Chapter 27

Chuvash and the Bulgharic languages

Alexander Savelyev

Abstract

Chuvash is the sole living representative of the Bulgharic branch, one of the two principal branches of the Turkic family. While extinct Bulgharic varieties are only limitedly accessible to linguists due to the scarcity of written sources, Chuvash has attracted a lot of attention as the most divergent Turkic language notable for both archaic features and innovations against the background of the Common Turkic branch. For a century, due to its salient features, Chuvash has played a key role in the Altaic and, further, Transeurasian debate. This chapter provides a historical overview and a grammar sketch of Chuvash as seen from both a genealogical and areal perspective. As shown in the chapter, Chuvash exhibits numerous features that are typical to Transeurasian languages in general and Turkic in particular. At the same time, some other traits that can be found in the modern Chuvash may have been adopted through contact with non-Transeurasian languages, most notably Uralic, such as Mari, and Indo-European, such as Russian.

Keywords: Chuvash, Bulgharic languages, Turkic languages, Transeurasian hypothesis, language contact

27.1 Introduction

Chuvash is the only modern representative of the Bulgharic branch, which is considered the earliest to split from the Proto-Turkic language. The traditional term “Bulghar(ic)” is preferred in this chapter to the label “Oghur(ic)” that has been used in the recent decades for the same grouping by some western scholars (Golden 1992: 17; Golden 1998: 18; Clark 1998a: 434; Johanson, this volume: Chapter 8)[1]. The name “Chuvash” is an adaptation of the endonym *čâ^{o}vaš*, and similar designations of the Chuvash people and their language are used by all their neighbors. Another term

vet'ke 'the Chuvash', *ved'eń* 'Chuvash (adj.)' of unclear but certainly early origin (Napol'skikh 2015: 69) survives in Erzya Mordva.

Chuvash is spoken primarily in the middle Volga River basin (the Volga–Kama region) in east-central European Russia. About half of the 1.5 million Chuvash population (figures for 2010) inhabit the Republic of Chuvashia where their language has an official status. Smaller Chuvash-speaking groups traditionally reside in neighboring provinces of the Volga–Kama region, particularly republics of Tatarstan and Bashkortostan as well as Samara, Ulyanovsk, Orenburg, Saratov, and Penza oblasts (see Figure 27.1).

<Insert Figure 27.1 here>

Figure 27.1 The geographical distribution of Chuvash in the Volga–Kama region

The number of Chuvash speakers was estimated as 1.04 million by the 2010 census, which makes it one of the largest minority languages in Russia. Having said that, it is a highly vulnerable language because of the dominant position of Russian in the region. From 2002 to 2010, Chuvash lost about a quarter of its speakers. Recent sociolinguistic studies (Alòs i Font 2015; 2016) indicate very rapid and extensive language shift in Chuvashia, not to mention Chuvash-populated areas outside the republic (Dolgova et al. 2004: 78–81).

The extinct relatives of Chuvash, the Bulghar varieties—referred to as "West Old Turkic" by Róna-Tas and Berta (2011)—were spoken mainly in East and South European steppelands from the mid-first to the early second millennium AD. They are known from very scarce and sometimes uncertain written sources, and remarkably better due to numerous borrowings in the neighboring languages (see Section 27.2). Proper Chuvash has been attested in Latin and Cyrillic scripts since the early 18th century; for overviews of pre-standard Chuvash sources, see Jegorov (1949), Sergejev (2004), and Savelyev (2014). A standardized Cyrillic-based writing system was created in the early 1870s by the Christian missionary and educator Ivan Yakovlev; with only minor

modifications, it has survived to this day.

Chuvash is one of the best-studied Turkic languages, with its first descriptive grammar issued as early as the mid-18th century (Sočinenija 1769). The cornerstone of both synchronic and historical Chuvash studies are the works by Nikolaj Ašmarin, who demonstrated close affinities between the modern Chuvash and medieval Bulghar varieties (1902), prepared detailed monographs on Chuvash grammar in general (1898) and syntax in particular (1903; 1923) and, finally, published the colossal 17-volume *Thesaurus Linguae Tschuvaschorum* (1928–1950). General surveys of Chuvash as seen in a broader Turkic and Altaic context were made by Ramstedt (1922), Poppe (1925), Jegorov (1971), and Fedotov (1980). Levitskaja (1976; 2014) and Mudrak (1993; 2002a) provided comprehensive studies on the evolution of Chuvash phonology and morphology. There are two etymological dictionaries of Chuvash, although both are somewhat outdated (Jegorov 1964; Fedotov 1996). More up-to-date etymologies for the inherited vocabulary of Chuvash can be found in Starostin et al. (2003), while Scherner (1977), Rédei and Róna-Tas (1982), Róna-Tas (1982), and Agyagási (2005) provide research into different layers of Chuvash borrowed vocabulary. The studies focused on the loans from Chuvash and extinct Bulgharic varieties into various other languages are, *inter alia*, Wichmann (1903), Gombocz (1912), Räsänen (1920), Dobrodomov (1974), Fedotov (1990), Dybo (2010), and Róna-Tas and Berta (2011). The system of Chuvash dialects was depicted by Kanjukova (1965), Sergejev (1969), and the group of authors behind the 5-volume *Materialy po čuvašskoj dialektologii* edition (Pavlov et al. 1960–1997). Krueger (1961) remains the only detailed description of Chuvash grammar available in English. The recently published grammars of Chuvash that need to be mentioned are Sergejev et al. (2012) (in Chuvash) and Pavlov (2014; 2017) (in Russian).

27.2 Historical connections: genealogy and contact

Early scholarship from the 18th century associated Chuvash with the Uralic languages, being unable to disentangle complicated areal phenomena in the Volga-Kama region (see, e.g., language

groupings in Pallas 1787–1789). The Turkic origin of Chuvash was proposed no later than by Klaproth in 1828 and convincingly proved by Schott in 1841. In 1863, Feizkhanov managed to read three grave epitaphs in the Volga Bulghar language based on his knowledge of the contemporary Chuvash. Strong arguments relating Chuvash to Volga Bulghar were summarized by Ašmarin in 1902; since then, the Volga Bulghar → Chuvash linguistic continuity has gained general acceptance in the field.

Together with its extinct relatives, Chuvash forms the separate Bulgharic branch of the Turkic family, which exhibits many differences from the so-called Common Turkic languages. To take the phonological level alone, Bulgharic is characterized by such striking features as rhotacism (Chu. *r* vs. CTk **z*), lambdacism (Chu. *l* vs. CTk **š*), palatalization of dental consonants (PTk **t*, **d* > Chu. *č*, PTk **s* > Chu. *š* before PTk **i*, **i̇*), retention of Proto-Turkic falling diphthongs (PTk **i̯V* > Chu. *yV*-), and extensive reshuffling of the vowel system (Dybo 2010: 83). Because of the complicated phonological changes over the history of the Bulgharic branch, it was long assumed that there were, in fact, no regular correspondences between Chuvash and Common Turkic. Although contradicting one of the basic assumptions of the historical comparative approach, this view enjoys some popularity even today. For a set of regular, even if rather complex, correspondences between the two major branches of Turkic, see Mudrak (1989; 1993; 2002a: 677–706), Dybo and Mudrak (2006: 9–227), Dybo (2007: 10–64).

It is generally agreed that the homeland of the Late Proto-Turkic language, the common ancestor of Proto-Bulgharic and Proto-Common Turkic, was located in the eastern part of the Eurasian steppes (see Robbeets et al., this volume: Chapter 43 for the details). The split of Proto-Bulgharic from Proto-Turkic can be dated approximately to between 2500 and 2000 years ago, with quantitative approaches pointing to around 100–0 BC (Dybo 2007: 66; Mudrak 2009: 181; Savelyev, this volume: Chapter 9). The earliest history of Proto-Bulgharic speakers is a matter of long-standing discussion; in particular, the question of whether Proto-Bulgharic was a major lexical donor for other languages in the region divides scholars into two camps. The first view stems from

the Proto-Turkic reconstruction that essentially favours Common Turkic data as opposed to the evidence from the Bulgharic branch, and it comes along with a skeptical eye on the Turko-Mongolic genealogical relationship. The proponents of this viewpoint consider Proto-Bulgharic as a primary source of non-Common Turkic loanwords found in such families as Mongolic and Samoyedic (e.g., Janhunen 2010: 293). The other view is based on paying special attention to Chuvash evidence in reconstructing Proto-Turkic and correlates with a pro-Altaicist position. In this perspective, it is Proto-Turkic rather than Proto-Bulgharic that is regarded as the source for the first layer of Turkic borrowings in Proto-Samoyedic (Dybo 2007: 135–154), while ancient lexical parallels between Turkic and Mongolic are explained by inheritance and—in some interpretations, e.g., Savelyev (2017: 136–137)—partly as Proto-Turkic loanwords in Proto-Mongolic.

The coming of the Bulgharic-speaking nomadic tribes to the European steppes is commonly associated with the Hunnic or immediately post-Hunnic period. While "Hunnic", i.e. the language of the European Huns of the late 4th and 5th centuries AD, is sometimes referred to as a Bulgharic variety, this connection is based mainly on historical rather than proper linguistic evidence. Extremely scarce remnants of "Hunnic"—a number of personal names and only a few common nouns—do not provide any conclusive proof and may actually be linguistically heterogeneous. Soon after the collapse of the Hunnic Empire in the mid-450s, different tribal confederations previously unattested in the European sources came on the scene; although being presumably Bulgharic-speaking, few of them left a real linguistic trace. One of the historical groups for which a Bulgharic affiliation has been widely accepted are the Avars, who arrived in the Ponto-Caspian steppe zone in the 550s, then conquered the Carpathian Basin and stayed there until the Hungarian conquest in 895–902 (Golden 1992: 108–112; Róna-Tas and Berta 2011: 36–38). Avar is known primarily from a few titles and personal names as well as from a short inscription in Greek letters. For its recent reading based on Chuvash data and for an overall interpretation of the Avar language (or at least one of the languages spoken by the Avars) as a Bulgharic variety, see Mudrak (2005a).

Similarly, little is known on the language of the Khazars who dominated the Pontic steppe and

the Northern Caucasus region from the late-7th to the 10th century. The list of Khazar titles, proper nouns, and rare common nouns displays, however, some apparently Bulgharic features, as in the classic example of *Šar-kel*. That is the name of a Khazar fortress in the lower reaches of the Don, referred to as *Бѣла Вѣжа* 'White Tower' in Old Russian sources and inevitably linked to the Chuvash words *šurə̂* 'white' (< VBulg. **šårï*, cf. CTk **sarïɣ* 'yellow') and *kil* 'home' (< 'house'; the word is only scarcely attested in Common Turkic).

The proper Bulghar dialects have their roots in the short-lived empire established by Kubrat in the 630s. This polity, known as Great Bulgharia to Greek sources, is traditionally localized along the Kuban River in the Northwest Caucasus region and in the Eastern Pontic steppe zone (Golden 1992: 244). Kubrat's empire dissolved soon after his death ca. 650 and the Bulghars split into five groups, each led by one of Kubrat's sons. Two of these groups are of particular interest from a historical linguistic perspective. The first, led by Asparukh, moved westwards and established the First Bulgharian empire in the Danube River basin ca. 679. The Turkic-speaking elite ruled over the kingdom for several centuries until they were assimilated by the Slavs. The Danube Bulghar language is poorly attested in titles, proper names, and several short texts written in the Cyrillic and Greek scripts; what is more informative, it is regarded as a major source of early Turkic borrowings in Hungarian and South Slavic varieties, including Old Church Slavonic. Another group of the Bulghars moved northwards and reached the Volga-Kama region in the 8th century (Zimonyi 1990: 180; Kazakov 1997: 36). The Volga Bulghars were first dominated by the Khazar Khaganate, but they achieved independence by the mid-10th century. Their empire gained control over vast territories along the Volga and Kama Rivers, with its metropole in present-day southern Tatarstan. The language of the Volga Bulghars is limitedly attested in epitaph inscriptions in the Arabic script dating from the 13th–14th centuries (Erdal 1993). These records clearly indicate Volga Bulghar as a Chuvash-type Turkic language, cf. VBulg. *sekir* 'eight' > Chu. *sakə̂r* 'eight' vs. CTk **sekiz*, VBulg. *tüweti-* > Chu. *tə̂ᵒvatə̂* 'four' vs. CTk **tȫrt ~ dȫrt*, etc. Volga Bulghar is also known due to borrowings in Permic and, to a lesser extent, Mordvinic languages (Róna-Tas 1988: 760–768). In

addition, it served as a lexical donor for Old Russian (see Agyagási 2016, including a detailed bibliography of the earlier research) and as a substratum for later Volga Kipchak dialects (Róna-Tas 1976). A language with a phonological system transitional from Volga Bulghar to the contemporary Chuvash was the source for the main layer of Bulgharic borrowings in Mari (Róna-Tas 1988: 768–771; Savelyev 2018).

Internal relations between the extinct members of the Bulgharic branch are problematic, and the ongoing attempts to classify them (cf. Robbeets 2015: 11) are based primarily on historical sources, not linguistic evidence per se. It should, however, be kept in mind that the names attested in early medieval chronicles, such as "Avar" or "Khazar", referred to multilingual polities, not to linguistically monolithic groups, and Bulgharic varieties labeled differently because of historical considerations may, in fact, have been very similar to each other. Judging by the scarce data available, there are relatively minor differences between the extinct Bulgharic languages, such as between Danubian Bulghar and Volga Bulghar (Mudrak 2005b; Dybo 2010).

The Chuvash language is generally regarded as an immediate—and only—descendant of Volga Bulghar. The replacement of the name of the people and their language was probably caused by the cataclysmic events that befell the Middle Volga region in the 13th–14th centuries: the Volga Bulgharian empire was destroyed in the Mongol invasion of 1236, after which its former domains were devastated by a series of Russian military campaigns, by the civil war in Golden Horde in 1359–1380, and by Timur in the 1390s. This period was marked by extensive migration of Bulgharic-speaking groups to the forest area in present-day Chuvashia, where they got into interaction with the Mari people. There is a controversy as to what extent the contact with the Uralic speakers, along with the long evolution of the Bulgharic branch independently from the other Turkic languages, may have contributed to the peculiarities distinguishing Chuvash from Common Turkic. For different interpretations and conclusions, see Agyagási (1998), Johanson (2000a), Mudrak (2002a).

The basic dialectal division of Chuvash is between Viryal (Upper) and Anatri (Lower) dialects.

The former occupies the northwest of Chuvashia, and the latter is spoken in its southeast. Outside of Chuvashia, the Anatri features are dominant, though there are Viryal-speaking areas as well. The split between the two dialects, which are mutually completely intelligible, can be provisionally dated to the late 16–early 17th century. Viryal is a more archaic variety, while Anatri is characterized by a few innovations, such as the change of Proto-Chu. **o* > *u* (see Section 27.3.2). Other dialectal varieties debated in the literature are Anat Yenči (Middle-Lower) and Poškărt (Northwestern). Anat Yenči is spoken in the northeastern part of Chuvasia and is described by some Chuvash scholars as a separate dialect. Yet, it is more often interpreted as a group of mixed or transitional varieties between Viryal and Anatri (Kanjukova 1965; Sergejev 1969). The highly idiosyncratic nature of the Poškărt variety, spoken at the northwestern extremity of Chuvashia, has been traditionally attributed to a Mari substrate imposed on Viryal speakers (Ašmarin 1898). However, it was recognized as a separate—and in some respects most archaic—variety of Chuvash in later studies (Mudrak 1993; Savelyev 2018).

Over the last thousand years, Chuvash has played a key role in interactions within the Volga-Kama Sprachbund, the term referring to extensive and recurrent contact relations primarily between Finno-Ugric (Mari, Permic, and Mordvinic) and Turkic (Bulgharic and Volga Kipchak) languages of the region (Bereczki 2005; Agyagási 2012). It was an important source of superstratum influence for the Finno-Ugric languages, but was itself severely affected by Volga Tatar. Conversely, evidence of substratum influences can be traced from Finno-Ugric to Chuvash and Tatar, and from Chuvash to Tatar (see also above in this section for contacts of Volga Bulghar and Section 27.6 for borrowed lexicon in Chuvash).

27.3 Phonology

27.3.1 Consonants

The native Chuvash consonant inventory includes 17 phonemes represented below in Table 27.1.

Table 27.1 Consonant phonemes

	Labial	Alveodental	Postalveolar	Palatal	Velar
Plosive	p	t		ť	k
Fricative		s	š	ś	x
Affricate			č		
Nasal	m	n		ń	
Approximant	v			y	
Trill		r			
Lateral		l		Ĺ	

All non-palatal consonants are realized as palatalized (to different extents, depending on the specific consonant) in a front-vowel context, cf.: *par-* [*par-*] ‘to give’ vs. *per-* [*p’ɛr’-*] ‘to throw, launch’.

As can be seen from Table 27.1, there is no opposition between voiced and unvoiced stops, which can be reconstructed to Proto-Turkic (Dybo and Mudrak 2006) and is widespread in the Transeurasian area in general (Robbeets 2017c: 592). However, all non-sonorant consonants have semi-voiced vs. voiceless allophones[2]. The semi-voiced variants occur intervocalically, which refers to the position between two vowels or after a sonorant before a vowel. The voiceless allophones appear in all other positions, i.e. word-initially, before a sonorant, before or after a non-sonorant consonant, or word-finally. There are also voiceless geminates that are frequent in the intervocalic position, particularly as an emphatic element within a non-verbal stem (e.g., in numerals in a substantive use, see Section 27.4.3, and in kinship terms as well as onomatopoeia) or at morpheme boundaries in nominal inflection (see Section 27.3.4). A distribution of voiced vs. voiceless resembling the Chuvash system was historically present in the neighboring Mari language (Kangasmaa-Minn 1998: 222), and it is likely that Chuvash has lost the original voiced/voiceless

distinction under the Mari substratum influence.

Word-initial sonorants are rare in Turkic languages and can hardly be reconstructed for Proto-Turkic. Chuvash develops */m-/* from PTk **b* before an old nasal, e.g., *mə̂ᵒyə̂ᵒr* ‘horn’ < PTk **buyŋuŕ*—a process that is common for all Turkic languages except for most Oghuz varieties (Dybo and Mudrak 2006: 148–149). The initial */v-/* derives from Proto-Turkic long rounded vowels, e.g., *var* ‘inside’ < PTk **ōŕ*. The initial */l-/* is still extremely rare in native Chuvash words, though it occurs in a few inherited stems due to the elision of the preceding vowel, e.g., *lar-* ‘to sit (down)’ < PTk **olur-*. The initial */y-/* comes from the glide element of Proto-Turkic diphthongs, e.g., *yal* ‘village, society’ < PTk **i̯el*. The sonorants */r/*, */n/*, */ń/* and */ĺ/* occur word-initially in borrowed lexemes only.

27.3.2 Vowels

The native Chuvash vowel system is presented in Table 27.2.

Table 27.2 Vowel phonemes

	front		mid-front		mid-back		back	
	unrounded	rounded	unrounded	rounded	unrounded	rounded	unrounded	rounded
close	i	ü					ï	u ụ
mid	e		ə	əᵒ	ə̂	ə̂ᵒ		(o)
open							a	

Chuvash distinguishes full and reduced vowels, the latter being described as the unrounded */ə/* and */ə̂/* for the standard language. However, the Viryal dialect and—as I showed earlier in Savelyev (2013) and Savelyev (2015) in contradiction to the previous descriptions—many Anatri varieties maintain the old opposition between rounded and unrounded reduced vowels, cf. *kər-* ‘to enter’ vs.

kəᵒrü ‘son-in-law’, *pârax-* ‘to throw’ vs. *pâᵒrụ* ‘calf’.

The contrast between full and reduced vowels is a distinctive feature of the Volga-Kama Sprachbund (Róna-Tas 1986; Tumaševa 1988). It applies across genealogical boundaries, being present in Tatar, Bashkir, Mari, and Mordvinic dialects. Although it was assumed that the Chuvash system had been copied from Mari (Agyagási 1998; Johanson 2000a), the direction of influence may well be the opposite (see, e.g., Bereczki 1994; Aikio 2014 against Mari as the source of the contrast in the Volga-Kama region).

The Viryal dialect has */o/*, which corresponds to */u/* in the Anatri dialect and the standard language. Both Viryal and Anatri vowel systems include the specific tense vowel */ụ/*, which is rarely mentioned in the literature but clearly contrasts with */o/* and */u/* respectively, cf. Viryal, Anatri *ụt* ‘riding horse’ vs. Viryal *otâ*, Anatri *utâ* ‘hay’.

The sources of */o/ ~ /u/* correspondence in the modern Chuvash are VBulg. **o* (< PTk **o*, **ō*) and VBulg. **å* (< PTk **a*, *ā*). The latter change illustrates another instance of converged evolution in the Volga-Kama region as the labialization of the ancestral **a* is a common feature in the local languages, cf.: PTk **a* > Kazan Tatar *å*, PUr **a* > Proto-Mari *å* (East *o*, West *a*) (Aikio 2014: 140). It is likely that this change has been induced by late Volga Bulghar/early Chuvash in both Mari and—due to substratum interference—Tatar (Čečenov 2002: 225; Dybo and Mudrak 2006: 224).

Chuvash exhibits front vs. back vowel harmony. In many cases, due to the drastic phonological changes, it has back-vowel stems where Proto-Turkic has front-vowel harmony, and vice versa, cf.: Chu. *kur-* ‘to see’ < PTk **gör-*, Chu. *pürne* ‘finger’ < PTk **bi̯arŋak*. Vowel harmony applies within stems and across morpheme boundaries. Most inflectional and derivational suffixes have two allomorphs in accordance with the stem, with */a/ – /e/*, */â/ – /ə/* and */ụ/ – /ü/* being in contrast (hereinafter represented by the capital *A*, *Ə* and *U* respectively). The high vowel */i/* is neutral to vowel harmony in the sense that it can occur after both a front and a back vowel in the previous syllable, cf. *pəᵒtt-i* [pregnant-NMLZ] ‘the pregnant one’ vs. *pâᵒtt-i* [porridge-POSS.3] ‘his porridge’.

Violation of palatal harmony in a back-vowel environment, caused by an adjacent palatal

consonant, is frequent in the Anatri dialect and in the standard language, while Viryal often retains the original phonetic structure, cf. Viryal *poťak* vs. Anatri, standard Chu. *putek* ‘lamb’.

There are some restrictions applying to the distribution of vowels. The high */ï/* appears in the first syllable only; however, in many dialectal varieties, it can occur in suffixes as the back-vowel counterpart of */i/*: Viryal (the speech of Xĕrlĕ Čutaj) *varr-ï* [middle-POSS.3] ‘(its) middle’. Viryal */o/* ~ Anatri */u/* as used in native words also appear in the first syllable only, though this restriction does not apply to */ụ/*. There is no phonological opposition between unrounded and rounded reduced vowels in non-initial syllables; reduced vowels may become labialized after a rounded vowel in the preceding syllable or near a labial consonant.

27.3.3 Syllable and word structure

The basic syllable structure is (C)V(C)(C). As is common in Transeurasian languages, Chuvash does not favour consonant clusters syllable-initially. However, they can appear in spontaneous speech due to the elision of the first-syllable reduced vowel, e.g., */âśta/* = [*əɕta ~ ɕta*] ‘where’, /*pərre*/ = [*p’ër’(:)ɛ ~ pr’ɛ*] ‘one’. Sonorant-obstruent sequences are the only consonant cluster type that is allowed syllable-finally; they occur rather frequently in verbal stems (typically at the boundary of a primary verbal root and a non-productive suffix, e.g., *vïr-t-* ‘to lie (down)’, cf. *vïr-ân* ‘place’) and quite rarely in native nouns (e.g., *śurt* ‘house’, *kâ^{o}vayt* ‘firewood; campfire’).

By far the majority of verbal stems in Chuvash are monosyllabic, while synchronically underived nouns tend to be disyllabic. Reduced vowels never occur at the end of verbal stems. At the same time, around a quarter of native non-verbal stems have the structure (C)VC(C)Ə, where the final reduced vowel may disappear due to prosodic or syntactic factors.

27.3.4 Morphophonology

Extensive morphophonological changes apply to stem-final vowels occurring before vowel-initial suffixes. In this position, */a/* and */e/* disappear: *kala-* ‘to speak’ > *kal-in* [speak-CONC]. Reduced

vowels behave the same way, and the preceding consonant undergoes lengthening: *sə̂ᵒnə̂ᵒ* ‘spear’ > *sə̂ᵒnn-i* [spear-POSS.3]. High /*i*/, /*ụ*/, and /*ü*/, which go back to Proto-Turkic combinations of a vowel and an approximant, prevent a hiatus in this environment by recovering their original shape, being represented as /*iy*/, /*â*ᵒ*v*/, and /*ə*ᵒ*v*/ respectively, cf. *śerśi* ‘sparrow’ > *śerśiy-əm* [sparrow-POSS.1SG], *sụ-* ‘to milk’ > *sə̂ᵒv-at-ə̂p* [milk-PRS-1SG], *pü* ‘body’ > *pəᵒv-ə* [body-POSS.3].

The stem-final or suffixal /*t*/ changes to /*č*/ in a number of cases when it is followed by an inflectional /*ə*/: *śunat* ‘wing’ > *śunač-ə* [wing-POSS.3]. As the /*t*/ > /*č*/ palatalization is a historical process that has only taken place in a certain phonological context (see Section 27.2), /*t*/ remains intact in many other forms, e.g., *śit-ə-p* [achieve-FUT-1SG].

27.3.5 Suprasegmentals

The accent system represented in standard Chuvash is close to that of the Viryal dialect. As a general rule, the last syllable with a full vowel is stressed: *aván* ‘good’, *sə̂ᵒrá* ‘beer’, *karlánkə̂* ‘throat’. In the words with reduced vowels only, the accent falls on the first syllable: *xárlə* ‘red’, *və́ᵒrə̂ᵒm* ‘long’. Most of Anatri varieties represent a more innovative system where the last syllable is usually stressed, regardless of the quality of the vowel: *karlankə́* ‘throat’, *xərlə́* ‘red’.

The accent system based on the opposition between etymologically low and high vowels, with the latter being often reduced as in Chuvash, is also found in the Uralic languages of the Volga-Kama region, including East Mari as well as Mordvinic and Permic dialects. While the ultimate origin of the system in the Volga-Kama Sprachbund is under discussion, it is likely that at least some of the local Uralic languages adopted it from Chuvash (Bereczki 1988).

27.4 Morphology

As is common in the Transeurasian area, Chuvash is a strongly agglutinative, suffixing language. Prefixal inflection is restricted to a few marginal cases only, such as the partial emphatic reduplication of adjectives and adverbs (see Section 27.4.4).

27.4.1 Nouns

Chuvash nouns are inflected for number, case, and possessive suffixes. The general order of suffixes in nominal forms is possessive—number—case: *tus-ə-sem-pe* [friend-POSS.3-PL-INS] ‘with his friends’. This fact contrasts Chuvash with all the other Turkic languages where the order of possessive and number markers is opposite.

27.4.1.1 Number

The number marking system is singular-plural. The singular is unmarked, and the plural is expressed with the suffix *-sem* / *-sen*. The former allomorph occurs in the nominative and in the innovative oblique cases, while the latter appears in the archaic oblique cases (see Section 27.4.1.2). The plural marker is not subject to vowel harmony, which reflects its late formation from an enclitic (< PTk **sayïn* ‘all’). The back-vowel variant has evolved secondarily in the Viryal dialect: *xola-sam* [city-PL]. The plural marker is also used for marking associative plurality: *Timər-sem* [‘male.name’-PL] ‘Timĕr and his relatives/friends’. There is no trace of the Common Turkic plural marker *-lAr* in Chuvash.

27.4.1.2 Case

According to the local grammar tradition (e.g., N. Andrejev 1961), the Chuvash case system includes eight cases: nominative, genitive, dative-accusative, locative, ablative, instrumental-comitative, caritive, and causal-final. The first five are archaic cases that are continuations of Proto-Turkic cases. The new cases—instrumental-comitative, caritive, and causal-final—are formed on the basis of Proto-Turkic postpositions that functioned as enclitics in Chuvash until as late as the 18th century (see, e.g., Savelyev 2014: 254–255) or even later.

The “traditional” Chuvash cases are presented and exemplified in Table 27.3.

Table 27.3 Case markers

Function	Marker	-C	-R	-V	-Ə
NOM	*-ø*	*yat* ‘name’	*xir* ‘field’	*ura* ‘foot, leg’	*śilə* ‘udder’
GEN	*(-V)-n, (-C)-Ən*	*yat-ân*	*xir-ən*	*ura-n*	*śill-ən*
DAT/ACC	*(-V)-nA, (-C)-A*	*yat-a*	*xir-e*	*ura-na*	*śill-e*
LOC	*(-V)-rA, (-R)-tA*, otherwise *(-C)-rA*	*yat-ra*	*xir-te*	*ura-ra*	*śilə-re*
ABL	*(-V)-rAn, (-R)-tAn*, otherwise *(-C)-rAn*	*yat-ran*	*xir-ten*	*ura-ran*	*śilə-ren*
INS/COM	*-pA*	*yat-pa*	*xir-pe*	*ura-pa*	*śilə-pe*
CAR	*-sƏr*	*yat-sâr*	*xir-sər*	*ura-sâr*	*śilə-sər*
CAUS/FIN	*-šƏn*	*yat-šân*	*xir-šən*	*ura-šân*	*śilə-šən*

The choice between the case allomorphs is determined by the final sound of the stem. In the table, the capital *C* represents any stem-final consonants except for *R*, which stands for one of the three dental consonants *r l n*. For diachronic reasons, *R*-final stems take specific locative and ablative allomorphs. The capital *V* represents stem-final consonants except for *Ə*, which stands for reduced vowels. Vowel-final stems as inflected with case markers are subject to the morphophonological rules described in Section 27.3.4. Besides, *Ə*-final stems as marked in the genitive or the dative-accusative behave morphophonologically like consonant-final and join a vowel-initial suffix. The consonant *-n-* in the dative-accusative and the locative is epenthetic; some Viryal varieties use *-y-* instead. The instrumental-comitative marker has the free variant form *-pAlA*. Yet another variant *-pAlAn*, possibly formed under the influence of Tatar *bilän* ‘together’, occurs in the Anatri dialect only.

In addition to the presented case markers, there are a number of “quasi-case” markers in the

sense that they are not usually included in the case system because of their compatibility with stems of certain semantic classes only or because their status as "true" case markers is otherwise debated. The approximative suffix *-(n)AllA* is formed on the basis of the dative-accusative marker *-(n)A* and the adverbial marker *-llA* (see Section 27.4.4): *vâºrman-alla* [forest-APPROX] 'towards the forest'. The terminative case is expressed by the suffix *-čče*n: *śičča-čče*n [seven-TERM] 'up to seven o'clock'. The egressive suffix *-RAnpa* comes from a combination of the ablative *-RAn* and the instrumental-comitative *-pA*: *śulla-RAnpa* [summer-EGRE] 'since summer'. The proprietive meaning can be expressed with the suffix *-(l)lə* (see also Section 27.4.4): *viśə laša-llâ śïn* [three horse-PROPR person] 'the person having three horses'.

There are also a number of fossilized case markers. The suffix *-Ən* is the marker of the old instrumental case (< PTk **-In*) that is no more productive when attached to nouns, but it can still be recognized in some adverbial forms: *ukśa-n* [money-INS] 'for money'. The non-productive vocative markers, which are frequent in kinship terms and personal names, include the suffixes *-ụk*, *-ay* (~ *-ey* ~ *-i*), and several more dialectal variants: *Naśt-ụk* ['female.name'-VOC] 'Nastya (Anastacia)', *kukam-ay* [mother's.mother-VOC] '(my) mother's mother'.

27.4.1.3 Possessive suffixes

The system of Chuvash possessive suffixes is presented in Table 27.4.

Table 27.4 Possessive suffixes

	SG		PL	
	DIR	OBL	DIR	OBL
1	*-Əm*		*-ƏmƏr*	
2	*-U*	*-Un-*	*-Ər*	
3	*(-C)-ə*	*(-C)-(ə)n-*	*(-C)-ə*	*(-C)-(ə)n-*

	(-V)-i	*(-V)-in-*	*(-V)-i*	*(-V)-in-*

Sources from as late as the end of the 19th century witness distribution of postconsonantal *-Ə* and postvocalic *-U* in the second-person singular marker (Ašmarin 1898: 133), but the former is not usually mentioned in synchronic descriptions from the 20th century (e.g., N. Andrejev 1961: 608). The third-person singular and plural markers are expressed with the same suffixes, both *-ə* and *-i* going back to the Proto-Turkic third-person possessive marker **-I* and, further, to the Proto-Altaic third-person personal pronoun (Schwarz et al., this volume: Chapter 32). The Proto-Turkic **-sI*, which is the postvocalic variant of **-I*, developed semantic specialization in Chuvash where *-(Ə)šə* is used with kinship terms and names of youngs of domestic animals: *kukk-âšə* [uncle-POSS.3] '(his) uncle', *pâ^{o}rụ-šə* [calf-POSS.3] '(his) calf'. Words for 'mother' and 'father', as well as kinship terms based on these two, have suppletive possessive forms: pre-modern Chu. **at(t)a* + the historical vocative suffix *-ay* > **at(t)ay* > Chu. *atte* '(my) father', Chu. *aśa* 'male', dial. 'father' > Chu. *aś-ụ* [father-POSS.2SG] '(your) father', pre-modern Chu. **aś-âšə* [father-POSS.3] > Chu. *aššə* '(his) father'; pre-modern Chu. **an(n)a* + the vocative suffix *-ay* > **an(n)ay* > Chu. *anne* '(my) mother', *ann-ü* [mother-POSS.2SG] '(your) mother', Chu. *ama* 'female', dial. 'mother' > *am-âšə* [mother-POSS.3] '(his) mother'.

27.4.2 Pronouns

The Chuvash personal pronouns are presented in Table 27.5.

Table 27.5 Personal pronouns

Person	SG		PL	
	DIR	OBL	DIR	OBL

1	*e-pə*	*man-*	*e-pir*	*pir-*
2	*e-sə*	*san-*	*e-sir*	*sir-*
3	*vâ^{o}l*	*â^{o}n-* [DAT/ACC], otherwise *un-*	*və^{o}sem* (< *vâ^{o}l-sem* [3-PL])	*və^{o}sen-*

Etymologically, Chuvash personal pronouns have a particularly archaic shape as compared to Common Turkic languages and are most probably inherited from Proto-Altaic or even Proto-Transeurasian (for this and other Transeurasian connections for Chuvash pronouns, see Schwarz et al., this volume: Chapter 32). There is no opposition between exclusive and inclusive first-person pronouns. The first- and second-person pronouns are reinforced by the deictic particle *e-* of likely Transeurasian origin, although Adamović (1984: 6) pointed to its possible connections with the deictic *e-* in some languages of Eastern Europe, cf. Ru. *ė-tot* 'this', Mord. *e-te* 'this'. The direct first-person plural and second-person plural stems occur without the *e*-element in the Northwestern dialect: *pẹr* 'we', *sẹr* 'you (pl.)'. The Chuvash second-person plural pronoun can be used for singular honorific reference, which can be a recent innovation driven by the contact with Russian.

Personal pronouns inflected in the genitive function as possessive pronouns. They are commonly used to reinforce the third-person possessive suffix but rarely occur together with first- and second-person possessive suffixes: *un-ân all-i* [3SG-GEN hand-POSS.3] 'his hand', (? *man-ân*) *all-âm* [1SG-GEN hand-POSS.1SG] 'my hand'.

The system of demonstrative pronouns as used for deictic reference consists of four elements expressing three degrees of distance, with proximal *ku* and *śakâ*, medial *śavâ* and distal *leš.* The opposition between *ku* and *śakâ* seems to reflect direct / emphatic vs. non-direct / non-emphatic deixis respectively (Ašmarin 1898; *pace* Sergejev 1994, interpreting the relations between the proximal demonstratives in the opposite way). All the deictic demonstratives can be used anaphorically, along with the third-person singular personal pronoun *vâ^{o}l* and the special anaphoric pronoun *xayxi*.

The reflexive pronouns presented in Table 27.6 are derived from the bound stem *xâ-* (*âγ-* in the

Northwestern dialect) corresponding to the Common Turkic emphatic particle **oq* ~ **uq* (Ramstedt 1922: 16; Räsänen 1957: 39; Mudrak 2009: 75).

Table 27.6 Reflexive pronouns

	SG	PL
1	*xam*	*xam-ârr*
2	*xu̯*	*xâ^{o}v-ârr*
3	*xây*	*xây-sem*

The basic interrogative pronoun stems are *kam* ‘who?’, *mən* ‘what?’, and the bound stem (PTk **qay-*) underlying the forms *xâśan* ‘when?’, *xâšə* ‘which?’, and *âśta* ‘where?’. The quantitative interrogative pronoun *miśe* is derived from *mən* ‘what?’ (cf. more archaic Northwestern and Viryal forms *mənče*, *minče*).

The interrogative pronouns serve as a basis for negative and indefinite pronouns. Negative pronouns are derived with the preposing negative marker *ni* (*ni-kam* ‘nobody’, *ni-mən* ‘nothing’), which is borrowed from Classical Persian negative prefix *ni-* or from Ru. *ni-* with the same meaning. In indefinite pronouns, the interrogative pronoun part is preceded by the indefiniteness marker *tA-* (*ta-kam* ‘somebody’, *te-mən* ‘something’) originating from the disjunctive conjunction *te* (cf. *te… te…* ‘whether… or...’).

27.4.3 Numerals

The Chuvash numeral system is decimal. The numerals from ‘one’ to ‘ten’, decimals from ‘twenty’ to ‘fifty’ and names for ‘hundred’ and ‘thousand’ are synchronically non-derived (see Table 27.7). All these numerals go back to Proto-Turkic, except for the numeral for ‘thousand’, which is a Common Turkic borrowing in Chuvash. The numerals for ‘sixty’ and ‘seventy’ are old derivatives

of ‘six’ and ‘seven’ formed with the suffix *-məl* (PTk **-mïl*): Chu. *ut-mâl* ‘sixty’ (cf. *ultâ* ‘six’), *śitməl* ‘seventy’ (cf. *śičə* ‘seven’). The numerals for ‘eighty’ and ‘ninety’ are multiplicative: *sakâr-vun(n)â* [eight-ten] ‘80’, *tâxâr-vun(n)â* [nine-ten] ‘90’. The other numerals are additive: *pin=te ik-śər vun-viśśə* [thousand=COORD two-hundred ten-three] ‘1213’.

The cardinal numerals from ‘one’ to ‘ten’ as well as ‘fifty’ appear in two variants depending on their syntactic function (Table 27.7). The full, or emphatic, form has a lengthened medial consonant and appears in substantive use: *tâᵒvattâ-ra* [four-LOC] ‘at four (o'clock)’. The short form has a single consonant instead and is used attributively: *tâᵒvatâ yulanut* [four horseman] ‘four horsemen’. Short variants can lose their final reduced vowel in accordance to the general tendency (see Section 27.3.3), termed as “clipped forms” by Clark (1998a: 442).

Table 27.7 Non-derived numeral lexemes

	short	full		short	full
1	*pər*	*pərre*	9	*tâᵒxâᵒr*	*tâᵒxxâᵒr*
2	*ikə, ik*	*ikkə*	10	*vunâ, vun*	*vunnâ*
3	*viśə, viś*	*viśśə*	20	*śirəm*	
4	*tâᵒvatâ, tâᵒvat*	*tâᵒvattâ*	30	*vâᵒtâᵒr*	
5	*pilək*	*pillək*	40	*xərəx*	
6	*ultâ, ult*	*ulttâ*	50	*alâ, al*	*allâ*
7	*śičə, śič*	*śiččə*	100	*śəᵒr*	
8	*sakâr*	*sakkâr*	1000	*pin*	

The formal difference between attributive and substantive cardinal numerals is unusual for the Turkic languages, although Levitskaja (1976: 45-47) refers to a similar distinction, involving the

long vs. short medial consonant opposition, in the West Siberian Tatar varieties. In the Volga-Kama region, a system distinguishing between the two syntactic functions can also be found in Mari. However, the contrast is achieved in its case by suffixing rather than consonant lengthening, cf. Western Mari *kok* ‘two’ (attributive), *kok-tï* [two-SUBST] ‘two’ (substantive).

Ordinal numerals are derived from the full form of cardinal numerals with the suffix *-məš*: *pilləkməš* [five-ORD] ‘fifth’.

Collective numerals can be divided into two types. General collective numerals are formed from the full form of cardinal numbers with the suffix *-Ən*, which is identical to the old instrumental case marker, and function as adverbial depictives: *ultt-ân kil-* [six-INS come] ‘to come all six’. Restrictive, or elective, collective numerals are derived with the combination of the archaic collective marker *-s-* (*-š-* for the third person) and a plural possessive suffix. The first- and second-person forms of restrictive collective numerals are based on the clipped forms of cardinal numerals: *ik-s-əmər* [two-COLL-1] ‘the two of us’, *tâ^{o}vat-s-âr* [four-COLL-2] ‘the four of you’. The third-person restrictive collective numerals are derived from the full forms of cardinal numerals: *śiččə-š-ə* [seven-COLL-3] ‘the seven of them’.

Distributive numerals are derived from the clipped forms of cardinal numerals with the suffix *-šAr* < PTk **łar*: *vun-šar* [ten-DISTR] ‘in twos’.

There are no obligatory classifiers in Chuvash. One of the most common counting words is *pə^{o}rčə* ‘grain, granule, little part’ as used in the third-person possessive form *pə^{o}rč-i* [grain-POSS.3]. It is associated with things consisting of numerous tiny items, such as hay, beeds, peas, etc.: *pər śüś pə^{o}rč-i* [one hair grain-POSS.3] ‘one hair’ (**pər śüś* would be grammatically incorrect).

27.4.4 Property words

Property words in Chuvash are adjectives and adverbs. As in other Turkic languages, many of the descriptive adjectives in Chuvash can function as nouns: *nayan* ‘lazy; a lazybone’, *ïrâ* ‘good; a good thing, a good spirit, etc.’. Other adjectives require a nominalizer to be used in a nominal function:

šurâ ‘white’ > *šurr-i* [white-NMLZ] ‘the white one’. Frequent denominal adjectival suffixes are *-(l)lƏ* with a proprietive meaning and *-xi* with a temporal meaning: *vây* ‘strength’ > *vây-lâ* ‘strong’, *śəᵒr* ‘night’ > *śəᵒr-xi* ‘night (adj.)’.

The comparative degree is formed with the suffix *-rAx*; the variant *-tArAx* is obligatory after *r* and allowable after *l n*: *śülə-rex* [high-COMP] ‘higher’, *yïvâr-tarax* [heavy-COMP] ‘heavier’, *avan-rax ~ avan-tarax* [good-COMP] ‘better’. The superlative degree is formed analytically, by adding the word *či* ‘most’: *načar* ‘bad’ > *či načar* ‘the worst’.

The intensive forms of adjectives are derived with partial reduplication, a phenomenon that is extremely widespread in the Turkic family and in the Transeurasian languages in general. The most common model is reduplication of the word-initial CV-sequence with an additional *p*, e.g. *xup~xura* [INTS~black] ‘very black, pitch-black’, *təᵒp~təᵒttəᵒm* [INTS~dark] ‘very dark, pitch-dark’. There are, though, numerous intensive forms that deviate from the regular model, e.g. *ś**a**p~śutâ* [INTS~light] ‘very light’, *kâᵒ**n**~kâvak* [INTS~blue] ‘very blue’, *yə**m**~yešəl* [INTS~green] ‘very green’. Certain adjectives retain a different pattern that does not apply reduplication and relies on specific (sometimes unique) intensive markers instead: *vəᵒr-śənə* [INTS-new] ‘brand new’.

A large number of adverbs are formally identical to adjectives, e.g., *avan* ‘good; well’. The two most frequent adverbial suffixes are *-(l)lA* and *-Ən*. The suffix *-(l)lA* derives mainly domain adverbs from nouns and adjectives, e.g., *mečək-le vïĺa-* [ball-ADV play] ‘to play ball’, *śənə-lle purân-* [new-ADV live] ‘to live in a new way’. The suffix *-Ən*, the old instrumental case marker (see Section 27.4.1.2), was interpreted by Ašmarin (1903: 143) as referring to a non-permanent property or pointing to subjective perception, cf. (1) and (2).

(1) *kun* *sulxân* *tâᵒr-aťˊ*
day cool stand-PRS.3SG
‘The weather is cool.’

(2) *kun* *sulxân-ân* *tâ^{o}r-aʹ*

day cool-ADV stand-PRS.3SG

‘The weather is [perceived by the speaker as] cool.’ (Ašmarin 1903: 144).

27.4.5 Verbs

Chuvash verbs are inflected for mood, tense, person and number, negation and potential suffixes. The basic order of suffixes in a verbal form is potential form—negation—mood (tense)—person and number: *kal-ay-ma-r-âm* [speak-POT-NEG-PST-1SG] ‘I was not able to say’. The bare verbal stem is identical to the second-person singular imperative form. Verbal inflection exhibits almost no irregularities; a rare exception are ten monosyllabic verbs that lose their final *-r-* in a number of forms, cf. *par-* ‘to give’ > *pa-nə* [give-PST.PTCP], but *čar-* ‘to stop’ > *čar-nâ* [stop-PST.PTCP].

27.4.5.1 Person and number

As well as other Turkic languages, Chuvash has two main paradigms of verbal person-number markers. As summarized by Dybo (2017: 128), paradigm I of the Turkic person-number markers coincides to a large extent with the possessive suffixes (see Section 27.4.1.3) and occurs in the verbal forms of non-participial origin. Paradigm II is based on personal pronouns (see Section 27.4.2) used as postpositional clitics and appears in the verbal forms that, ultimately, go back to participles. In Chuvash, paradigm I occurs in the past tenses as well as in the conditional mood, while paradigm II appears in the future and—with some further development—the present tense as well as the concessive and the optative moods. The two paradigms of verbal person-number markers in standard Chuvash are presented in Table 27.8.

Table 27.8 Verbal person-number markers

Paradigm	I	II

	SG	PL	SG	PL
1	*-Əm*	*-ƏmƏr*	*-Əp*	*-(Ə)pƏr*
2	*-Ən*	*-Ər*	*-Ən*	*-Ər*
3	*-ʼə*	*-ʼəś*	*-ə*	*-əś*

As well as Old Turkic and most Siberian Turkic languages, Chuvash copied the first-person plural verbal marker from the possessive paradigm, while most of the other contemporary Turkic languages make use of the specific first-person plural marker *-*UK* (Mudrak 2009: 29–30; Dybo 2017: 128–129). The Chuvash development seems to represent an innovation against the background of such branches as Kipchak and Oghuz Turkic. The Chuvash second-person markers as represented in paradigm II are copied from paradigm I. Mudrak (2012: 710–712) reconstructs the original paradigm II second-person markers -(*Ə*)*s* (singular) and -(*Ə*)*sƏr* (plural) based on the archaic concessive forms in the Northwestern dialect. He further attributes the partial merger of the two paradigms in Chuvash to an early influence of Kipchak Turkic languages, which underwent a similar process. The difference between the third-person markers in the two paradigms is that *-ə* in paradigm I causes palatalization of the preceding **t*, while its counterpart in paradigm II does not, cf. *kur-t-âm* [see-PST-1SG], *kur-č-ə* [see-PST-3SG] vs. *sut-ə* [sell-FUT.3SG].

For the specific paradigm of person-number markers in the imperative, see (ii) in the next section.

27.4.5.2 Tense and mood

(i) Indicative

The indicative is the only mood to have tenses, including the present, the future and several past tenses. All tenses are expressed with finite verbs and take person-number agreement, except for the indefinite past and the indefinite pluperfect.

The present tense is marked by the suffix *-At-* (with the cumulative *-aẗ* / *-et* for the third-person

singular and -*Aśśə* for the third-person plural) and expresses both habitual (3, 4a) and durative meanings (4b). It is sometimes referred to as the non-past tense as the same markers are used to express scheduled future events (4c).

(3) *śər* *xə̑y-ən* *tənəl-ə* *tavra* *śavrə̑n-aʹ*
Earth REFL.3SG-GEN axis-POSS.3 around rotate-PRS.3SG
'The Earth rotates on its axis.'

(4) *epir* *ĭyt-at-pə̑r*
1PL ask-PRS-1PL
a. 'We ask.'
b. 'We are asking.'
c. 'We will [definitely] ask.'

The future tense expresses a hypothetical future situation (5). Benzing (1959: 740, 747) and Levitskaja (1976: 63–65) have pointed out correctly that the Chuvash future tense marker is -*Ə*-. That contradicts a widespread view (e.g., Jegorov 1957: 190) that Chuvash has no overt future tense marker (in this interpretation, the reduced vowel is regarded as a part of the person-number suffix).

(5) *epir* *ĭyt-ə̑-pə̑r*
1PL ask-FUT-1PL
'We will [probably] ask.'

The simple past tense is also referred to as the "categorical" past tense in the Chuvash grammar tradition (e.g., N. Andrejev 1961: 618). It is typically used when the speaker knows for sure that something has taken place, because he witnessed the event or for other reasons. The simple past

tense markers are *-t-* (*-č-* for the third person) after *-R-*, otherwise *-r-* (for *-R-*, see the paragraph under Table 27.3) (6). The Northwestern dialect also uses the converb marker *-sa* to express the simple past tense. The simple past in Chuvash contrasts with the so-called "indefinite" past tense, which is used to express the typologically frequent combination of the perfect (resultative) and indirect (reported) past meanings (7, 8). The "indefinite" past (perfect–indirect past) tense is formed by using the past participle marker *-nƏ* in a finite position.

(6) *vâ^{o}l* *kay-r-ə*
3SG go-PST-3SG
'He went (away).'

(7) *vâ^{o}l* *kay-nâ*
3SG go-PST.PTCP
'He is [apparently] gone.'

(8) *veś-ex* *tụ-nâ*
everything-EMPH do-PST.PTCP
'[Someone] has [apparently] done everything.' = 'Everything is (has been) done.'

The imperfect tense is marked by the suffix *-Att-* (*-Atč-* for the third person): *purân-att-âmâr* [live-IPRF-1PL] 'we used to live'. The imperfect tense marker is historically analyzable as a combination of *-At-*, which is identical to the present tense marker of ultimate participial origin, and the defective existential verb PTk **e(r)-* inflected with the past tense marker **-t-* (Levitskaja 1976: 65).

There are several ways to express pluperfect meanings. The pluperfect tense marked by *-sAtt-* (*-sAtč-* for the third person) is referred to as the "long-past categorical" tense by the Chuvash grammarians (e.g., N. Andrejev 1961: 619) and is used in various pluperfect functions for events

that were witnessed by the speaker (9). The marker *-sAtt-* is analyzable as a combination of the converb marker *-sa* and the existential verb PTk **e(r)-* in the past tense (Levitskaja 1976: 72). In addition, different past-tense forms can be inflected with the clitic *-(č)čə*, which goes back to PTk **e(r)-t-i* [be-PST-3SG]) and functions as the so-called retrospective shift marker (Plungian and van der Auwera 2006: 344). The combination of the past participle marker *-nƏ* with *-(č)čə* is usually interpreted as expressing another pluperfect tense that indicates reported pluperfect situations and the resultative (perfect) in the past (10). The retrospective shift marker can reinforce the pluperfect forms in *-sAtt-*, while some dialects also allow the use of *-(č)čə* after the simple past and imperfect tense forms. Furthermore, the retrospective shift marker is compatible with non-verbal predicates, such as nouns, adjectives, and the existential-possessive predicates *pur* 'there is' and *śuk* 'there is no' (11). In the modern Chuvash, the synthetic pluperfect forms are gradually replaced by the analytic construction of the main verb and the auxiliary *pul-* 'to be', both inflected with the past participle marker *-nƏ* (12).

(9) *âśta* *kay-ma* *tux-satt-ân*
where go-INF come.out-PLUPRF-2SG
'Where were you going?' (lit. 'Where had you come out to go?', cancelled prospective)

(10) *pəltər* *sad-a* *mulkač* *pï-nâ-ččə*
last.year garden-DAT/ACC hare come-PST.PTCP-RETR
'Last year, a hare [apparently] came (lit. 'had come') to the garden.' (reported remote past/resultative in the past)

(11) *ača* *čux-n-e* *man* *ünerśə* *pul-as* *ᵊməᵊt* *pur-ččə*
child time-POSS.3-DAT/ACC 1SG.GEN artist be-FUT.PTCP dream EXIST-RETR
'As a child, I had a dream to become an artist.'

(12) *pəčək Atner škul-a vəᵒren-me ulttâ-r-ax kay-nâ*

little ‘male.name’ school-DAT/ACC study-INF six-LOC-EMPH go-PST.PTCP

pul-nâ

be-PST.PTCP

‘Little Atner started school when he was as young as six years old.’

(ii) Imperative

There are no tenses in the imperative mood. The imperative (functionally, the hortative in the first person and the jussive in the third person) forms are inflected with a special set of person-number markers (Table 27.9) that are not identical to the corresponding markers in the indicative and other moods.

Table 27.9 Imperative person-number suffixes

	SG	PL
1	*-Am*	*-Ar*
2	*-Ø*	*-Ər*
3	*-tƏr*	*-ččƏr*

The imperative forms can be reinforced by various clitic markers. The retrospective shift marker *-(č)čə* as attached to forms in the imperative functions as a politeness marker, and the emphatic imperative markers include *-xa* (a discourse particle with a large number of functions), *-xalə* (< *xalə* ‘now’ of Arabic origin), and -s*Am* (linked to the second-person singular conditional form PTk *-s*An* in Levitskaja 1976: 77) (13).

(13) *kala-sam-ччə ača-m*

tell-EMPH-POL child-POSS.1SG

'Please tell me, my child!'

The prohibitive paradigm makes use of free negative markers derived from existential verbs. The first-person prohibitive forms are built on imperative forms followed by *mar*, a particle that is historically the negative aorist participle form of the defective existential verb PTk **e(r)-* (Levitskaja 1976: 64; Mudrak 2002a: 710). For the second and third person, Chuvash uses the prepositional particle *an*. Its ancestral form, VBulg. **en* (~ **eŋ*), has been traditionally interpreted as a loan from Permic **e-n* [NEG.EXIST-2], cf. Udm., KZ *en* (Wichmann 1903: 148; Jegorov 1964: 26–27). The proposal is based on the fact that the Permic word is of Uralic origin (< Proto-Uralic negative verb **e-*), while parallels for Chu. *an* in other Turkic languages are scarce and somewhat unreliable (? Oghuz *eŋ ~ aŋ* 'no' attested in Kāšġarī's *Dīwān*). Another old argument is that preverbal particles are "an absolutely anti-Turkic feature", as formulated recently by Manzelli (2015: 642–645). However, a borrowing of the Permic prohibitive marker into Chuvash seems to be quite unlikely. First, there are no other grammatical markers in Chuvash reliably borrowed from Permic (on the very limited Permic influence on Chuvash, which is restricted to the lexical domain only, see Section 27.6). Second, despite the lack of reliable cognates in Turkic, Chu. *an* may still be of genuine origin as it has plausible parallels in the other Altaic branches, cf. the negative verbs PTg **e-*, PMg **e-* as reconstructed by Robbeets (2015: 192–202).

(iii) Conditional

The conditional mood is typically used to indicate events that are dependent on irreal conditions as well as counterfactual situations (14). The standard Chuvash conditional mood marker *-Ətt-* (*-əччə-* for the third person) is analyzable as a combination of the original future tense marker *-Ə-* and the reflex of the Proto-Turkic existential verb **e(r)-* inflected with the past tense marker **-t-*

(Levitskaja 1976: 80–81). This grammaticalization pattern has been further recycled by some varieties of the Viryal dialect, which use the retrospective shift marker *-(č)čə* (< PTk **e(r)-t-i* [be-PST-3SG]) as attached to the future tense forms in order to express conditionality: *par-ə̂-p-čə* [give-FUT-1SG-RETR] ‘I would give/have given’.

(14) *pəśer-se* *par-ə̂tt-ə̂m=ta* *śə̂ᵒnə̂ᵒx* *śuk*
cook-CVB give-COND-1SG=EMPH flour NEG.EXIST
‘I would have cooked [pancakes] for [you], but there is no flour.’ (G. V. Zotova, “Śuralnă kuna ujavlani”)

(iv) Concessive and optative

The concessive mood can indicate subjunctive situations in addition to the proper concessive relations. Ašmarin (1898: 308–309) compared the Chuvash concessive mood marker *-in* with the Erzya Mordva *ino* ‘so be it’ and explained both as loans from the Russian dialectal particle *in*, *ino* ‘so be it’. This proposal requires further investigation. In standard Chuvash, the concessive mood marker is attached to the future tense forms only (15), but “present-concessive” and “past-concessive” forms are also attested in early records of the Chuvash dialects: *pïr-at-s-ə̂n* [go-PRS-2SG-CONC] ‘(even) if you go’, *tiv-r-in* [touch-PST.3SG-CONC] ‘(even) if he had touched’.

(15) *xam-ə̂n* *var-ə̂m-ran* *śural-nə̂-sker-e*
REFL.1SG-GEN womb-POSS.1SG-ABL be.born-PST.PTCP-NMLZ-DAT/ACC
par-ə̂p-in *śeś*
give-FUT.1SG-CONC only
‘[Will the Lord] only [be pleased] <…> if I give those born from my own womb?’ (Micah 6: 7)

In the second-person singular and plural, the concessive displays an archaic feature in that it makes use of the original paradigm II second-person markers *-(Ə)s-* and *-(Ə)sƏr-*. The concessive mood marker *-in* is modified as *-Ən* in the second-person singular—most likely, under the influence of the paradigm I second-person marker *-Ən*—and is omitted in the plural (but preserved as *-iń*/*-ẹn* in the Northwestern dialect) (Mudrak 2002a: 711).

The optative mood is based on the retrospective shift marker *-(č)čə* cliticized to the concessive mood form. Forms in the optative are typically used in subordinate clauses to indicate wishful irreal conditions (16).

(16) *ex* *Ilempi* *man* *kin-əm* *pul-in-ččə*
INTJ ‘female.name’ 1SG.GEN daughter.in.law-POSS.1SG be-CONC-RETR
tâ^{o}van *xər-e* *yurat-nâ* *pek* *yurat-ât-âm*
own daughter-DAT/ACC love-PST.PTCP alike love-COND-1SG
‘Oh if only Ilempi had been my daughter-in-law, I would have loved her as I loved my own daughter!’ (Je. G. Selivanova, “Ilempi”)

In the contemporary Chuvash, the synthetic concessive and optative forms are extremely rarely used. Similar functions are performed by analytical constructions with different forms of the verb *pul-* ‘to be’, including *pul-in* [be-CONC] ‘so be it; although’, *pul-san* [be-COND.CVB] ‘if’, *pul-san=ta* [be-COND.PTCP=EMPH] ‘even if’.

27.4.5.3 Negation

Finite verbs make use of the negative marker *-mA-* in all tenses of the indicative and in all other moods except for the imperative, see (ii) in Section 27.4.5.2 for the details. In the present and imperfect tenses, negation is marked by the suffix *-mAs-*, which is historically the negative form of the future participle marker *-as*, cf.: *vĭĺa-ma-r-âm* [play-NEG-PST-1SG] ‘I did not play’, *vĭĺa-mas-t-âp*

[play-NEG-PRS-1SG] ‘I do not speak’. Non-finite verb forms, see Section 27.4.5.6, make use of non-verbal negators, i.e., the negative particle *mar* (< PTk **e(r)-me-r* [be-NEG-AOR.PTCP]), the negative existential-possessive predicate *śuk*, and the caritive case marker *-sƏr*, cf.: *parân-malla mar* [give.up-NEC.PTCP NEG.PTCL] ‘one should not give up’, *kur-as śuk* [see-FUT.PTCP NEG.EXIST] ‘one is not supposed to see’, *vaska-ma-sâr* [hurry-INF-CAR] ‘without hurry’.

27.4.5.4 Potential form

Both affirmative and negative verb forms can be inflected with the specific potential form marker *-Ay-*: *pïr-ay-at-pâr* [come-POT-PRS-1PL] ‘we are able to go’, *kil-ey-me-rə* [come-POT-NEG-PST.3SG] ‘he was not able to come’. The affirmative potential form is rarely used in the modern Chuvash, while the negative form is more frequent. Instead or in combination with the synthetic potential marker, analytical constructions with *pul-tar-* [be-CAUS] ‘to be able, be possible’ can be used to mark physical capability to do something or epistemic possibility: *kâtart-ma pultar-at-âp* [show-INF be.able-PRS-1SG] ‘I am able to show’, *pul-ma pultar-ay-mas-ť* [be-INF be.possible-POT-NEG-PRS.3SG] ‘that cannot happen, that is impossible’.

27.4.5.5 Aspect

Chuvash does not have aspect as a separate grammatical category. Some aspectual concepts are covered by tense marking, e.g., the imperfect and the pluperfect. In rare cases, aspect can be marked by specific suffixes, such as the attenuative-iterative marker *-kAlA*: *vula-* ‘to read’ > *vulakala-* ‘to read a little from time to time’. As well as in other Turkic languages, aspect in Chuvash is most commonly expressed with analytical, in particular, converb constructions, e.g., *üs-se pïr-* [grow-CVB go] ‘to be growing’ (gradative), *kan-sa il-* [rest-CVB take] ‘to take a rest for a while’ (delimitative), etc. For a recent comprehensive study of the functional domain of aspect/aktionsart in Chuvash, see Lebedev (2016).

27.4.5.6 Non-finite verbs

The system of non-finite (nominal) verb forms includes infinitives, participles, and converbs. The infinitive markers are *-mA* < PTk **-maq* and *-mAškƏn* (< pre-modern Chu. **-mAk-šƏn*, the form in PTk **-maq* inflected with the causal-final case marker). The two infinitive markers are mostly interchangeable, but *-mAškən* has a more specific focus on supine functions.

The most common participle forms, all of genuine origin, are the past participle in *-nƏ*, the present participle in *-AkAn* (along with the less productive form in *-An*), the future participle in *-as*, and the necessitive participle in *-mAllA* (the infinitive form in *-mA* inflected with the adverbial *-llA*). The less frequent participle markers include the potential *-i*, the simulative *-Anśi*, and the sufficientive *-mAlƏx*. Participles can be used as both attributes and predicates, cf.: *kay-malla śul* [go-NEC.PTCP road] ‘the road that one has to go along’, *un kay-malla* [3SG.GEN go-NEC.PTCP] ‘he has to go’. All participle forms make use of the nominalizer *-i* to derive verbal nouns that often function as predicates of subordinate clauses (17).

The most common converb in *-sA* expresses simultaneity (17), sequentiality, or manner of action. Other frequent converb markers include the simultaneous *-A* (occurs typically in aspectual converb constructions and, as a fossilized suffix, in postpositions), the temporal/conditional *-sAn* ~ *-sAssƏn*, and the terminative *-iččen* (the potential participle form in *-i* inflected with the terminative case marker).

(17) *Kuĺuk pər šarla-ma-sâr kəneke vula-n-i-ne itle-se*
‘female.name’ one make.noise-INF-CAR book read-PST.PTCP-NMLZ-DAT/ACC listen-CVB
lar-sa čâºlxa śïx-aťˊ
sit-CVB stocking knit-PRS.3SG
‘Without making a sound, Kuĺuk is sitting and knitting stockings and listening to [a friend of her] read a book.’ (M. D. Trubina, “Xĕllexi kaşsenče”)

27.5 Syntax

Most syntactic features of Chuvash are shared with Turkic and Transeurasian languages, which does not exclude some internal and contact-induced innovations. As early as a century ago Ašmarin (1903: IV–V) noticed that some of the syntactic patterns that occur in Chuvash, especially in oral spontaneous speech, were affected by or even copied from Russian. Uralic influence on the Chuvash syntax is also more than probable in view of long-term and tight contact relations between the Uralic and Turkic languages in the Volga-Kama region, but this question requires further investigation. Possible inter-family influence from Tatar, which has significantly affected Chuvash in general, is difficult to detect because of the original close resemblance between the Chuvash and Tatar syntactic patterns.

27.5.1 The nominal group

Chuvash noun phrases are normally head-final, i.e. all types of modifiers precede their head. Modifiers do not agree in number, person, or case with head nouns. The basic order of modifiers within noun phrases is case-marked nominal modifier—demonstrative—relative clause—quantifier—classifying adjective—qualitative adjective.

Chuvash makes use of several kinds of noun-noun combination that are prototypical for Turkic languages. Genitive constructions (N-GEN N-POSS.3) are possessive noun phrases marking the possessor as referential: *laša-n xür-i* [horse-GEN tail-POSS.3] ‘the tail of the horse’. In contrast to that, possessive compounds (N-∅ N-POSS.3) refer to such a relationship of possession where the possessor is non-referential: *laša xür-i* [horse tail-POSS.3] ‘the horse tail’. Possessive compounds can be also used appositively: *Atâl šĕv-ə* [‘river.name’ river-POSS.3] ‘the river Volga’. In a number of cases, nominal combinations can be formed by two nouns in juxtaposition (N-∅ N-∅), the first one functioning as an adjective, e.g., when the dependent noun denotes the material the head noun is made of: *čul śurt* [stone house] ‘stone house’.

In apposition, the proper name typically precedes the single-word appellative but follows the

appellative expressed by a noun phrase: *Almuš patša* [‘male.name’ king] ‘the king Almuš’, *pâᵒlxar patš-i Almuš* [Bulghar king-POSS.3 ‘male.name’] ‘the Bulghar king Almuš’.

As in other Turkic languages, the head noun does not bear plural marking when preceded by a quantifying modifier: *ultâ xəś* [six sword] ‘six swords’, *numay tâᵒšman* [many enemy] ‘many enemies’.

Some noun phrases with a quantifying or pronominal modifier in Chuvash deviate from the standard Turkic patterns. The determiner *pur=te* [EXIST=EMPH] ‘all’ as used together with a personal pronoun takes the final position: *vəᵒsem pur=te* [3PL EXIST=EMPH] ‘they all’. The reflexive pronoun *xâ-* is often pleonastically preceded by a personal pronoun: *vâᵒl xâj* [3SG REFL.3SG] ‘he himself’. Both cases can presumably be attributed to late Russian influence, cf. Ru. *oni vse* [3PL all.PL] ‘they all’, *on sam* [3SG REFL] ‘he himself’.

Forms with possessive suffixes inflected for the first and second person (see Section 27.4.1.3) are infrequent in the contemporary Chuvash (particularly, in the Viryal dialect). Analytical constructions with possessive pronouns are used instead: *san-ân laša* [2SG-GEN horse] ‘your horse’. A similar tendency has been attested in the neighboring Mari language (Galkin 1964: 71). Synchronically, non-third-person possessive suffixes in Chuvash express a sort of inalienable possession relations. In the second person, they are typically attached to nouns denoting body parts or kinship: *all-ụ* [hand-POSS.2SG] ‘your (sg.) hand’, [younger.sister-POSS.2PL] ‘your (pl.) younger sister’. Using the first-person possessive suffixes means, in addition, that the speaker exhibits a deep emotional connection with the possessed entity: *pičče-m* [older.brother-POSS.1SG] ‘my (beloved) older brother’, *śəršïv-âmâr* [land-POSS.1PL] ‘our (beloved) motherland’, *puś-âm* [head-POSS.1SG] ‘my (poor) head’.

As in other Turkic languages, direct objects with a low referential status remain unmarked, i.e. expressed with the default nominative form, while definite and specific indefinite direct objects are marked by the dative-accusative: *kəneke vula-* [book read] ‘to read a book’, *kəneke-ne vula-* [book-DAT/ACC read] ‘to read the book’. Along with the proper accusative and dative functions, the

dative-accusative also marks destination: *Muskav-a an=kuś* [‘city.name’-DAT/ACC NEG=move.IMP.2SG] ‘do not move to Moscow!’. Besides, the dative-accusative occurs in the expressions of purpose and various temporal concepts, and it can also function as *dativus ethicus*. The genitive does not only indicate the possessor in possessive constructions but also marks the subject in some modal expressions (18). The locative and the ablative are used in their proper functions for both location and time. As in many Turkic and Transeurasian languages, the ablative is also used in comparative constructions: *yur-tan šurâ śeśke* [snow-ABL white bloom] ‘the bloom that is whiter than snow’. The instrumental-comitative is not only used for means of the action and accompaniment but also indicates reason, among other semantic roles: *xïrâm-ə pit vïś-n-i-pe* [belly-POSS.3 very be.hungry-PST.PTCP-NMLZ-INS/COM] ‘because he [is] very hungry’. The caritive expresses the absence of something: *yânâš-sâr* [mistake-CAR] ‘without mistakes’. In addition to reason and purpose, the causal-final case can also indicate benefactive relations: *ača-sem-šən uyav* [child-PL-CAUS/FIN festival] ‘children’s festival’ (lit. ‘the festival for children’).

(18) *un-ân* *śiy-es* *kil-et*
3SG-GEN eat-FUT.PTCP come-PRS.3SG
‘He wants to eat.’

Chuvash makes much use of constructions with postpositions, including where other Turkic languages prefer synthetic case marking. The choice between case suffixing or postpositional strategy in Chuvash depends on various factors. For example, postpositions, not case suffixes, are used in personal locational constructions, i.e. when spatial relations refer to an animate referent rather than to a place: Chu. *yal-a kil* [village-DAT/ACC come.IMP.2SG] ‘come to the village!’ vs. *man pat-a kil* [1SG.GEN vicinity-DAT/ACC come.IMP.2SG] ‘come to me!’, cf. Turkish *ba-na gel* [1SG-DAT come.IMP.2SG]. Ašmarin (1903: 20) questioned whether the extensive use of postpositions in Chuvash might have been driven by contact with the Indo-European languages, which tend to use

analytic declension (with prepositions) in order to express case relations.

Within a noun phrase, modifiers are typically coordinated by the clitic *=tA*, which is widely spread across the Transeurasian languages: Chu. *xərlə=te šurə̂ čečək-sem* [red=COORD white flower-PL] ‘red and white flowers’. The basic way to coordinate two noun phrases is via the instrumental-comitative suffix *-pA*: *vatə̂ śïn-pa viləm* [old person-INS/COM death] ‘the oldman and the death’ (the name of a fairytale). Noun phrases can also be coordinated by *=tA*, as well as by the free coordinator *tata* (a reduplication of the former).

27.5.2 The verbal group

Most topics related to the verbal group syntax were covered in Section 27.4.5 and other sections of the chapter; in this section, I discuss specifically verbal valencies. The number of arguments controlled by a predicate may vary from zero to three, and the valency of the verb is changed by voice suffixes. The valency can be reduced by the markers *-l-* and *-n-* that have passive, automative, or reflexive functions (19, 20). Clauses with verbs inflected with these markers may include a reason of action or an instrument (both expressed with the instrumental case), but never an agent (Salo 2013: 233) (21).

(19) *ača* *śura**l**-aťˊ*
child be.born-PRS.3SG
‘The child is being born.’

(20) *kalav* *śə̂ᵒmə̂ᵒl* *vula**n**-aťˊ*
short.story easily be.read-PRS.3SG
‘The short story is read with ease’.

(21) *yə̂və̂ś-sem* *śil-pe* *avkala**n**-aśśə*

tree-PL wind-INS/COM bend-PRS.3PL

‘The trees are bending in the wind.’ (Ašmarin 1928–1950, I: 36)

The highly productive causative marker *-(t)tAr* and the non-productive causative suffixes *-(Ə)t*, *-(Ə)r*, *-Ar, -At* (all distributed according to the historical morphophonological context, see Mudrak 2014) increase the valency of the verb by adding the causer argument, the causee being expressed in the dative-accusative (22).

(22) *śak* *temâ-pa* *man-a* *universsittet-ra* *Iraida.Gennadyevna*

this topic-INS 1SG-DAT/ACC university-LOC ‘female.name’

soччineni *śir**tar**-čə*

composition make.write-PST.3SG

‘Iraida Gennadyevna made me write a composition on this topic at the university.’ (“Čăvaš xalăx sajčĕ”)

27.5.3 The clause

The basic word order is Subject–Object–Verb (23). The Object–Subject–Verb structure occurs when the subject takes up the focus position (24). The word order Object–Verb–Subject is common when the subject is a non-emphatic pronoun used as clitic (25). Other combinations are also possible in spontaneous speech, folklore, and literature.

(23) *epə* *san-a* *yurat-at-ə̂p*

1SG 2SG-DAT/ACC love-PRS-1SG

‘I love you.’

(24) *Yakur-a* *Petər* *śap-r-ə*

Yakur-DAT/ACC Petĕr hit-PST-3SG

‘It is Petĕr who hit Yakur.’ (Ašmarin 1903: 90)

(25) *ə̂ᵒn-a* *pasar-ta* *kur-t-ə̂m* *epə*

3SG-DAT/ACC market-LOC see-PST-1SG 1SG

‘I saw him on the market.’ (Ašmarin 1903: 90)

Existential and possessive predication constructions make use of nominal predicates, the affirmative *pur* and the negative *śuk*. The former can be omitted in many cases. Both *pur* and *śuk* can be inflected with the retrospective shift marker *-(č)čə* to express existential-possessive relations in the past (11). In addition, all non-present tenses make great use of the verb *pul-* ‘to be, exist’ in ‘being’- and ‘having’-constructions. The possessor in predicative possession constructions is normally marked by the genitive case marker (11). Temporary possession is expressed with the locative construction (26). The possessor is cross-referenced on the possessee (the subject of the predicative possession clause) via possessive suffixes or not, depending on a number of factors such as alienable/inalienable distinction, referentiality, and dialectal variation (Ašmarin 1903: 119–121).

(26) *un-ə̂n* *ukś-i* *aššən-če*

3SG-GEN money-POSS.3 father.POSS.3-LOC

‘His money is with his father.’ (Ašmarin 1903: 286)

Untypically for the Turkic languages, Chuvash does not mark nominal predicates by person-number suffixes: *epə čə̂ᵒvaš* [1SG Chuvash] ‘I am Chuvash’, *čüreče uśə̂* [window open] ‘the window is open’.

Clausal negation in Chuvash makes use of both suffixal and analytical negators, see Section 27.4.5.3 for the details. Chuvash does not have special morphological means for marking topic or

focus. As noted above, the topic/focus relations are typically marked by word order. Yes/no and alternative questions are marked with the clitic *=i* attached to the predicate (27). Some Anatri varieties use the interrogative particle *=A* (28) instead. The latter has been compared to Udm. *=a* since Wichmann (1903: 148), but the exact relations between the dialectal Chuvash and the Permic interrogative marker remain unclear. There are several other interrogative markers in Chuvash with more specific functions, including the emphatic interrogative *=ši* and the mirative *=im*.

(27) *śakân-t-i* *śəršïv* *san-a* *kâ^{o}mâ^{o}l-a* *kil-et=i*
this-LOC-NMLZ land 2SG-DAT/ACC mood-DAT/ACC come-PRS.3SG=Q
kil-mes-t=i
come-NEG-PRS.3SG=Q
'Do you like the local area or not?' (Ašmarin 1928–1950, XIII: 100)

(28) *es* *šïv-a* *kə-me* *kay-r-ən=a*
2SG water-DAT/ACC enter-INF go-PST-2SG=Q
'Did you go swimming?' (Ašmarin 1928–1950, I: 3)

As is common in Turkic and other Transurasian languages (see Malchukov and Czerwinski, this volume: Chapter 35), clause subordination in Chuvash makes great use of subordinate clauses based on non-finite predicates. According to the general left-branching tendency, subordinate clauses precede the independent clause. The order of constituents in embedded clauses remains basically intact as compared to the equivalent finite constructions. Nor is there specific strategy for case-marking on the embedded predicate arguments, i.e. they are marked by the same case as they would be in the equivalent independent clause. In this respect, Chuvash deviates from the standard Turkic model with a Transeurasian background, where the embedded subject is commonly marked with an oblique case.

Complement clauses are primarily based on participles, along with the existential nominals *pur* and *śuk*. Examples (29), (30), and (31) illustrate complement clauses taking up the position of the subject, the direct object, and the oblique object. Most types of complement clauses require further nominalization of the non-finite predicate, commonly by means of the suffix *-i*. The nominalizer originates from the third-person possessive suffix *-i*, but it is used regardless of the subject deictic reference. Complement clauses marked with the first- or second-person possessive marker occur only in archaic constructions that are of very limited use in the modern Chuvash (32).

(29) [*pir-ən* *un-pa* *kiləš-mell-i*] *śeś* *yul-č-ə*
1SG-GEN 3SG-INS/COM agree-NEC.PTCP-NMLZ only remain-PST-3SG
'It only remained for us to agree with him.' (I. Andreev 1961: 211)

(30) [*čâᵒvaš-sen* *tâᵒmpay* *sâᵒmax=ta* *purr-i-ne*]
Chuvash-PL.GEN 'tămpaj' word=EMPH EXIST-NMLZ-DAT/ACC
pəl-m-eśśə=ši
know-NEG-PRS.3PL=EMPH.Q
'Don't they really know that there is the word *tămpaj* in Chuvash?' (A. P. Kazanov, "Tămsaj je tămpaj")

(31) [*śemye-ren* *uyrâm* *purân-n-in-če*] *nim*
family-ABL separate live-PST.PTCP-NMLZ-LOC nothing
layâxx-i=te *śuk*
good-NMLZ=EMPH NEG.EXIST
'There is no good in living separately from one's family.' ("Xïpar" newspaper, 03.11.2010)

(32) [*man* *san-a* *temən* *čul* *sâᵒmax* *kal-ass-âm*]

1SG.GEN 2SG-DAT/ACC some that.much word say-FUT.PTCP-POSS.1SG

pur

EXIST

‘I have a lot to tell you.’ (I. Andreev 1961: 166)

Adverbial clauses are based on converbs or nominalized participles in oblique cases (33) and express modifiers of the main clause predicate. Non-finite verb forms with different semantics, see Section 27.4.5.6, take up the position of the embedded predicate and form adverbial clauses of manner, purpose, reason, condition, etc. In addition, Chuvash makes extensive use of analytic constructions based on a participial form followed by a postposition (34).

(33) [*ep* *mən-šən* *kil-n-i-pe*] *čuxlan kâštax*

1SG what-CAUS/FIN come -PST.PTCP-NMLZ-INS/COM think a.bit

‘Think for a bit about why I came here.’ (A. A. Tarasov, “Sutnă pürtri julaški kaś”)

(34) *Vańuška* *kallex* [*niməś-sem* *vâᵒrman* *tâᵒrâᵒx yïtâ-sem-pe* *śüre-me*

‘male.name’ again German-PL forest along dog-PL-INS/COM walk-INF

pultar-ass-i *śinčen*] *šuxâšla-r-ə*

be.able-FUT.PTCP-NMLZ about think-PST-3SG

‘Vańuška thought once again that the Germans might have been walking with dogs around the woods.’ (L. Ja. Agakov, “Ïltăn văčăra”)

Relative clauses are based on non-nominalized participles and take up the prenominal slot as modifiers of the main clause argument: *numay kalaś-akan śïn* [much talk-PRS.PTCP person] ‘the person that talks [too] much’. Among several patterns of non-subject relativization found in Turkic and other Transeurasian languages (see Malchukov and Czerwinski, this volume: Chapter 35),

Chuvash uses that called “attributive clause” in Comrie (1998: 254) and “unmarked relative clause” in Pakendorf (2012: 274). This pattern implies that the relative clause subject in not cross-referenced on the participle, nor on the head noun. As a result, Chuvash subject and direct object relative clauses are structurally identical in that the complex clause [*arslan vəᵒler-nə*] *pəl* [lion kill-PST.PTCP elephant] can be translated both as ‘the elephant that was killed by a lion’ and ‘the elephant who killed a lion’. The same pattern is applied in relativization of oblique objects and adjuncts: *epə kišər pa-nâ śïn* [1SG carrot give-PST.PTCP person] ‘the person whom I gave a carrot’, *pičče-m purân-as śurt* [older.brother-POSS.1SG live-FUT.PTCP house] ‘the house in which my older brother will live’. In using the unmarked strategy of non-subject relativization, Chuvash clusters with the Kipchak languages, including Volga Tatar (Pakendorf 2012: 264), so this may well be an areal phenomenon.

Relativization of possessors of subjects and objects is another available option in Chuvash: *ïvâl-ə əner man-a śap-nâ śïn* [son-POSS.3 yesterday 1SG-DAT/ACC hit-PST.PTCP person] ‘the person whose son hit me yesterday’*, ïvâl-n-e əner epə śap-nâ śïn* [son-POSS.3-DAT/ACC yesterday 1SG hit-PST.PTCP person] ‘the person whose son I hit yesterday’.

Clause linkage based on the so-called “adverbial conjunctions”, see Malchukov and Czerwinski (this volume: Chapter 35), can be illustrated by the construction with the *sa*-converb form of the citation marker *te-*. This conjunction can introduce reported direct speech (35) or, due to further grammaticalization, clauses with a purposive or causal meaning (36). Some of the other Chuvash conjunctions seem to be (half-)calqued from Russian, cf., e.g., Chu. *mən-šən tesen* [what-CAUS/FIN say-COND.CVB] and Ru. *potomu čto* ‘because’.

(35) *pir-ən* *pâᵒrâᵒv-a* *vəᵒsem śâᵒl-sa* *xâᵒvar-nâ* *te-se*
1PL-GEN calf-DAT/ACC 3PL save-CVB leave-PST.PTCP say-CVB
kala-r-ə *vət* *anne*
say-PST-3SG EMPH mother

‘Mother said that they had saved our cattle.’ (V. M. Kuz’mina, “Jămra ačisem”)

(36) *epə* *un-ran* *və̂^{o}l* *Izraiĺ* *patš-i* *an=pul-tə̂r* *te-se*

1SG 3SG-ABL 3SG Israel king-POSS.3 NEG=be-IMP.3SG say-CVB

pə̂^{o}rə̂^{o}n-t-ə̂m

turn.away-PST-1SG

‘I have rejected him as king over Israel.’ (Lit. ‘I have turned away from him saying: “Let him not be the king of Israel”!’) (1 Samuel 16: 1)

As is common in Transeurasian languages, Chuvash makes use of constructions based on the simultaneous/sequential converb (*-sa*) in order to coordinate clauses. Long clause chains based on forms in *-sa* are possible, e.g., in a narrative context. Other frequent means of clause coordination include the coordinating marker *=tA* cliticized to the first conjunct and simple clause sequencing.

Along with the presented strategies of clause linking, Chuvash, especially in its spoken non-standard form, makes use of clauses based on relative pronouns and of other Indo-European-like patterns that are most likely calqued or half-calqued from Russian.

27.6 Lexicon

Most of the Chuvash vocabulary used in daily communication (around two-thirds, according to my estimate) is of ultimate Proto-Turkic origin. In the domain of basic vocabulary, the ratio of inherited lexemes reaches 90–95%. The major differences between Chuvash and Common Turkic basic vocabularies originate in semantic shifts accumulated over the two thousand years’ evolution in parallel, cf. PTk **yūŕ* > Chu. *śə̂^{o}var* ‘mouth’, CTk **yʉ̄z* ‘face’ (CTk **aɣïz* ‘mouth’), PTk **sį̄ārïɣ* > Chu. *šurə̂* ‘white’, CTk **sarïɣ* ‘yellow’, (CTk **āq* ‘white’), etc.

The Chuvash basic vocabulary exhibits many of the semantic distinctions that are not universal cross-linguistically but were characteristic to Proto-Turkic and can also be found in Common

Turkic languages, such as the ‘arm’/‘hand’ distinction, the lexical distinction in terms for older and younger siblings (see van Berlo, this volume: Chapter 38), and the 2D/3D distinction in terms for ‘round’ and ‘thin’ (but not for ‘thick’, as in other Turkic languages).

Lexical isoglosses connect Chuvash with different subgroups of Common Turkic. For example, an archaic isogloss that is known from Old Turkic and is preserved in Chuvash and South Siberian Turkic as geographically peripheral Turkic languages is the word for ‘wife, woman’: Chu. *arə̂m* < pre-modern Chu. *av-r-ə̂-m* < VBulg. **ɛv-r-i-m* [house-LOC-NMLZ-POSS.1SG], lit. ‘the she-who-is-in-the-house of mine’, cf. Siberian Turkic *ep-či*, a derivate of *ep* ‘house’ with a similar meaning (originally ‘housewife’). A remarkable number of lexical items, including those in the domain of basic vocabulary, relate Chuvash specifically to the Oghuz subgroup, cf. Chu. *kə̂ᵒvapa* and Oghuz **gōpek* ‘navel’ of the same origin vs. reflexes of **kīn-dük* in the other Turkic languages. Besides, shared lexical and semantic isoglosses connect Chuvash with the Volga Kipchak languages, which is apparently a result of the extensive co-evolution of Turkic languages in the Volga-Kama region.

Borrowed vocabulary of Chuvash goes back to different sources. An early layer of borrowing is represented by some Old Ossetic (“Alanic”) loans from the Volga Bulgharian period or even prior to it (e.g., Chu. *kəle* ‘heel’). Contrary to a commonly cited misconception, the number of loanwords of substratum Uralic origin in Proto-Chuvash as revealed so far is not large at all, with around 50 reliable borrowings from Mari (e.g., *kə̂ᵒtkə̂ᵒ* ‘ant’) and a dozen terms of Permic origin (e.g., *pə̂ᵒši* ‘elk’). The Viryal and particularly Northwestern dialect adopted some more Mari lexemes due to relatively recent interactions. Since the 14th century and until the most recent times, Chuvash and particularly its Anatri dialect have been crucially affected by Tatar, with at least several hundred borrowings (e.g., *iš-* ‘to swim’). In addition, Tatar served as a mediator for most of the Arabic and Persian loanwords in Chuvash (e.g., *layə̂x* ‘good’, *čun* ‘soul’) and partly for words of Mongolic origin, while others were borrowed directly form Middle Mongolian to the ancestor of Chuvash (e.g., *pusaxa* ‘threshold’). Russian has increasingly influenced Chuvash since the mid-16th century, resulting in an enormous amount of borrowings belonging primarily to the semantic domains of

administrative rule, society, and particularly modern life.

Acknowledgements

The research leading to these results has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 646612) granted to Martine Robbeets. I am grateful to all the colleagues who provided their comments on the first draft of this paper. Klára Agyagási, the reviewer of this chapter, drew my attention to some less-accessible literature on contact relations of Chuvash. Artem Fedorinchyk and Maria Kholodilova shared their expertise on different aspects of descriptive grammar of the contemporary Chuvash, especially on the syntax part. Ruslan Idrisov offered his feedback on my description of the Chuvash phonology. A special thanks goes to Elena Savelyeva, who created a map of the Chuvash-speaking area especially for this chapter.

---

[1] I use the label "Bulgharic" as a classificatory term for one of the two major branches of Turkic, while "Bulghar" is applied only to those extinct Bulgharic varieties that are associated with the historical Bulghar tribes.

[2] My transcription does not mark this allophonic alternation for the sake of simplicity.

This is a draft version of a chapter that appears in Robbeets, M. and A. Savelyev (eds). *The Oxford Guide to the Transeurasian Languages* (OUP, 2020), see https://global.oup.com/academic/product/the-oxford-guide-to-the-transeurasian-languages-9780198804628. The research leading to these results has received funding from the European Research Council under the Horizon 2020 Program/ ERC Grant Agreement n. 646612 granted to Martine Robbeets.