous
nom	nominative	vr	verbalizer
npst	non-past		

References

Aikhenvald, Alexandra. 2008. Semi-direct speech: Manambu and beyond. *Language Sciences* 30. 383–422.

Anand, Pranav & Andrew Nevins. 2004. Shifty operators in changing contexts. In *Proceedings of Semantics and Linguistic Theory XIV*, ed. Robert B. Young, 20–37. Ithaca, NY: CLC Publications.

Anand, Pranav. 2006. De *de se*. Doctoral Dissertation, MIT.

Ašmarin, Nikolaj I. 1898. *Materialy dlja izsledovanija čuvašskogo jazyka*. Kazan: Tipo-litografija imperatorskago universiteta.

Chappell, Hilary. 2008. Variation in the grammaticalization of complementizers from verba dicendi in Sinitic languages. *Linguistic Typology* 12. 45–98.

Chisarik, Erika & Wim van der Wurff. 2003. From 'say' to 'because': Grammaticalisation and reanalysis. Paper presented at the Conference on Comparative Diachronic Syntax, Leiden. Press. http://www.let.leidenuniv.nl/ulcl/events/compdiachr/programme.htm

Culy, Christopher. 1994. A note on logophoricity in Dogon. *Journal of African languages and linguistics* 15:113–125.

Curnow, Timothy Jowan. 2002. Three types of verbal logophoricity in African languages. *Studies in African Linguistics* 31. 1–25.

Deal, Amy Rose. 2017. Shifty asymmetries: universals and variation in shifty indexicality. Available at: http://linguistics.berkeley.edu/~ardeal/papers/Deal-indexicals-2017.pdf

Deal, Amy Rose. 2018. Indexiphors: notes on embedded indexicals, shifty agreement, and logophoricity. In Rodica Ivan (ed.): *UMOP 40: The Leader of the Pack: A Festschrift in Honor of Peggy Speas*, pp. 59-86. Amherst, MA: GLSA.

Dimmendaal, Gerrit J. 2001. Logophoric marking and represented speech in African languages as evidential hedging strategies. *Australian Journal of Linguistics* 21. 131–157.

Evans, Nicholas. 2006. View with a view: Towards a typology of multiple perspective constructions. In Rebecca T. Cover & Yuni Kim (eds.), Proceedings of the thirty-first annual meeting of the Berkeley Linguistics Society, 93–120. Berkeley: Berkeley Linguistics Society.

Evans, Nicholas. 2013. Some problems in the typology of quotation: A canonical approach. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), Canonical morphology and syntax, 66–98. Oxford: Oxford University Press.

Farkas, Donka F. 1992. On obviation. In Lexical matters, ed. I.A. Sag and A. Szabolcsi, 85–110. Stanford, CA: CSLI Publications.

Goldberg, Adele E. 1995. Constructions: A Construction Grammar Approach to Argument Structure. Chicago: University of Chicago Press.

Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Güldemann, Tom. 2008. *Quotative indexes in African languages: A synchronic and diachronic survey*. Berlin: Mouton de Gruyter.

Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.

Heine, Bernd & Tania Kuteva. 2007. *The genesis of grammar: A reconstruction*. Cambridge: Cambridge University Press.

Huitink, Janneke. 2012. Modal concord: A case study of Dutch. *Journal of Semantics* 29. 403–437.

Kaufmann, Magdalena, and Claudia Poschmann. 2013. Embedded imperatives: Empirical evidence from colloquial German. *Language* 89. 619–637.

Khanina, Olesya. Ms. Sentencial'nye aktanty čuvašskogo jazyka.

Kortmann, Bernd & Ekkehard König. 1992. Categorial reanalysis: The case of deverbal prepositions. *Linguistics* 30: 671-697.

Kožemjakina, Nastja. 2017. Nekotorye dannye o konstrukcijah s sentencial'nym aktantom v chuvašskom jazyke (malokaračinskiij dialekt). Paper presented at the fieldwork seminar in Poshkart.

Landau, Idan. 2015. *A Two-Tiered Theory of Control*. Cambridge, MA: The MIT Press.

Maier, Emar. 2015. Quotation and unquotation in free indirect discourse. *Mind and Language* 30:345–373.

Maier, Emar. 2016. A plea against monsters. *Grazer Philosophische Studien* 93:363–395.

Matić, Dejan, and Brigitte Pakendorf. 2013. Non-canonical SAY in Siberia: Areal and genealogical patterns. Studies in Language 37:356–412.

Mayol, Laia, and Robin Clark. 2010. Pronouns in Catalan: Games of partial information and the use of linguistic resources. *Journal of Pragmatics* 42:781–799.

Messick, Troy. 2016. The Morphosyntax of Self-ascription: A Cross-linguistic Study. Doctoral Dissertation, University of Connecticut.

Nikitina, Tatiana. 2012. Personal deixis and reported discourse: Towards a typology of person alignment. *Linguistic Typology 16*. 233–263.

Novelli, B. 1985. *A grammar of the Karimojong language*. Berlin: Reimer.

Özyıldız, Deniz, Travis Major & Emar Maier. 2018. Communicative reception reports as hearsay: Evidence from indexical shift in Turkish. In Proceedings of West Coast Conference on Formal Linguistics 36.

Patel-Grosz, Pritty & Grosz, Patrick G. 2017 Revisiting Pronominal Typology. *Linguistic Inquiry* 48. 259–297.

Podobryaev, Alexander. 2014. Persons, imposters, and monsters. Doctoral Dissertation, MIT, Cambridge, MA.

Polinsky, Maria. 2015. Embedded finite complements, indexical shift, and binding in Tsez. *Languages of the Caucasus* 1:1–37.

Remizova, Tatiana. Ms. Posessivnye markery i posessivnye konstrukcii v malokaračkinskom govore čuvašskogo jazyka.

Schlenker, Philippe. 2011. Indexicality and *de se* reports. In *Semantics: An international handbook of contemporary research*, ed. K. von Heusinger, C. Maienborn, and P. Portner. Berlin: de Gruyter.

Sells, Peter. 1987. Aspects of logophoricity. *Linguistic Inquiry* 18. 445–479.

Shklovsky, Kirill & Yasutada Sudo. 2014. The Syntax of Monsters. Linguistic Inquiry 45. 381–40.

Spronck, Sjef & Tatiana Nikitina. 2019. Reported speech forms a dedicated syntactic domain. *Linguistic Typology* 23. 1–41.

Sudo, Yasutada. 2010. The Syntax and Semantics of Indexical Shifting in Modern Uyghur. Generals paper, MIT, Cambridge, MA.

Sundaresan, Sandhya. 2018. An alternative model of indexical shift: Variation and selection without context-overwriting. Available at: https://ling.auf.net/lingbuzz/004115

Traugott, Elizabeth C. & Graeme Trousdale. 2013. Constructionalization and Constructional Changes. Oxford: Oxford University Press.

Yıldırım-Gündoğdu, Hilal. 2017. The structure of *diye*-clauses in Turkish. MA Thesis, Boğaziçi University.