

Erasmus Mundus Intensive Program
Turkic languages and cultures in Europe (TLCE)

Nutshell Chuvash

by

András Róna-Tas
Szeged, Hungary

1. The place of Chuvash within the family of the Turkic languages

1.1. General remarks. Proto-Turkic is the most ancient form of the Turkic languages which can be reconstructed by all available historical and present data of all Turkic languages. Ancient Turkic is the name of the earliest period of the changes of Turkic languages, which lasted until the beginning of Old Turkic in the 4th-6th centuries A. D. Proto-Turkic is a reconstructed set of linguistic features and their relation. Ancient Turkic was a living language about which our knowledge is limited. Ancient Turkic is a set of formulae, Ancient Turkic is a language the contours of which we know from its descendants. Proto-Turkic is our invention which has to be reformulated after new discoveries, Ancient Turkic has to be discovered. Ancient Turkic has to be divided into two chronological layers. In Early Ancient Turkic no stabilized dialects can be assumed. In Late Ancient Turkic the ‘ancestor dialects’ of the later Turkic language groups were formed. The speakers of these ‘ancestor dialects’ lived in a more or less continuous area. Though the speech of groups living further from each other became more and more different, their speakers understood each other. On the other hand the territory inhabited by all speakers was crossed by isoglosses. Lines which connected points where the same linguistic features were used.

1.2. The separation of the Old Chuvash from the other Turkic languages. One of the important isoglosses of Late Ancient Turkic was the line which connected the places where the Turkic word for ‘calf’ was pronounced as **bizaġu* and the line which connected the place of speakers who pronounced the same word as **biragu*. It has to be remarked that according to the results of modern dialectology most isoglosses of the similar cases fall together, but some lexical items may ‘go astray’, may have an isogloss which is not everywhere following the common line. Another important isogloss was the line which connected the places where the word for ‘stone’ was pronounced as **tāš* and another which connected the places where it was pronounced as **tāl*. Interestingly enough the isoglosses of the oppositions *r:z* and *l:š* coincided, but some words transgressed this line. E.g. the Turkic word for ‘dream’ *tüš* did not observe the *š/l* line and the form *tül* appears in Old Uigur otherwise an *š*-language. This Late Ancient Turkic was the time when the ancestor of the Chuvash language separated from the other Turkic languages. The speakers of the ancestor of the Chuvash language moved away from the other Turks. Here we have to make one more remark. We now know for sure that a smaller group, or groups of the ancestors of the present Chuvash language remained in the East and joined other Turkic peoples. This is of general importance. People did not split according to their dialects. The dialects were formed with the people who separated.

1.3. The two main groups of the Turkic languages. The most simple way to name the two groups is to call them *Ogur* and *Oguz*, and the languages *Oguric* and *Oguzic*. To the first of them pertains Chuvash, to the second all other languages. The Oguzic languages are sometimes also called Common Turkic languages. The SW Turkic languages later separated from the Old Oguz group, and are also called in some works *Oguz*. Present Chuvash has preserved old traits not preserved in other Turkic languages and has innovations not appearing in other Turkic languages. Features disappeared in Chuvash which have been present in Ancient Turkic and are present in all other languages. We shall give only a few characteristic example for each.

1.3.1. The pronominal stem. The stem of the pronoun first person singular ‘I’ was **be* in Ancient Turkic (AT). This was preserved in Chuvash and later a deictic prefix *e-* was added to it. Now it is written as <epě> and pronounced as /eBě/. In all other Turkic languages AT **be* got from the oblique stem the so called “pronominal” *-n*, that is *-n* was generalized (*be+n*). Parallely the closed *-e-* changed to an open *ä*. Later *bän* changed in some languages by nasal assimilation

to *män*. The genitive case of the same pronoun was in AT **beniŋ*. This changed in later Chuvash to **mänin* and to-day it is *manän*. This does not basically differ from the other Turkic languages, where we have *bänin*, *bänün*, *mänin* etc., though in early OT texts we find *beniŋ* etc. Thus we see, that the ancestors of the Chuvash separated from the other Turkic languages before the *-n* was generalized and appeared also as the final of the stem of the nominative case.

1.3.2. The plural. All Turkic languages but Chuvash have the plural marker *+lAr*. This itself seems to be a later generalization, because OT had some other plurals as well, as it can be seen in such cases as *oġul* ‘son’ > *oglan* ‘sons’, or *är* ‘man’ > *ärän* ‘men’ further *be(n)* ‘I’ and *biz* ‘we’, etc. The suffix *+lAr* may originally have been the marker of individual live beings. The *+lAr* marker of plurality does not exist in Chuvash. The Chuvash plural suffix is *+sem*. This marker does not obey the vocalic harmony, it is also co-occurring with back vowels as e.g. in *värman* ‘forest’ (< *orman*) > *värmansem* ‘the forests’. In some Chuvash dialects the back vocalic form of *+sem* can be observed. The oblique stem of *+sem* is *+sen*.

1.3.3. The possessive marker of the third person. In all Turkic languages the marker of the 3rdPSing is after vocalic finals *+si*. This is not present in Chuvash. See e.g. *pulă* ‘fish’ > *pulli* (< *pulă+i*), *laša* ‘horse’ > *laši*. However we find the suffix in a limited type of the lexicon. In words denoting family relationship the suffix *+si* > *šě* is present: *appa* ‘elder sister’ > *appašě* ‘his/her elder sister’. This may, but not necessarily does, point to the fact that *+si* was originally a suffix joining only words denoting living persons. In OT and all Turkic languages the suffix *+I* of the 3rdP possessive relation is assimilated according to vowel harmony. Front vocalic *+i* appears after front vocalic stems and back vocalic *+i* appears after back vocalic stems. Later, in some Turkic languages there is also an assimilation according to labial and non labial vowels in the stem. In Chuvash this is not the case. *+i* > *ě* is front vocalic after all types of stems, also after back vocalic ones: *laša* > *laši*, *värman* ‘forest’ > *värmaně*. In this case Chuvash preserved the Ancient Turkic situation yet reflected by some of the Old Turkic Runiform inscriptions.

1.4. The phonological structure of Chuvash

1.4.1. The vowels. Chuvash has according to quantity two types of vowels: “normal” and “reduced”. In the Anatri dialect and in the literary language based on it, there are only two reduced vowels */ă/* and */ě/*. In the Viryal dialect two reduced labials are yet preserved */õ/* and */õ̃/*. There are three “normal” or full front vowels: */e/*, */i/*, */ü/* and three back vowels: */a/*, */i/* and */u/*. The system is not perfectly balanced, because */e/* is mid-low while */a/* is a low vowel. Thus the vocalic system of Chuvash has eight vocalic phonemes in the Anatri and ten in the Viryal dialects.

The vowels:

High (closed)	/u/	/i/ /i/	/ü/
Mid-high (mid-closed)	(/õ̃/) /ă/		/ě/ (/õ̃/)
Mid-low (mid-open)			/e/
Low (open)		/a/	

1.4.2. The consonants. Under the influence of the Finno-Ugrian languages spoken in the Volga area Chuvash has lost the voiced feature of most of its consonants. All stops, affricates and sibilants are voiceless strong consonants, however in intervocalic and sandhi position they

become voiceless weak ones. The consonants are the following (in brackets the weak allophones): /p/ (~B), /t/ (~D), /k/ (~G), /č/ (~j), /š/ (~ž), /ś/ (~z), /χ/ (~γ), /v/, /m/, /n/, /l/, /r/, /y/. Under the influence of Russian all consonants before and after front vowels are palatalised. This feature is stronger in the case of stops than with other consonants.

1.4.3. The word stress. The main rule is that on a reduced vowel there cannot be stress unless the word has only reduced vowels. In this case the first vowel is stressed. If all vowels are full vowels, the stress is on the last syllable. In mixed words the last full vowel has the stress.

1.5. The main morphological features of Chuvash

1.5.1. The morphophonological types of nominal stems. A nominal stem can have as final phonological segment either a consonant or a vowel. However the final full labial vowels surface before suffixes as diphthongs: E.g. *šu* ‘water’ > *šiuě* <šivě> ‘its water’, also *šěršiv* ‘country’ (cf. OT *yersub*), *χü* ‘bosom of cloth’, *χěvě* ‘the bosom of her cloth’ (cf. OT *koy* < *kōñ > *koyun*). The final reduced vowels change into full vowels if the suffix begins with or consists only of a vowel. E.g. the Accusative–Dative has the marker +A. If the word ends with a full vowel the suffix is +nA: *laša* (*laža*) ‘horse’ > *lašana* ‘the horse (Acc)’ or ‘to the horse’. But if the word ends with a reduced vowel the preceding consonant becomes long and the final a full vowel: *pulă* ‘fish’ > *pulla* ‘the fish (Acc.)’ or ‘to the fish’.

The stems ending in -t change their -t to -č before original *i: *yat* ‘name’ > *yače* (*yaže*) ‘his name’ (OT **āti*). Stems ending in -l, -n and -r have the locative marker +tA instead of +rA. See *pulă* ‘fish’ > *pulara*, but *tetel* ‘fishing net’ > *tetelte* (← Chermis *tetele*).

1.5.2. Order of suffixes. Nouns get suffixes in the following order: derivation+ possessive+ number+ case. e.g. *kuś+lăχ+ămăr+sem+pe* (*kuźlăγămărzemBe*) ‘with our glasses’ (**köz+lik+imiz+ler+birle* instead of **közliklerimizbile*).

1.5.3. The case suffixes of the nominal stems. Old Turkic cases: Genitive: +(Ā)n, Accusative-Dative + (n)A, Locative +rA/tA, Ablative +rAn/tAn, Instrumental (Ā)pA. New cases: Privative: +sĀr, Causal-Purposal +šAn. For the last two see OT +*siz* and *üčün*.

1.5.4. The pronouns and pronominal stems: *epě/man-*, *esě/san-*, *văl/un-*, *epir/pir-*, *esir/sir-*, *věsem/věsen*.

1.5.5. The possessive suffixes: Singular: 1stP +(Ā)m, 2ndP+u/ü 3rdP + ě/i, Plural: 1stP. +m(Ā)r, 2ndP +(Ā)r, 3rdP +ě/i.

1.5.6. The numerals. The Chuvash has two types of numerals. The first series is used in adjectival phrases, the second in nominal ones. As in: “I have *three* children, How many children do you have? *Three*”. The numerals are the following: (in brackets the emphatic ones): 1. *pěr(ě)*, (*pěrrě*), 2. *ikě* (*ikkě*), 3. *višě* (*viśšě*), 4. *tăvată* (*tăvattă*), 5. *pilěk* (*pillěk*), 6. *ultă* (*ulttă*), 7. *śičě* (*śiččě*), 8. *sakăr* (*sakkăr*), 9. *tăχăr* (*tăχčăr*), 10. *vun(ă)* (*vunnă*).

1.5.7. The morphological types of the Chuvash verbal stems. The Chuvash verb may have vocalic, diphthongual and consonantal finals. There are 10 verbs which lose their final -r in certain conjugational forms e.g. *par-* ‘to give’ > *pamasčě* ‘he does not give’.

1.5.8. Temporal and personal markers of finite verbs

	1.Present	2.Future	3.Simple past, Imperfect	4.Pluperfect	5.Imperative Prohibitive
Marker/negation	-t/-mAs-t-	ϕ/-mĀ-	-r(-t-*)/-mA-r-	-As(s)A-tt- mAsĀ-tt-	ϕ an (before verb)
1Psg	-Āp	-p	-Ām	-Ām	-Am
2Psg	-Ān	-n	-Ān	-Ān	ϕ
3Psg	(palatalization)	ϕ	-ĕ**	-ĕĕ**	-tĀr
1Ppl	-pĀr	-mĀr	-ĀmĀr	-ĀmĀr	-Ar
2Ppl	-Ār	-r	-Ār	-Ār	-Ar
3Ppl	-śĕ***	-ś	-ś	-ĕś**	-ĕĕĀr

* after consonants, **preceding -t- becomes -ĕ-, ***preceding -t- becomes -ś

1.5.9. Modal markers. The *conditionl marker* is -ĀttĀ- (3rdP -ĕĕĕĕ-) with the personal suffixes of Type 3. The *concessive marker* is -in-. In the second person singular -sĀn, plural -sĀr. In the 1st and 3rd person -in follows the personal markers of Type 1. The *optative marker* is -ĕĕ added to the concessive form.

1.5.10. Verbal nouns in adjectival function. They may have either temporal or modal function. Temporal function: Present time -AkĀn, Past time -nĀ, Future: -As. Modal functions: Necessity -mĀllĀ, Satisfaction: -mĀlĀχ, Possibility: -i, Pretence: -Anśi.

1.5.11. Verbal nouns in adverbial function. Also they may have temporal and modal functions. Different temporal relations, as simultaneity, anteriority, posteriority may be expressed by the verbal noun ending in -sĀ. In some cases the following finite verb may preserve its original meaning, in some cases the verbal noun carries the semantic load and the finite verb loses its original meaning. Conditional function has the verbal noun ending in -sĀn.

1.5.12. Adverbs. Chuvash has simple adverbs as *irĀ* 'well' and adverbs derived with the suffix +lĀ.

1.5.13. Postpositions. The postpositions can be grouped according to their semantics and according to the case which they govern. They can denote space, time, cause and aim, comparison, uncertainty, instrument etc. Some of them stand with the nominative, other with accusative-dative, or with ablative. A few can be used with either nominative or accusative-dative.

1.5.14. Connectives. Special words may be used for connecting co-ordinated phrases or sentences. Simple connection: *ta, tata*, adversary connection: *anĕāχ, śapaχ, śapaχ ta*, separating sentences of different choices *e...e, pĕrre...tepre*, mutually excluding sentences: *te...te*. Others are separating subordinate sentences. They are divided into consecutive (*mĕnśĕn tesen*), concessive (*pulin te, pulsĀn ta, śapaχ*), conditional (*enĕen, eχer, pĕr*) and purposal (*tese*).

1.5.15. Emphatic words, are words which emphasize a part of the sentence. They can stand before or after the word emphasized. Most frequent are those which stress the importance of the word, as -aχ, -ta, *mĕn*.

1.6. Lexical peculiarities of Chuvash. The lexical stock of the Chuvash language consists of three groups: 1. Words inherited from Ancient Turkic, 2. Words formed in the separate life of Chuvash by inner word formation and 3. Loanwords. All groups have several subgroups. Most interesting are the loanwords. Chuvash has early loanwords from other Turkic languages, e.g. *pusäk* 'big' ← *bäzük* < *bädük*. The Middle Chuvash form **bazuk* was borrowed by Tatar as *bazik* 'thickset, large'. Later Chuvash copied many words from Tatar and Bashkir. At least three chronological layers of Volga Kipchak loanwords can be distinguished. All Middle Mongolian loanwords came through Tatar into Chuvash. Chuvash borrowed words from the neighbouring Finno-Ugrian languages as Cheremis. There is a layer of Chuvash lexicon which originates from a yet unknown substratum. There are Russian loanwords from the 10th century until the most recent times. Some of them were copied from or through Tatar. Interesting are the words of Iranian origin. A few of them were copied in a time, when Chuvash speaking groups contacted the Alans in the South, a few others were borrowed through Finno-Ugric or Volga-Kipchak languages. Chuvash has a very interesting group of words of Arabo-Persian origin, from a time when the Chuvash people joined the Muslim faith.

2. The relation of Chuvash to Mongolian. The Altaic hypothesis

2.1. General remarks. Some scholars claim that the Turkic language family pertains to a greater one. According to the restricted hypothesis Turkic, Mongolian and Manchu-Tunguzian are genetically related. According to the extended hypothesis, the genetically related Altaic languages comprise also Korean and Japanese. According to the so called Nostratic hypothesis Altaic is a member of an even more extended group of genetically related languages. To them pertain the Indo-European and the Semitic languages, the Dravidian, the Kartvelian group of Caucasian languages and the Uralic languages.

In the following we shall deal only with the restricted hypothesis and even within this hypothesis only with the relationship of the Turkic and Mongolian languages. In the research of the Turko-Mongolian genetic relationship a key role was attributed to Chuvash. According to the two great scholars of the Altaic hypothesis G.J. Ramstedt and N. N. Poppe Chuvash preserved many traits of the common Altaic proto-language. These traits are only present in Chuvash and Mongolian, while changed, or are not preserved, in all other Turkic languages, called also Common Turkic. Recently S. Starostin (died in 2005), A. Dybo and O. Mudrak published an etymological dictionary of the Altaic languages in three volumes. In many aspects this work and the other publications of these authors represent a new phase in the study of Turko-Mongolian relations. While Ramstedt and Poppe tried to set up linguistic rules, according to which certain changes occurred in the Altaic languages, and proposed for each rule a few examples, Starostin and his team selected a stock of the lexicon comprising about 2800 "etymologies". They tried to formulate rules according to which the words, pertaining to the alleged proto form can be genetically connected. This resulted in about 1.500 new etymologies and a great set of new rules. The hypothesis met sceptical opinions from the earliest time. Some of the scholars as Clauson, Doerfer and Ščerbak tried to prove that the Turkic and Mongolian languages are not related. From the methodological point of view their position was weaker, because it is impossible to prove that something does not exist. However they and other scholars brought a vast material to surface, which did not support the hypothesis of the genetic relationship. The opinion of Róna-Tas is that the genetic relationship of Turkic and Mongolian is not yet proved. This, however, does not exclude the possibility of a genetic relationship. First we have to separate those facts which are due to loan-relations, and then what remains has to be studied.

2.2. The so called rhotacism and lambdaism. The two terms are derived from the respective Greek names of the letters <ρ> *rho* and <λ> *lambda*. The term rhotacism was first used in the description of the Latin language. Some words ending in *-s* as e.g. *mōs* ‘will, use, tradition, usage etc.’ form their oblique stem changing the *-s-* through *-z-* to *-r-*, *mōs* > *mōris-*. Other, as e.g. *vās* ‘vessel, pot’ do not change their final *-s*: *vās* > *vāsis*. Those *-s*, which change to *-r-* were originally *-*z-*. Thus *mōs* < **mōz* > *moris*, but *vās* < **vās* > *vasis*. In case of Turkic *buzagu* ‘calf’ WOT *buragu* > Chuvash *pǎru* (→ Hungarian *borjú*) the original idea was that the PT *-z-* changed to *-r-*. In Mongolian we find : *birāgu(n)* and Middle Mongolian *bura’u* ‘calf in its second year’. The idea of Ramstedt and Poppe was that Mongolian and Chuvash preserved an old *-r-*, while the common Turkic languages changed the original *-r-* to *-z-*. T. Tekin therefore changed the terminology and called this feature: zeticism (Greek zeta ζ). However Ramstedt had to solve the problem, that why do we find *-r-* in the common Turkic languages, as in *kara* ‘black’, if all *-r-* sounds changed to *z*. The solution was, that in Proto Altaic there existed two different types of *r*. They were called simply *r₁* and *r₂*, claiming that the first remained *r* and the second changed to *z*. Mongolian and Chuvash preserved the second as *-r-*, while common Turkic changed them to *z*.

Altaic	<i>r₁</i>		<i>r₂</i>		
	Mongolian <i>r</i>	Turkic <i>r</i>	Mongolian <i>R</i>	Chuvash <i>r</i>	Turkic <i>z</i>
	<i>kara</i>	<i>kara</i>	<i>buragu</i>	<i>buragu</i> > <i>pǎru</i>	<i>buzagu</i>

This argumentation has serious difficulties. 1. There is no Turkic or Mongolian language in which there would exist two types of phonologically opposed /r/ sounds. 2. Though there were experiments to attach to one or the other /r/ sound special phonetic traits, they did not succeed. In fact, we know, by the hypothesis, that the two sounds were in phonematic opposition. 3. Though according the Altaic hypothesis Mongolian and Chuvash preserved the “Altaic” opposition *r₁ : r₂* this is not the case, there is no opposition of two kinds of /r/ or any similar consonant which would have been evolved from one of them in Chuvash or Mongolian. 4. /z/ is a well known phoneme in common Turkic. It corresponds in some cases in Chuvash to *ś* (see e.g. OT *kōz* ‘eye’ ~ Chuvash *kuś*) and in Mongolian to *s*, cf. Turkic *kez-* ‘to travel’ ~ Mongolian *kesü* ‘to travel’. This allows the conclusion that Chuvash and Mongolian had no /z/ or they lost them. If Chuvash had never a /z/ and all other Turkic languages had it, this would need an explanation, but there does not exist such. It remains as a possibility, that Chuvash lost its *z* or changed it into another sound. 5. Since we do not know what was the phonetic character of *r₂*, but know that it is represented by *z* in the majority of the Turkic languages, what is the argument against a hypothesis, that it was exactly the phoneme /z/? The fact that it is represented in Mongolian and Chuvash by /r/ needs the presupposition that the opposition *r₁ : r₂* was neutralized (they have only one *r*). 6. Finally, if the *r₁ : r₂* opposition existed it should have exist in the time of the Altaic proto-language. Some of these arguments may be questioned, but all of them together have at least to give thought for looking to another solution.

Now we have some data for the chronology of the rhotacism. For the word denoting the stirrup we find in Chuvash the word *yārana*. This form goes back to an earlier form **irāñä*. One part of the Turkic languages have a form which goes back to **üzengü* (SW, NW and Turkestan), the other group has forms going back to **izāñä* (Baraba, Khakas, Tuvan, Yakut and Yellow Uigur). There can be no doubt that the name of this very important and special part of riding and fighting has the same origin. Mongolian *dörüge* ‘stirrup’ is an independent Mongolian word formed from Mongolian *dörü* ‘iron or rope nose-ring, lead rope (fore cattle), rope handle for basket’ and has no connection with the Turkic word. It can be seen, that the Chuvash and the Siberian/Yellow

Uigur words go back to a form *iXäjä* where X may be either /r/ or /z/. Bang connected the word with OT *üze* ‘above’. Though this was many times repeated, this is morphologically unclear. In any case we have to depart from the earliest reconstructed form **üXeje* or **üXeñü* and from this, with assimilation was formed **iXeje*. The change of X to Chuvash *r* and common Turkic *z* had to occur after the word became a term. Now we have excellent archaeological data on the appearance of the stirrup, which was a very important invention in nomadic warfare. The earliest datable stirrups are not earlier than the 3rd century A.D. The stirrup was not known in the soldiers of the Roman Empire, and first stirrups were used in Europe by the Avars, who arrived from Inner Asia in the middle of the 6th century. Somewhat earlier it was used in South Siberia, in China and in Korea. The stirrups in the archaeological finds were metal stirrups, but we know from clay figures and written sources that before the metal stirrup there existed wooden stirrups and rope stirrups. Originally these stirrups served for making easier the mounting of the horse for sick or pregnant persons and were used only on one side of the horse as “ladders”. In some languages the word for stirrup has until now also the meaning ladder e.g. Turkic *baskač*. The detailed analysis of the history of the stirrup shows that its history had two phases. In the first phase it was only used as a support for mounting on one side of the horse, and in the second phase, used on both side of the horse, it became the help for riding and fighting (mainly for shooting backwards the arrows). Taking into account all these facts, we can conclude that the stirrup, as such, appeared in the last centuries B.C. Since the sound X had to split into *z* and *r* after the appearance of the stirrup, we can now fix the date of the split of X. This could only have been in the last hundred years B.C or the first ones A.D. If an Altaic protolanguage comprising Mongolian and Turkic ever existed, it had desolved thousand years earlier. The first conclusion is that the split of X, whatever it was, occurred in the separate life of the Turkic language. The formation and separation of Chuvash occurred simultaneously with this split. Since in Chuvash $r_1 : r_2$ opposition was simultaneously neutralized, that is: both phonemes became *r* (*kara* > *χura*. *buXagu* > *pǎru*) this is a Chuvash innovation. Chuvash did not preserve the original Altaic feature. But if Chuvash had **buXagu* why has Mongolian **buragu*? For the same reason as Hungarian has **buragu* (>> *borjú*). Both are loanwords. If both are loanwords, then Chuvash had already *-r-* on place of X at the time of the Chuvash Mongolian, and later at the Chuvash Hungarian contacts. All the above mean that whatever the phonetic peculiarities of r_2 were, may it have been a kind of /r/, of /z/ or a third sound, the question of rhotacism does not support the Altaic hypothesis.

According to Róna-Tas r_2 was in fact /z/. The OT word for ‘copper’ later for ‘brass’ is *yez*. This existed in Old Chuvash and has been borrowed by Moksha Mordvin, where it sounds now as *šerä*. The early Mongolian form is present in the first word of the composition *jer jebseg* ‘weapons, arms’. These words go back to a Tocharian form **yäs*, see TocharianB *yasa*, A *wäs*, a cognate of Latin *aurum* < **ausum* etc.

The question of rhotacism has many other aspects. Important is that in many common Turkic words we find a change of /z/ to /r/, as e.g, *tīz* ‘knee’ and *tirsgek* ‘elbow’, *semiz* ‘fat’ *semri-* ‘become fat’. T. Tekin wrote more important papers on this feature. Another change is that of the intervocalic *-d-*. It became in Chuvash *-r-*. This could happen only through $d > \delta > z > r$. Benzing claimed that the two types of rhotacism occurred at the same time. This is untenable. The second rhotacism occurred later, after the 10th century A.D.

In place of common Turkic *š* we frequently find *-l-* in Chuvash: common Turkic *bēš* ‘five’ ~ Chuvash *pil-ěk*. This is, however not as consequent as the case of /r/ on place of /z/. Common Turkic *baš* ‘head’ is in Chuvash *puś* and not **pul*. The explanation of Ligeti for this feature was that the change of *š* to *l* began, but was interrupted. The reason may have been the western migration of the Proto Chuvash. On phonological grounds lambdaism has the same problems as rhotacism, and cannot be used as a proof of the Altaic hypothesis.

3. The sources of the study of the history of Chuvash

3.1. The earliest sources of the history of the Chuvash or more precisely the Oguric language are of linguistic nature. In the Common Samoyed language the word for ‘hundred’ is **yür*. This reflects Chuvash **yür* or **jür*. In the second case we have to reckon with the fact that in Common Samoyed, the common ancestor of all Samoyedic languages, had no *ǰ*- and substituted *y*- in place of *ǰ*- (present Chuvash: *śer*). The dissolution of Common Samoyed, that is the *terminus ante quem* is according to P. Hajdú the first centuries around the birth of Christ. The word for ‘pine’ is in Chuvash *χir(ǎ)*. The Proto Samoyed form can be reconstructed as **kadi* < *kâât- âëÿâ*. This went through the second rhotacism in Chuvash and is a good witness of the fact, that the Chuvash speaking people came from South Siberia. In the Kejilig *χobu* (Jenisei No. 45, Vasilev 1983:31) inscription written in runiform letters we find the name of a certain *Kümüł bodun*, whose chief was *Kümüł üge*. He became *üge* in his 30th year (*otuz yašimga*) and died in his seventieth year (*yetmiš yašimga*). The text is clearly written in a common East Old Turkic language, but the word *kümüł* ‘silver’ is with lambdaism (cf. Chuvash *kēmēl*, common Turkic *kümüš*). One has the impression that the “Silver people” were former speaking a Chuvash-type language from which in the 8th-9th centuries only scattered names have been preserved. The Silver people appear also in the Kizil Čira inscription as *kümüł bodun*. The same people appear in the Chinese sources. The Kitan, a people who spoke a Mongolian language was the ruler of the steppe from the 10th to the 12th century after the collapse of the Uigur khaganate. They occupied the northern part of China and founded the Liao dynasty. In the *Liao shi*, the work on the history of the Kitans, we find the a tribe name written as *Wu-gu-li*, with other characters *Yu-gu-li*, written also as *Yu-jue-lyu* and *Yu-jue-li*. This tribe lived in Siberia. Its earlier name can be reconstructed as *Ugur(i)*, later it became *Yügur(i)*. This means that parts of the Ogur tribal confederation, originally speaking a Chuvash-type language at least preserved their name in the form *Ugur* until the 10th-12th centuries, when it changed to *Yugur*. Masudi, the traveller, who died in 956, mentioned in his geographical book among the Western Siberian people a group called *Yigur*, evidently the same group. And the Khwarezm shah in a letter dated 1182 mentioned a certain *Yugur-zadagan*, who came to his court from his people. These scattered data from linguistic and written sources show that the Chuvash speaking groups split in South Siberia, their major part came to Europe and later became part of the Onogur, the Bulgar and the Volga Bulgar polities. But a part of them remained in Asia and assimilated to the Turks living there. The greater part of the early Chuvash moved to the Kazak steppe and appear there in the 4th century A.D. They move further to the West in the middle of the 5th c. From this on they frequently occur in the written sources of the West.

The move of the bulk of the Oguric tribes to South-West and then to West made ended their impact on the Samoyed, Paleo-Siberian and Mongolian neighbours. This impact, which lasted until about the 3rd-4th centuries A.D. was great. It is mirrored by the Oguric loanwords of these languages. Words of the type *buzagu* ~ *buragu* → Mongolian *buragu* were copied in the time before the 3rd-4th centuries. They offer a rich material for the reconstruction of this phase of the history of the Oguric languages.

3.2. According to the Byzantine sources after the death of Attila (453), the king of the Huns, and the withdrawal of the Huns to the North of Caucasia, new people appeared on their horizon. Priskos call them *Onogur*, Theophilaktos Simokattes *Unnugur*. This was the name of a tribal confederation of (originally) ten *Ogur* tribes, as *Tokuz Oguz* the confederation of the Nine *Oguz* tribes. But the name of the single tribe was *Urog* (Priskos), *Ugur* (Menandros), *Ugur*, *Ogór* (Theophylaktos). Behind these data we can discover the name *Ogur* > *Ugur*. This is nothing else than the *r*-form of the ethnonym *Oguz*. This ethnonym had the word stress on the second (last)

syllable. This we know from the fact that already in the tenth century we meet forms like *Guzz*, which may have evolved only from *Ogúz* > *Ugúz* > *Guz(z)*. The same happened with the name *Ogur* which changed in certain position to *Gur*. We find this form in the name of the *Sara-gur* (Priskos), *Utur-gur* (Utigur, Utrigur,) (Prokopius, Menandros), *Kutur-gur* (Kutrigur, Kotrager, Kurturgur etc.) (Prokopius, Menandros, Theophilaktos Symokattes). In other words from the middle of the 5th century we find the Ogur tribes in the region North to Caucasia. Their name is clearly *r*-Turkic.

3.3. Loanwords in Hungarian. From the 6th century we find a greater variety of sources on the history of the Oguric languages. This is the time, when the Ogur-Hungarian contacts began. Hungarian has about 420-450 Old and Middle Turkic loanwords, which were copied before the Ottoman invasion of Hungary in the middle of the 15th century. Some of them are of Pecheneg and Cuman origin, but most loanwords pertain to the Oguric type of languages, see e.g. *buzagu* ~ *buragu* → Hungarian **burgu* > *borjú*.

3.4. Loanwords in Alanian-Ossetic. The only group of people who spoke an Iranian language and remained in this part of the steppe were the Alans or As people. Their language has been preserved by the two Ossetic groups the Iron and the Digor. In the early 13th century, together with Cumanians, a group of the Alan-As entered Hungary, and were known as *Yās* (*Jász*). The Oguric tribes and the Alans had close contacts until the Pecheneg invasion at the end of the 9th century. Ogurs are mentioned in the heroic epos of the Ossetians, the Nart epos. The name of 'reed' in Ossetian is *qamīl* a loan from Oguric *qamīl* (> Chuvash *χāmāl*), cf. common Turkic *qamīš*.

3.5. Common botanical names in Turkic, Ossetic and Hungarian. There exist about 15 botanical terms the most of which appear in Turkic, Ossetian and Hungarian. Some of them are of Caucasian origin, other are Turkic, and some are Alanian. The 15 botanical terms (for 'ash tree', 'cornel', 'pear', 'hop', 'blackthorn', 'bulrush', 'beech', 'hawthorn', 'hornbeam', 'cornel', 'walnut', 'hazelnut', two names for 'oak', and 'reed') point to the Don-Kuban region and the long lasting contacts of the Oguric, Alanian and Hungarian speaking groups.

3.6. Early Ogur-Slavic contacts. The slow eastwards migration of the Slavic tribes reached around 900 the line between the present Nizhnij Novgorod and Rjazan, that is the Oka river. They remained in the forest region but had trade and other contacts with the steppe and the people living there. These contacts became stronger in the 7th-8th centuries. The Southern Slavs living on the Balkan and the East Slavs living in what is to-day the Ukrain, Belarus and Russia, had close linguistic contacts with the Oguric tribes. These contacts are reflected by a few dozen loanwords in Old Slavic. Among them e.g. *šaran* 'a type of carp' ~ Oguzic *sazan* 'carp' shows two different Oguric features while Chuvash *saśan*, *sāśan* are copies of Tatar words. Interesting is Ossetic *sarzan* (~ *sazan*) 'carp' which may be simply a copy of Turkic *sazan*, but may preserve an older form with *-r-* which later has been contaminated with *sazan*.

3.7. Sources on the language of the Avars. The Avars, surely a people of different origins, passed the North of Caucasia in 555 and occupied the Carpathian Basin in 568. The name of their first ruler was Bayan. Some scholars connected this name with the Mongolian language, but now it is clear that the name is an Oguric one (see Chuvash *puyān* 'rich' (< *bayan*), all Oguzic languages have *bay* 'rich'. Between 670 and 700 a Bulgar group joined the Avars in the Carpathian Basin. We know about a dozen Avar inscriptions in a runiform script. This script pertains to the same system as the runiform inscriptions of the Treasure of Nagyszentmiklós. All

are very short, the longest is the inscription on a needle case found near Szarvas (SE Hungary). Names and titles used in the Avar Empire are noted in Latin and Greek sources.

3.8. Sources on the language of the Bulgars. In some sources Bulgars are called Onogurs and Onogurs are called Bulgars. However after the foundation of the Bulgar polity of Kuvrat in 630, the name Bulgar is used more and more. The reign of Kuvrat and his sons lasted until the seventies of the 7th century. These 40 odd years are too short that we may reasonably expect linguistic data in the sources. Some names as that of Kuvrat (see OT *kuvrat*- ‘to collect (people)’) or that of his eldest son (*Bat*) *Bayan* are surely Turkic. On the latter see above. We know that the Kutrigurs were from the same stock as the Bulgars (Nicephoros, Theophanes). We know more about the language of the Danube Bulgars and the Volga Bulgars.

3.8.1. There exist several sources on the language of **the Danube Bulgars**, called also Proto-Bulgars to distinguish them from the Bulgar Slavs. They fled from Bulgaria of Kuvrat under the leadership of one of his sons Asparuh. To the first group of written sources pertain inscriptions written with Greek letters. Most of them are in Greek language, but contain several Danube Bulgar names, titles and different objects, connected with taxes or tax exemption. We dispose also of a few inscriptions in the local language. An interesting part of these inscriptions can be seen on the famous Treasure of Nagyszentmiklós (now in Vienna, *Kunsthistorisches Museum*). The Treasure contains 16 golden vessels. There is one inscription in Greek letters and in Greek language, one inscription in Greek letters but in the local language and 14 small inscriptions in a uniform type of script. There exists an interesting List of the Rulers of the (Danube) Bulgars (*Protobulgarische Fürstenliste* according to one of its editors, O. Pritsak). This remained to us in three copies partly translated into Church Slavic and what remained from the original text is very difficult to decipher. The second group of sources are the written sources mainly in Middle Greek, but some also in Latin and later in Old Slavic. The third group is the present Slavic Bulgarian language which preserved a few words from Danube Bulgarian. The last group contains words copied into the language of the neighbours.

3.8.2. Sources of the language of **the Volga Bulgars** are early loanwords in the Permian languages (Zyryan, Permyak and Votyak) and a few words which were copied by the Tatar and Bashkir languages from the languages of the Volga Bulgars. The descendant of one dialect of the Volga Bulgars is the Chuvash language. Some of the Kazan Tatars are Kipchakized Volga Bulgars. Among the written sources important is the Travelogue of Ibn Fadlan, who was member of the embassy sent by the caliph to the Volga Bulgars in 921/922. The report contains several local names and words. Ibn Fadlan was an Arab and worked with two interpreters who joined the group. One of them was a Bulgar. A very important group of sources on the Volga Bulgar language are the Volga Bulgarian inscriptions. These Muslim tomb inscriptions are in Arabic, but many words and the dates are given in Arabic script but in the Volga Bulgarian language. The earliest is dated 1281 the last 1357, some have not complete dates.

3.9. Sources on the language of the Kazars. The Kazar kingdom was surely polyethnic and in Kazaria more languages were spoken. It is even possible that more Turkic languages were present, since the Kazars lived in the 7th century under the rule of the western branch of the East Turkic Kaganate. This lasted formally until about 670. Speakers of Oguzic languages may have appeared in Kazaria from the 8th century on. Therefore the logical question is not “What type of Turkic language was spoken by the Kazars”, but “What types of Turkic languages were spoken in Kazaria”. In the Arabic, Persian, Armenian and Georgian sources there are a few names and titles of Kazar dignities. The problem with this staff is, that titles and names are not always from the

language which the people speak. German *Kaiser* is not German (*Caesar*), Hungarian *király* ‘king’ is not Hungarian (through Slavic mediation German *Karl*). The same is with the names. The name of one of the sons of Kuvrat, who led the Danube Bulgars to the right bank of the Lower Danube was Asparuch, a name of Iranian origin. This does not mean that other titles and names could not be from the language of the people in question. However in the case of the Kazar, beside the usual problems (miswriting, misreading, transcription problems etc.) a serious factor which have to be taken into consideration is the presence of Oguzic groups among the Oguric groups in Kazaria. Finally we have to mention that though no argument has been voiced against the Turkic character of the Oguric language of Kazaria, yet the origin of the Kazars is obscure. Their name seems to be connected with the eastern transmission of the name and later title *Caesar*. Some names may pertain to a pre Kazar time before the Turkization of the Kazars. In view of these the slowly growing stock of Kazar words have to be revisited. A special case is the name of the fortress Sarkel or Šarkel, called in Greek “the White House”. There does exist a small library of papers written on the name. The second part, which is of Iranian origin and means ‘house’ (Greek *οίκημα*). This word is present as such only in Chuvash, where *kil* ‘house’ is the regular correspondence. (The quasi suffix *-gil* in Azeri and Turkish may have the same origin but different function). The first word may be read either with *s-* or with *š-*, the second was not noted in Greek, and noted with diacritics in Hebrew, which may have been omitted. Allegedly there exist a Hebrew inscription on the Crimea where there is written a Hebrew dotted Shin, but this data has to be verified. Whether the initial has to be read with *S-* or *Š-* may be left open. But Chuvash is the only language where *šură* < *sāri(γ)* has the meaning ‘white’ and not ‘yellow’. A geographic name of interest is the name of a river on the Crimean peninsula. In the glosses of an anonymous scholiast added to the *Notitiae Episcopatum* composed between 733-746, we find the river name on the Crimean peninsula: *χαρά σιου* (translated as *μαῦρον νεγόν* ‘black water’). Of importance is, that instead of the expected *kappa* we find *χ-* and instead of the Oguzic *sub* or *suv* we find *siu*. The *χ* in back vocalic words appears in other Turkic languages, but not so early. In Chuvash it is a rule, and ‘black’ is now *χură*. The second word is Volga Bulgarian *šiu*, and now in Chuvash *šiv*, which are perfect matches of *siu*, which may have reflected a form *šiu*. In an itinerary preserved by Constantine VII in his *De Administrando Imperii* we come through the river *χαρά κουλ*, i.e. *χara kül* or *χara köl* the Black River (“in which they fish for sturgeon”). It flows into the Meotic Bay.

We have rich sources on the Kazarian-Alanian contacts. From this point of view the name of the ‘ash tree’ is of importance. In Ossetic it sounds as *kaerz*. The reconstructed Turkic form is **kebrüč*, but the copied form had to be **kevrš* → **kers* > *kerz*. If the Turkic form would have retained its final affricate, it would be reflected in Ossetic by *-c* < **č*, that is **kerc*. See Turkic *burč* ‘pepper’ → Ossetic *byrc*, *burcae* ‘černyj perec’. As we see *kaerz* reflects a later phase, which is also present in Hungarian *kőris* (kö:riš).

There exist a group of yet undeciphered runiform inscriptions in East Europe. The territory, where they were found is identical with the Kazar and Bulgar territory. It may be, that some of them are using the script for non Turkic language or languages (e.g. Alanian), but we have to suppose that most of the rendered Turkic language or languages. The deciphering of the inscriptions is hampered by the fact, that we do not dispose of a good edition of all these, about now over 100 inscriptions. None of them is longer than 60-70 signs (Inscription of Szarvas, the inscription on a scull of a bull from Elista), most are consisting of a dozen or less signs. We do not have bilingual inscriptions or inscriptions written with the runiform alphabet and another one. The alphabet is a carved type, or remained to us only in this form. It does not have any visible trait with the East Turkic, Nordic or Etrusk carved scripts.

3.10. Loanwords in Chuvash. After the migration to the Volga-Kama region in the 8th-9th century the Volga Bulgars became a mighty people in a relatively loose dependence from the Kazars.

3.10.1. Permian loanwords in Chuvash. The Zyryan gave way to the Bulgars and moved to the north, but we find some Permian words in Chuvash. Such is the prohibitive particle *an* used before the imperative of the verbs *an kala* ‘don’t speak’. Votyak, Zyryan *en*. It is not to be connected with the Mongolian stem *e-se*. Chuvash *pileš* ‘rowan berry and tree’ goes back to a Permian **pileš*, Chuvash *pāši* ‘elk’ to Permian **pušey*.

3.10.2. Persian and Arabic loanwords. Trying to get rid of the Kazar power the Volga Bulgars began to embrace the Muslim faith and sent ambassadors to Bagdad for help. As we have seen above the embassy visited the Volga Bulgars in 921/922 and Ibn Fadlan wrote a record on it. The embassy found that there are already some traits of the Muslim faith present among the Bulgars, but the full and pure form of the faith has to be established. This is happening with the help of the embassy and later by visiting and adjoining imams. Most of these persons came from Khwarezm. This means that from the beginning of the 10th century on, we can observe the influx of Persian and Arabic words originally mainly connected with Islam. These words give us help for the reconstruction of the history of Middle Chuvash. The word for ‘man, person, human being’ is in Chuvash *šin*. This word goes back to Persian *jān* ‘soul, vital spirit’ and not to Mongolian *jon* ‘people’ (→ Siberian Turkic languages). This and a very few other words entered the ancestor of the Chuvash language in a very early phase. The same word was copied a second time with many other Persian and Arabic words as *čun* ‘soul, spirit’. This happened in a time, when PT *č*- already changed to *ś*- thus the affricate did not join other words which began with *č*-□. The first was borrowed in a time before the PT illabial /a/ changed to /i/ or became a labial /â/. The word for ‘week’ and earlier also for ‘Friday’ is in Chuvash *erne*. This is already on several Volga Bulgarian inscriptions present as *erne*. The word goes back to Persian *adina* ‘Friday’. This word was earlier copied then the *d > δ > z > r* change was closed. The second or third step can be seen in Hungarian *búza* ‘wheat’ ← T **būza* < **buza* < *bugda(i)* it began before the end of the 9th century. It was closed in the 12th century. From the year 1230 we find in the Russian sources already *-r-* as in **turun* (*trnove*) a title which corresponds to Turkic *tudu’n*.

3.10.3. Early Russian loanwords. As we have seen above, the eastward migration of the Slavs reached the Oka river around 900. It continued in the next centuries. The Slave merchants reached the Volga and established there settlements for trade, which later became Russian cities. The earliest contacts between Slavs and the Volga Bulgars were trade connections. But soon the contacts became closer and more and more Russian words were used in the Volga region. One of these words is Chuvash *kānčālā*, *kēnčēlē* ‘a bunch of flax for spinning, also spindle’. The Russian word is now *kužel’*, but it had a nasal and sounded in Proto Slavic *kǫžьlb* (= *konžālā*). It was copied into Finnish as *kuontalo* and is present in all languages of the Volga region. The nasal disappeared from East Slavic before the end of the 10th century. Another early loanword is Chuvash *šap* ‘flail’ < **čäp* ← Russian *cep* → Hungarian *csép*. Here, however the word came under the influence of the Chuvash word *šap-* ‘to strike, beat’. Old Russian /ts/ is sometimes substituted by Chuvash /ś/ as in Russian *nemec* → Chuvash *niměś*, but there is no trace of Old Russian *č* → Chuvash *č* > **š* > *ś*. There exist some dialectal representations which seem to be due to inner Chuvash processes as Russian *kalač* ‘a kind of white loaf of bread’ → Chuvash *kulača*, dialectal *kulaś*. That means that the change had to be closed before the bulk of the Russian loanwords were copied into Chuvash. While the existence of the voiced affricate /dž/ <j>

is clearly reflected by the Volga Bulgarian inscriptions, there is no direct trace of the unvoiced /č/. Nevertheless we have to reckon with different Chuvash-Bulgar dialects.

3.10.4. Tatar loanwords. Speaker of different Kipchak languages came to the Volga-Kama region after the Mongolian invasion which began in 1235. Now they pertain to the Kazan Tatar and the Bashkir languages respectively. For a short time the Kipchak lived together with some Chuvash speaking Volga Bulgar groups. Other Chuvash groups settled in the forest region of the Volga bend. The Kipchak population used the Kwarezmian Turkic as their literary language, which came soon under heavy local influence. This hybrid Khwarezmian Turkic and Kipchak language can be observed in one part of the Volga Bulgarian inscriptions. These inscriptions are sometimes called Inscriptions Pertaining to the First Style. The Second Style contains the Chuvash traits. There are also mixed inscriptions showing some traits of both the Bulgarian and the Kipchak-Khwarezmian language. Most interesting is a tomb stone, where the epitaph of the deceased lady is written on the one side in Style One and on the other side in Style Two.

We can now separate at least three different chronological layers of the Tatar loanwords. To the earliest layer pertain such Chuvash words as *χurāntaš* ‘brethren, relatives “those from the same womb”’ ← *karindaš*, cf. Chuvash *χiram* ‘belly, womb’ < *qarın*. To the second layer pertains e.g. Chuvash *χapχa* ‘gate, door’ ← Tat *kapka* and to the third e.g. *karāntāk* ‘peritoneum’ ← Tat *karindik*.

3.10.5. Mongolian loanwords through Tatar mediation. In 1235/6 winter the Mongolian army crushed the army of the Volga Bulgars and occupied the territory. They settled mainly Kipchak Turkic troops to rule the regions, who later became members of the Golden Horde and afterwards formed the Khanate of Kazan. The Russians ended the rule of the Kazan khans in 1551. Mongolian words were spread over the whole Mongolian Empire of the Chingisids. Most of the Mongolian words came into Chuvash *via* Tatar. They are important because they help to give a chronology of some of the phonological changes in Chuvash. Mongolian *karčiga* ‘falcon’ → Tat *karčiga* → Chuv **χarčiga* > *χurčka*. Middle Mongolian *malaqai* ‘cap’ → Tatar *malaxai* → Chuvash **malaxai* > *mulaχχay*. Even more interesting is the fate of Mongolian *ja’uči* ‘matchmaker, go-between’. The earlier form is preserved only as a Chuvash loanword in Cheremis *šaus*, *šavus* etc. ‘wedding groomsman, bridesman’ this reflects a Chuvash **šavus* which disappeared from Chuvash. It copied later the same word as *yevčě* which is a copy of Tatar *yäuči* itself a copy of the Mongolian word.

3.10.6. Cheremis loanwords in Chuvash. While the Volga-Bulgars had early contacts with the Permian languages (Zyryan, Permyak and Votyak), after one part of the speakers of the Chuvash language settled down in the forest region of the Volga bend they became into contact with several other languages. Among them is the Cheremis or Mari. This language is full of Chuvash loanwords, but Chuvash also copied some words from Cheremis. Among them is Chuvash *yaxtā* ‘pine tree’ ← Cheremis *yakte*, see Chuvash *χirā* id. ← Samoyed (above) or Chuvash *karāš* ‘corncrake, landrail’ ← Cheremis *karš* or Chuvash *vāy* ‘strength’ ← Cheremis *vij*.

3.10.7. Loanwords from an unknown substrate language. The present land of the Chuvash was surely not empty when they settled there. The geographical names of the region points partly to speakers of the Permian group. The element *-kar* in *Šupaškar* ‘Cheboksary’ is surely Permian *kar* ‘city’. Cf. Zyryan *kar*, Votyak *kar*. But there exist a group of words which are present in both Cheremis and Chuvash, sometimes in other Volga languages, and have neither Turkic, nor Iranian or Finno-Ugric etymology. The Chuvash name of the ‘beetle’ is *nārā*. The word occurs also in Cheremis as *nor*, *nār* in the same meaning, and it is present also in Bashkir as *nor*.

However the word cannot be connected with any of the known language groups. The study of these type of words promises new insights into the history of Chuvash.

3.11. The Volga Bulgar and Chuvash loanwords in other languages. After arriving to the Volga-Kama region the Volga Bulgars came into contact first with the Permian languages (Zyryan, Permyak and Votyak). The Cheremis-Chuvash contacts began later, in the 13th century. Though there are only very few Bulgar loanwords in Mordvin they have to be very early because of geographical-historical reasons.

3.11.1. Bulgarian loanwords in Mordvinian. Among the few Mordvinian words of Bulgarian origin we may mention Moksha Mordvin *ayār* ‘cool (of weather)’ ← Bulgar *ayar*, see Turkic *ayaz* id. or Moksha Mordva *šerä* ‘brass’ ← Bulgar **še,r* cf. Turkic *yez*.

3.11.2. Volga Bulgarian loanwords in the Permian languages. The Zyryans moved to the East to give way to the Bulgars in the 10th-11th centuries. After a short time the connections among Zyryan on one side and Permyak and Votyak on the other, ceased to exist. That means that all words which are common to Permian have to be earlier than the 11th century. Such are e.g. Zyryan *kiš*, Votyak *kiš* ‘Weberkamm < Proto Permi *kiš* ← Bulg *χiš* (> Chuvash *χěš*) < **kilč*. See Turkic *kilič*, *kilč*. Interesting is the following word: Zyryan *kuže* ‘Waldgeist’, Votyak *kužo* ‘Wirt, Hausherr’ < Proto Permian **kuja* ← Bulgar *χuja* (> Chuvash *χuša*) ← Persian *χwaja*. For the chronology of the change of *č* >> *š* interesting is the Zyryan word *čarla* ‘Sichel’ which reflects a PP **čarla* ← Bulgarian *čarla*, while Votyak *šurlo* is a copy of Bulgar **šurla* (> Chuvash *šurla*). See Hungarian *sarló* (šarlō).

3.11.3. Volga Bulgarian loanwords in Russian. The Volga Bulgarian empire was in the 11th-12th century the main power north to the Kazars and East to the Rus. Many of the early Turkic elements in Russian pertain to this period. We mention only two words which are playing an important role in present Russian. The word *tovar* ‘goods, wares, article, commodities etc.’ comes from Volga Bulgarian **tavar*. Though it is absent from present Chuvash it is also present in Hungarian in the form *tár* ‘container’, also in such compositions as e.g. *kinčstár* ‘treasure house’. From this word was formed (but the derivation is not entirely clear) the word *tovarišč* ‘comrade’ which, or the original form of which is present also in Hungarian as *társ* ‘companion, fellow’. The other Russian word is *hozjain* more precisely its stem *hozja* which existed independently in Old Russian. Persian *χwaja* was mediated by Volga Bulgarian (cf. Chuvash *χuša*). According to some sources the famous district of Moscow, the Arbat was originally the place of the Volga Bulgarian traders (later that of the Tatars). The name should go back to the Volga Bulgarian form (**arbat*) of the ultimately Arabic word *rabat* ‘suburb of a city, quarter’ (see the name of the capital of Marocco: Rabat). Less probable is the Greek etymon which connects the name with *ropation*, itself also of Arabic origin.

3.11.4. Volga Bulgarian loanwords in Tatar and Bashkir. The bulk of the Kipchak population appeared in the Volga-Kama region with and after the Mongols. However we cannot exclude that some groups could have already joined the Bulgars in a somewhat earlier time. In any case the present Tatar and Bashkir language has a small group of words which point to a Volga Bulgar origin. I would mention only a few. Tatar and Bashkir *izge* has the meaning ‘saint, holy’. This word is also present in Kazak as *ezgi* ‘good, sacred’ and in some other Turkic languages. These forms come from Volga Bulgarian **ezgi* < *ezgü*. The word is a Volga Bulgarian form of OT *edgü* ‘good’. The word was copied into Hungarian where it has the meaning ‘saint’. The other word is Tatar *kazak* ‘das Pfund’. This is an old metric unit used for weight. It is the same word as *kadak* ‘nail’, see Radlov *kadak* ‘der Nagel, das Pfund (weil die Pfunde auf den hölzernen Staabwagen

Bessmen) mit Nägeln bezeichnet sind'. The *-d-* is represented by *-z-* in the same way as in *izge* etc. above. Tatar dialectal *šoklan-* 'to alienate' comes from Volga Bulgarian **šoklan-* see Chuvash *šuk* 'it is not' < PT *yok*.

3.11.5. Chuvash loanwords in Cheremis. Cheremis has more hundreds of loanwords copied from Chuvash. This material is a great help for the reconstruction of the history of Chuvash. The word for 'pig' is now in Chuvash *sīsna*. The fact that this is a late form due to assimilation from **sīsna* we learn only from Cheremis, where in some dialects the initial *ś-* is preserved. The initial goes back to *ǰ-*, as we see in Hungarian *disznó* < **ǰisnau*. Sometimes Cheremis preserved a form which had to be present in Chuvash but is not. The name of the 'hemp breaker, swingle' is in Chuvash *tilă*. The form **tulă* has been preserved only in Cheremis *tule* (cf. Hungarian *tiló*). In case of the old Turkic *-ŋ-* Chuvash has *-m-*. This is seen in Cheremis *tüme* 'Hügel' ← Chuvash *těme* < **tūjek*. Some words disappeared from present Chuvash and Cheremis preserved them. Such a word is Cheremis *čij* 'schell, stark, häftig' which presupposes a Chuvash **čij*. The word is present in Kazak *tij* 'strong, brave' and occurs in some Siberian languages. See Mongolian *čing* 'firm(ly)', *čingga* 'strong'. It may be the same as the prefix expressing 'very' in Chuvash *čăn*, e.g. in *čăn layăχ* 'very good'. An other word preserved only in Cheremis is *yiră* 'Buttermilch'. This reflects a Chuvash form **yiră* < **irag* < *ayrag* (not **yadağ*) and is present in Hungarian as *író*.

3.11.6. Multiple reborrowings frequently occur. Chuvash *šărpă* 'splitter' is a copy of the Tatar word *širpî* itself a loan from Volga Bulgarian *širpî* which goes back to PT **čirpî*. Cheremis *šărpă* may be either from Tatar or from Chuvash.

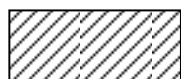
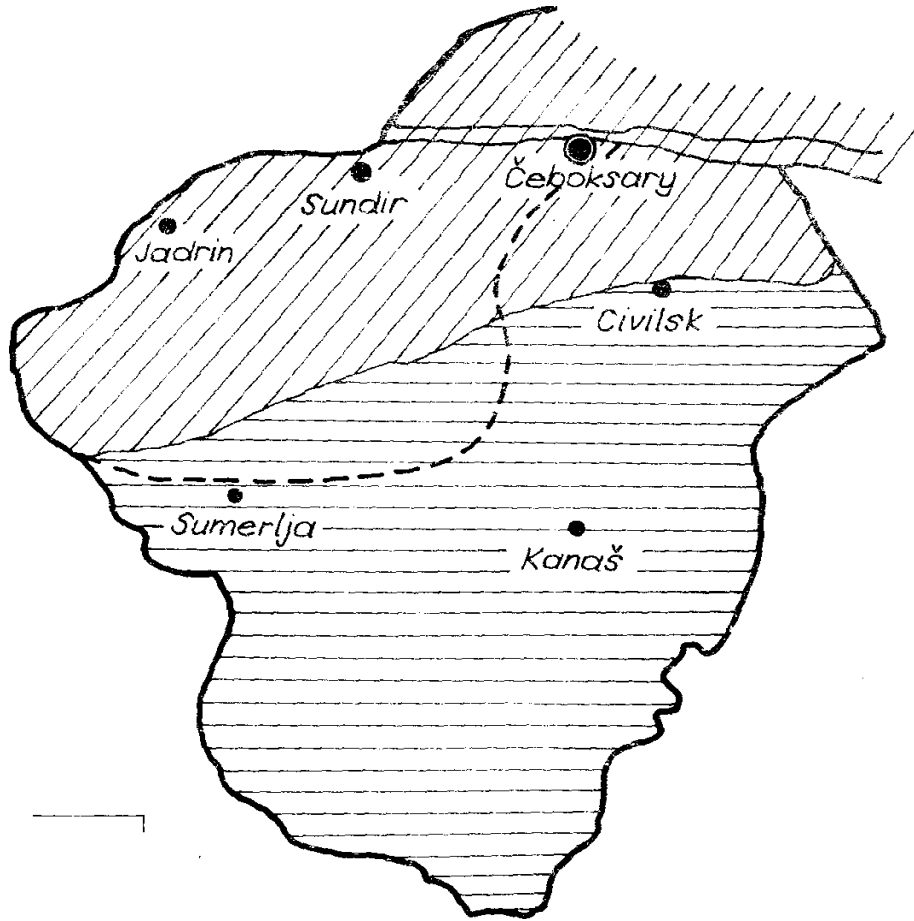
4. The present Chuvash language

4.1. The dialects. Chuvash has two main dialects. The northern dialect is called *Viryal* from *vir* 'upper' (PT *ör*) and *yal* 'village, place of living' (PT *ēl*). It is also called *Turi* from *tu* 'mountain' and the suffix *+ri*. The southern dialect is called *Anatri* from *anat* 'lower part' < *an-* 'to descend' (< *än-*). *Viryal* is the more archaic dialect. It preserved the labial, non labial opposition in case of the reduced vowels, thus it has four reduced vowels as */ă/, /ě/, /ǒ/, /Ǔ/*. This opposition is very important for the etymology of the Chuvash words, because the labial reduced vowels preserved the +labial feature of the original. The two labial reduced lost their +labial feature in the *Anatri* dialect and the literary language. A very clear difference between the *Viryal* and *Anatri* dialects is the representation of the full labial. There are two non reduced labials in Chuvash, one of them appears as [o] in *Viryal* and as [u] in *Anatri*. The other is /ü/. There are a few words which have *u* also in *Viryal* as *ut* 'horse', *tutlă* 'sweet'. Until now no reasonable cause of this feature has been suggested. In some subdialects of *Viryal* the diphthongs are represented by triphthongs, e.g. *-ăva-* is appearing as *-oa-* e.g. *tăvat* 'four' *Viryal toat* (< *tört*). Important is, that the plural *-sem/sen-* is not obeying the vowel harmony in *Anatri*, but in *Viryal* subdialects we find already *-sAm/sAn-*. The isoglosses of the different dialectal features are not always coinciding. In the contact zone of the two dialects a mixed situation can be observed. Some consider this zone to be a different dialect, but this is hardly acceptable. Below we see the main isoglosses for *Viryal* and *Anatri*:

Important is that the literary language is based on *Anatri* (see below), but the capital *Cheboksary* is within the territory of the *Viryal* dialect. The education is in the literary language and the educated people in *Cheboksary* speak the literary language, in fact a normalized variant of the *Anatri* dialect. The newspapers, the radio and the television uses the literary language, and it is therefore now very difficult, and becomes more and more difficult to find native speakers

who speak the old Viryal dialect without the impact of the literary language. Therefore all data on the Viryal dialect are of high value.

The two main features of the Chuvash dialects:



/o/



/u/



isogloss of the plural

The Chuvash dialects

4.2. The formation of the literary language. The earliest monument of the Chuvash language can be found in the work of Philip Johan Strahlenberg collected in 1723 and published in 1730. Some small wordlists were published by G.F. Miller (collected in 1733, published in 1758), J. E. Fischer (collected in 1747, remained for a long time in manuscript and seems to contain the same Chuvash material of Miller, Fischer is now published by Gulya in 1995), by Pallas (collected in 1768/69, published in 1771) and others. All these lists were written with the Latin alphabet. In the second half of the 18th century the activity of the Orthodox Christian missionaries grew. To help the translation work of Christian texts into Chuvash, grammatical sketches were worked out. These were written in Russian and the Chuvash material was also written by the Cyrillic script. The priests of Chuvash origin began to write in their own language using the Cyrillic script. From the end of the 18th century on we have several poems written for festive occasions in Chuvash. Parallel to the growing activity in the episcopal seminars of Kazan and Nizhnij Novgorod Catharine II ordered to collect all languages spoken in her empire. Among them also Chuvash has been collected and many local learned people began to collect material and send it to Pallas, who was in charge of the collection. We know about a Chuvash wordlist of about 3.000 words most probably from 1785. An archbishop, by name Damaskin, compiled a Russian-Tatar-Chuvash-Mordvin dictionary in 1785, which remained in manuscript, but was many times cited. In 1863 was published the famous *Načertanije pravil čuvaškogo jazyka i slovar´* a grammar which was compiled under the direction of V. P. Višnevskij for the theological schools of the Kazan bishopry. This and the first Chuvash translation of the Bible were used by W. Schott for his *De lingua Tschuwaschorum* published in Latin in Berlin, 1841.

In the next phase the activity of I. Ja. Jakovlev was fundamental. He wrote many articles on the Chuvash language and in 1871 he reformed the Cyrillic orthography used for Chuvash. His alphabet was based on phonological principles. With this, without using the modern terminology he was the first who applied the phonological principle to a Turkic language. The present Chuvash alphabet and orthography got its final shape in 1938. Jakovlev worked in the College of Simbirsk, and the Simbirsk school, which trained the teachers for the Chuvash schools, played a decisive role in shaping the Chuvash literary language. Jakovlev, a great scholar who never turned to Marxism, died in 1930. His close personal friendship with the father of Lenin saved him and his scholarly work. The great Chuvash poet Konstantin Vasilev *Ivanov*, known also as *Kestenttin* (1890-1915) wrote his epic poem *Narspi* in the Anatri dialect and with the orthography of Jakovlev. Most of the writers following him were former pupils of the Simbirsk school, but even those, who were not, followed the same way. With this the literary Chuvash language got its shape which was, of course, influenced by later activities, but its basic features were ripe before the Bolshevik revolution in 1917.

4.3. Areal features in Chuvash. The languages of the Volga-Kama region, Chuvash, Tatar, Bashkir, Cheremis and Votyak form a language area with special traits common to the languages spoken there. In some cases it is easy to show the direction of the changes, in other it remained and perhaps remains unclear. One of the most common feature is the opposition of the quantity of the vowels. The basic opposition of reduced and fully formed vowels appears in all languages of the area. Another feature of the vowels system is the presence of the labial /â/. This is the sound which changed later in Chuvash to /o/ and in the Anatri dialect to /u/. They are opposed to the non labial /a/ which changed from open /ä/. These two features effected the word stress and the vowel harmony. There are important morphological and phraseological similarities among the Volga-Kama languages. Their study presupposes deep knowledge of the Finno-Ugric and Turkic languages of the area.

4.4. Graphemes, orthography and graphotactics. The present Chuvash language is written with the Cyrillic script adapted to the needs of the Chuvash language. The basic principle is phonologic and not phonetic. A word written as /ака/ ‘plough’ is automatically read as [aGa], since phonetically all stops and affricates are weakened in intervocalic position. The sound [G] is an allophone of the phoneme /k/. All stops and affricates are palatalised before and after the front vowels. This is not marked by the orthography, e.g. /теце/ ‘saying (as saying something)’ is pronounced [tʰezʰe]. Consonants are weakened not only in intervocalic position but before and after sonants, or in sandhi position. Thus *шурамныç* ‘dawn’ is pronounced as *šuramBuś*, *нăрçа* ‘flea’ sounds as [părzá], and *хара кыç* ‘black eye(s)’ sounds [χaraGuś]. In case when for some reasons the strong quality of the consonant should be kept, the consonant letter is doubled. ‘Two’ is *икĕ* and pronounced as [íGĕ], but the emphatic version is written as *иккĕ* and pronounced as [íkĕ].

The marker of the 3rdP singular form of the indicative verb is *tʰ*. Since all consonants are palatalised in front vocalic words, the palatalised quality of the *-t* is not marked in front vocalic words. Thus *килет* ‘he/she comes’ is pronounced [kʰilʰet]. In the back vocalic words the palatalization of *-t* has to be marked, because it is not automatic. The word *калать* ‘he/she says’ is pronounced as [kalatʰ] therefore is written with *-мь*, though the palatalization is the same as in the front vocalic words.

5. Summary

Chuvash is one of the Turkic languages. It has to be studied on its own right, but also for the sake of the reconstruction of Ancient Turkic. However to find those traits which are in one or the other form preserved from Ancient Turkic, can be separated only if we study the complex history of Chuvash. The history of Chuvash gives us help to reconstruct not only Ancient Turkic, but also the contacts of the Oguric languages with other Turkic and non Turkic languages. The case is somewhat similar to the problem of “Altaic”. We can reconstruct “Altaic” only after we dispose of a thorough study on the history of the Turkic languages and reconstructed as much as possible of Proto-Turkic.